UGR

2011

EVALUACION DE CONOCIMIENTOS Y RECURSOS DIDÁCTICOS EN LA FORMACIÓN DE PROFESORES SOBRE PROBABILIDAD CONDICIONAL

TESIS DOCTORAL

DEPARTAMENTO DE DIDÁCTICA DE LA MATEMÁTICA

JOSÉ MIGUEL CONTRERAS GARCÍA
UNIVERSIDAD DE GRANADA
Directoras: Carmen Batanero Bernabéu
Carmen Díaz Batanero

Editor: Editorial de la Universidad de Granada
Autor: José Miguel Contreras García
D.L.: GR 2388-2011
ISBN: 978-84-694-3603-5

UNIVERSIDAD DE GRANADA

Departamento de Didáctica de la Matemática

EVALUACION DE CONOCIMIENTOS Y RECURSOS DIDÁCTICOS EN LA FORMACIÓN DE PROFESORES SOBRE PROBABILIDAD CONDICIONAL

TESIS DOCTORAL

José Miguel Contreras García

Directoras:

Dra. Carmen Batanero Bernabeu

Dra. Carmen Díaz Batanero

2011

EVALUACION DE CONOCIMIENTOS Y RECURSOS DIDÁCTICOS EN LA FORMACIÓN DE PROFESORES SOBRE PROBABILIDAD CONDICIONAL

TESIS DOCTORAL

MEMORIA realizada bajo la dirección de las
Doctoras Carmen Batanero Bernabeu y Carmen Díaz Batanero
que presenta D. José Miguel Contreras García
para optar al grado de Doctor

Fdo: José Miguel Contreras García

V^{o} B^{o} V^{o} B^{o}

Dra. Carmen Batanero Bernabeu Dra. Carmen Díaz Batanero

Trabajo realizado en el marco de la beca FPI BES-2008-003573, Proyectos SEJ2007-60110/EDUC, MEC-FEDER y EDU2010-14947 (MICINN-FEDER) y Grupo PAI FQM126 (Junta de Andalucía).

Agradecimientos

En primer lugar quiero agradecer a mi directora de tesis, la Dra. Carmen Batanero, por la oportunidad que me brindó en su día de trabajar con ella. Su sabiduría, trabajo y, sobre todo, paciencia ha marcado este trabajo de investigación y al que lo firma. Gracias Carmen eres un ejemplo a seguir.

También a mi otra directora, Carmen Díaz, por toda su ayuda.

Quiero agradecer especialmente a tres profesores del Dpto. de Didáctica de la Matemática. A Juan D. Godino, por su simpatía y su apoyo; a José María Cardeñoso por su amistad y consejos y sobre todo a Rafael Roa por sus consejos, buenos ratos y sobre todo por cederme su espacio y tratarme como uno más.

Quiero acordarme de Angustias Vallecillos que siempre me aconsejó muy bien, y me trató mejor.

Agradecer a Pedro Arteaga y a Gustavo Cañadas su compañerismo y sus consejos. A los profesores que me ayudaron con los cuestionarios.

También quiero agradecer a Antonio Arcos, María del Mar Rueda y Andrés González su empujón en mis comienzos.

Agradecer a mis padres por permitirme estudiar en los malos momentos y por todo lo que me han inculcado. A mis tres hermanos todo lo que han hecho por mí y sobre todo acordarme de los que ya no están y tanto les debo: gracias tito Pepe, gracias tita Manola y gracias abuela.

A mis amigos que tanto me han aguantado, gracias Linares, Paco, Pepe y Wes.

Y sobre todo agradecer a Elena ya que todo lo que he conseguido y conseguiré de aquí en adelante te lo debo a ti.

INTRODUCCIÓN

Aunque en los últimos años se ha producido una gran expansión en la investigación sobre formación y desarrollo profesional del profesor de matemáticas, este esfuerzo apenas ha tenido en cuenta el caso particular de la probabilidad y la estadística. Como consecuencia la enseñanza de la estadística y probabilidad en la Educación Primaria ha sido casi inexistente hasta muy recientemente, y en la Educación Secundaria recibía un peso mucho menor que el que se trata de dar en las nuevas directrices curriculares.

Esta situación está cambiando pues, tanto en los Decretos de Enseñanzas Mínimas en España, como en las orientaciones curriculares de muchos otros países, observamos un cambio en la tendencia de la enseñanza de la estadística, que se inicia desde la Educación Primaria y se intensifica en todos los ciclos. Ello muestra el interés de investigar sobre los conocimientos y creencias de los profesores que son responsables de esta enseñanza.

Entre todos los posibles temas que se podrían haber elegido para este estudio nos hemos centrado en la probabilidad condicional por varios motivos. La probabilidad condicional fue incluida en la lista elaborada por Heitele (1975) de las diez ideas fundamentales sobre las que se asienta el desarrollo histórico del cálculo de probabilidades. Dichas ideas son generalmente contra intuitivas y según este autor han de ser educadas. Además, es posible enseñarlas con distintos grados de formalización desde la escuela primaria a la universidad. La probabilidad condicional es también base de muchos otros temas en probabilidad y estadística, por lo que su comprensión facilitará la de otros conceptos posteriores.

La probabilidad condicional se incluye explícitamente como tema a tratar en los Decretos de Enseñanzas Mínimas de Educación Secundaria y Bachillerato y es parte de las asignaturas de estadística en todas las carreras universitarias. Aunque no aparezca explícitamente en los Decretos de Educación Primaria, está implícita en la lectura e interpretación de tablas de doble entrada, que si se incluyen en los mismos. Por otro lado, su comprensión y la resolución de problemas elementales relacionados con la probabilidad condicional se requiere en la toma de decisiones en la vida cotidiana y en la interpretación de información que aparece en los medios de comunicación, y otras materias del currículo, en especial en muchos otros temas estadísticos incluidos en Bachillerato. Es también importante en el diagnóstico y evaluación, que forma parte de las tareas profesionales del profesor.

Todas estas razones nos han llevado a interesarnos por la formación de los profesores y a llevar a cabo una serie de estudios exploratorios relacionados entre sí, y con una temática común. Más concretamente, abordamos una serie de estudios relacionados con la evaluación de las necesidades formativas de los profesores de los niveles no universitarios en relación a la probabilidad condicional y los objetos matemáticos que con ella se relacionan, tales como la probabilidad compuesta, el teorema de Bayes, independencia o dependencia. Cada estudio tiene sus propios objetivos, hipótesis y metodología, aunque todos comparten el mismo marco teórico. La Memoria se organiza en los siguientes apartados:

En el primer capítulo, una vez introducido el problema de investigación y resaltado su interés, se describe el marco teórico en que se apoya este trabajo, que es el Enfoque Ontosemiótico propuesto para la didáctica de la matemática por Godino y colaboradores. Seguidamente se realiza un análisis de los objetos matemáticos relacionados con la probabilidad condicional y del contenido que sobre este tema aparece en las orientaciones curriculares para la enseñanza no universitaria. El capítulo finaliza con la presentación de los objetivos de esta investigación y la descripción de los diferentes estudios que la componen y un resumen de sus características metodológicas.

En el segundo capítulo presentamos los antecedentes que constan de dos partes principales. La primera sintetiza una breve perspectiva de la investigación relacionada con la formación de profesores de matemáticas para enseñar probabilidad. Se comienza con una breve reseña de algunos trabajos que definen categorías del conocimiento del profesor de matemáticas, para pasar a analizar los modelos específicos planteados en el campo de la educación estadística y que podrían ser de utilidad en nuestro trabajo. La segunda parte del estado de la cuestión sintetiza los trabajos relacionados con la comprensión o la enseñanza de la probabilidad condicional.

En el capítulo 3 se presenta un estudio de evaluación de algunos conocimientos matemáticos y didácticos de una muestra de 183 futuros profesores de educación primaria. En el Estudio 1, algunos aspectos del conocimiento común sobre el contenido matemático se evalúa a partir de las respuestas a una tarea abierta, en la que los estudiantes han de calcular la probabilidad simple, compuesta y condicional a partir de la lectura de una tabla de contingencia. El análisis semiótico detallado de las respuestas permite clasificar las estrategias y describir los conflictos semióticos de los participantes. Una segunda parte de la tarea, donde los estudiantes han de identificar los objetos matemáticos utilizados en la resolución de la primera parte, sirve para describir una parte del conocimiento especializado del contenido matemático. Finalmente una pregunta abierta sobre sugerencia de variables de tarea que podrían cambiarse en el problema permite una aproximación a su conocimiento del contenido y la enseñanza.

En el capítulo 4 se describe un estudio de evaluación de los conocimientos matemáticos en una muestra de futuros profesores de secundaria constituida por 95 estudiantes de los últimos cursos de la Licenciatura de Matemáticas y 101 estudiantes del máster de secundaria (especialidad matemáticas). Las respuestas al cuestionario RPC elaborado por Díaz (2007) permiten evaluar simultáneamente la presencia de los principales sesgos sobre este razonamiento que se describieron en el estudio de los antecedentes (capítulo 2) y su conocimiento matemático formal. En el Estudio 2 se realiza un estudio detallado de las respuestas de estos estudiantes a cada ítem del cuestionario, comparando los dos grupos que componen la muestra entre sí, y comparando la muestra global con los resultados de Díaz (2007) en estudiantes de psicología.

Los Estudios 1 y 2 descritos anteriormente muestran las necesitadas formativas de los futuros profesores, tanto en los aspectos matemáticos, como didácticos relacionados con la probabilidad condicional. Basándonos en los resultados de dichos estudios, en el capítulo 5 presentamos dos estudios de recursos que podrían utilizarse para llevar a cabo esta formación. En el Estudio 3 se clasifican y analizan recursos en Internet, mostrando posibles soluciones a las tareas relacionadas con ellos, dificultades previsibles en el trabajo con los recursos, objetos y procesos matemáticos trabajados e idoneidad didáctica de los recursos. Un estudio similar se lleva a cabo en el Estudio 4 respecto a algunas paradojas clásicas de la teoría de la probabilidad relacionadas con las ideas de probabilidad condicional o independencia.

En el Estudio 5 se presentan los resultados de un breve taller formativo cuyo diseño está basado en una de las paradojas descritas en el Estudio 4 y que tiene en cuenta los sesgos en el razonamiento de los profesores evaluados en los Estudios 1 y 2, descritos también en los antecedentes del trabajo. El taller se experimentó con profesores en formación y en ejercicio en varios congresos y jornadas dirigidas al profesorado. El análisis semiótico de los informes escritos sobre las estrategias iniciales y finales en un juego permite evaluar su conocimiento común del contenido y su evolución como consecuencia del taller. El estudio de las respuestas proporcionadas en un análisis didáctico posterior de la actividad por parte de los profesores permite evaluar diferentes componentes de su conocimiento didáctico.

La serie de estudios descritos proporciona información detallada sobre algunos conocimientos matemáticos y didácticos de los futuros profesores y profesores en ejercicio en relación a las tareas propuestas. Aunque el campo de la probabilidad condicional es muy amplio y este conocimiento puede ser diferente en tareas más complejas, nuestros estudios aportan una información inicial en un terreno en que la investigación previa es casi inexistente. Todos estos resultados se han recogido en una serie de publicaciones que se describen a lo largo de la memoria.

CAPÍTULO 1

EL PROBLEMA DE INVESTIGACIÓN

1.1. INTRODUCCIÓN

Este capítulo se dedica a presentar el problema abordado en esta investigación, que es el de la formación de profesores con relación a la enseñanza de la probabilidad condicional. Comenzamos el capítulo analizando la importancia de la probabilidad condicional, tanto desde el punto de vista de su influencia en el desarrollo del Cálculo de Probabilidades y Estadística, como de la formación del alumno para la toma de decisiones en su vida cotidiana.

Seguidamente se presentan algunos elementos del Enfoque Ontosemiótico para la didáctica de la matemática, que es el marco teórico utilizado para fundamentar la investigación: en concreto, utilizaremos en este trabajo las ideas de práctica matemática y objetos ligados a las prácticas, función semiótica e idoneidad didáctica. Las ideas de prácticas y objetos servirán para analizar los objetos matemáticos relacionados con la probabilidad condicional, con el fin de describir un significado de referencia de dicho concepto en nuestro trabajo. A partir de la función semiótica determinaremos algunos conflictos semióticos relacionados con la probabilidad condicional. Usaremos el concepto “idoneidad didáctica” para valorar los recursos educativos útiles en la formación de profesores sobre este tema.

Posteriormente describiremos los contenidos relacionados con la probabilidad condicional en el currículo español de Educación Secundaria Obligatoria y Bachillerato, junto con un análisis de las recomendaciones de algunos organismos internacionales sobre la enseñanza de la probabilidad condicional.

El capítulo se finaliza exponiendo los objetivos del trabajo y haciendo un breve resumen de la organización del mismo y de las características metodológicas de los diversos estudios que lo componen.

1.2. IMPORTANCIA DE LA PROBABILIDAD CONDICIONAL

Uno de los conceptos más fundamentales en la comprensión de la inferencia, tanto en su perspectiva clásica, como bayesiana es el de probabilidad condicional, que permite incorporar cambios en nuestro grado de creencia sobre los sucesos aleatorios a medida que obtenemos nueva información (Heitele, 1975). La razón es que una probabilidad no siempre se asigna en ausencia de información, sino que a menudo tenemos conocimientos sobre los sucesos implicados, que podemos utilizar para mejorar nuestra asignación inicial de probabilidades.

Así, cuando un médico establece el diagnóstico de un enfermo, no sólo valora las probabilidades de incidencia de diversas posibles enfermedades que puedan corresponder a la misma patología en la población general, sino que también tendrá en cuenta la edad y género del paciente y su historial de enfermedades previas. Ejemplos similares aparecen en muchos contextos cotidianos, como el pronóstico del tiempo, resultados de partidos o votaciones, calificaciones en exámenes o veredictos en juicios.

Díaz (2004) indica que la probabilidad condicional está en la base de muchos conceptos inferenciales, incluyendo los que tienen mayor dificultad de comprensión por parte de los investigadores:

- En relación con el test de hipótesis, el concepto de probabilidad condicional interviene en la definición del p-valor, nivel de significación, distribución muestral del estadístico, región de aceptación y crítica, así como en la aplicación de la lógica del contraste, en la definición de los errores tipo I y II, y de la potencia.
- En el método fiducial lo usamos para definir la función de verosimilitud.

- En la inferencia Bayesiana, tanto las probabilidades a posteriori como las verosimilitudes se definen en término de la probabilidad condicional. Es también el núcleo del teorema de Bayes y del concepto de región de credibilidad.
- También en la metodología de intervalos de confianza frecuenciales, tanto los conceptos de nivel de confianza, como la construcción del intervalo se basan en la probabilidad condicional.

Muchas interpretaciones erróneas descritas en estos conceptos se deben, bien al intercambio de términos en la probabilidad condicional (Falk, 1986), bien a la confusión entre probabilidad condicional y simple o conjunta (Pollatsek, Well, Konold y Hardiman, 1987) o a no comprender un razonamiento basado en una lógica condicional. Pensamos, por tanto, que sería preciso ayudar a profesores y estudiantes a mejorar su comprensión de este concepto.

Además, la probabilidad condicional interviene también en el estudio de la correlación y regresión, que se definen basándose en este concepto y que se extienden posteriormente a muchos procedimientos, tales como los modelos lineales, el análisis factorial o discriminante y la asociación de variables cualitativas. Creemos que no es exagerado decir que la probabilidad condicional es posiblemente uno de los temas estadísticos más relevantes. Por otro lado, el lenguaje y notación usadas durante la enseñanza no es siempre el más sencillo, como indica Feller (1973, p. 127): *"La noción de probabilidad condicional es un instrumento básico de la teoría de probabilidades y, por desgracia, su gran simplicidad se ve a veces oscurecida por una terminología singularmente inadecuada"*.

En consecuencia de la relevancia citada, pensamos que no se ha prestado la debida atención a este concepto en la enseñanza secundaria ni universitaria. Todo ello justifica el interés de una investigación sobre el tema.

1.3. MARCO TEÓRICO

En este trabajo vamos a utilizar algunas nociones teóricas desarrolladas por Godino y colaboradores en el Enfoque Ontosemiótico (EOS). Este marco teórico propone tres dimensiones en el análisis de la enseñanza y el aprendizaje de las matemáticas, que son: la epistemológica, la cognitiva y la instruccional. Cada una de ellas se aborda con herramientas teóricas agrupadas en tres modelos teóricos: teoría de los significados institucionales y personales de los objetos matemáticos (Godino y Batanero, 1994, 1998a y b), teoría de las funciones semióticas (Godino, 2002; Godino, Batanero y Roa, 2005) y teoría de las configuraciones didácticas (Godino, Contreras y Font, 2006). Además, recientemente se interesa por definir un modelo de conocimientos del profesor (Godino, 2009). En nuestro trabajo utilizaremos tan sólo parte de los conceptos que forman este marco teórico.

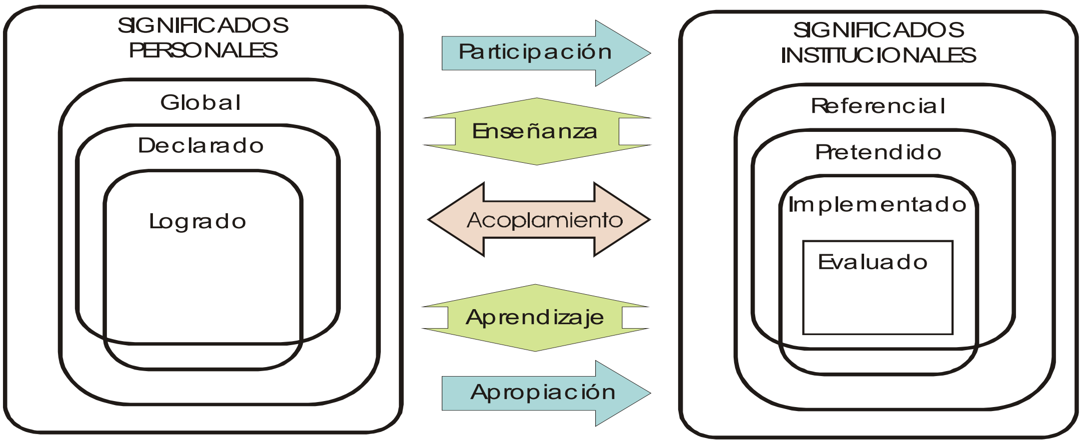
1.3.1. LA ACTIVIDAD MATEMÁTICA Y OBJETOS LIGADOS A ELLAS

En este marco teórico se considera la matemática desde un triple punto de vista: (a) como actividad de resolución de problemas (internos o externos a las matemáticas) socialmente compartida; (b) como un lenguaje simbólico propio en que se expresan las ideas matemáticas; y (c) como un sistema conceptual lógicamente organizado. Para hacer operativo este triple carácter de la matemática se toma como noción primitiva la de situación-problemática, a partir de la cual definen los conceptos de práctica, objeto (personal e institucional) y significado, que describimos a continuación (Godino, 2002).

Los autores de este marco teórico conciben la actividad matemática como un conjunto de prácticas, de las cuáles surgen los objetos matemáticos. Una *práctica matemática* es "*toda actuación o expresión (verbal, gráfica, etc.) realizada por alguien para resolver problemas matemáticos, comunicar a otros la solución obtenida y validarla o generalizarla a otros contextos y problemas*" (Godino y Batanero, 1994, p. 334). En el estudio de las matemáticas, más que una práctica particular ante un problema concreto, interesa considerar los sistemas de prácticas de las personas en su actuación ante tipos de situaciones problemáticas. A la pregunta, qué es un objeto matemático se propone como respuesta: "*el sistema de prácticas que realiza una persona (significado personal), o compartidas en el seno de una institución (significado institucional) para resolver un tipo de situaciones-problemas ligadas con el objeto*" (Godino y Batanero, 1994, p. 335).

Puesto que los significados dependen de los contextos sociales y de los sujetos, su carácter es relativo. Su uso en el análisis didáctico lleva, en consecuencia, a introducir la tipología básica de significados que se resume en la Figura 1.3.1 (Godino, 2003, p. 141). Respecto al significado institucional, se diferencia entre el global (qué significa el objeto en una institución), referencial (qué significado del objeto se considera en una enseñanza o investigación), pretendido (qué se pretende enseñar), implementado (qué se logra enseñar) y evaluado (qué parte se evalúa).

Figura 1.3.1 Tipos de significados institucionales y personales



Respecto al significado personal, se diferencia entre global (todo lo que un sujeto conoce), evaluado (lo que podemos evaluar a priori de su conocimiento), declarado (lo que, pasada la evaluación, hemos conseguido evaluar) y logrado (la parte del conocimiento que está de acuerdo con el significado institucional).

En la parte central de la Figura 1.3.1 se representan las relaciones entre enseñanza y aprendizaje, que supone el acoplamiento progresivo entre los significados personales e institucionales. La enseñanza implica la participación progresiva del estudiante en la comunidad de prácticas que soporta los significados institucionales, y el aprendizaje implica la apropiación por el estudiante de dichos significados (Godino, 2003).

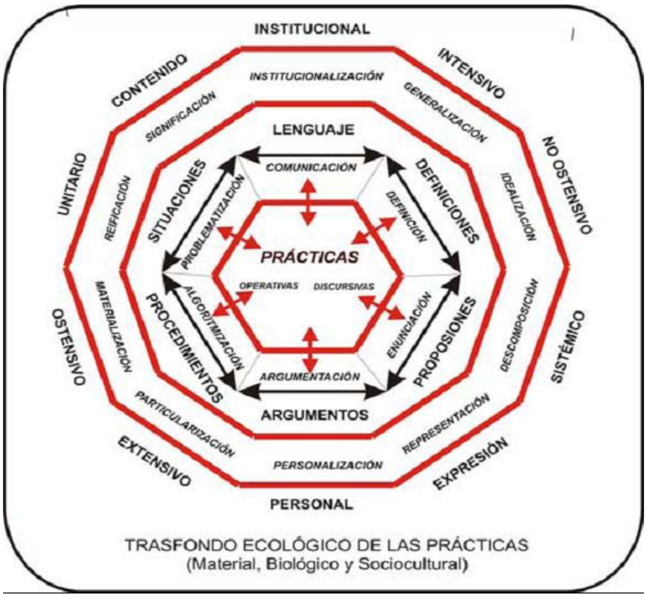
1.3.2. OBJETOS EMERGENTES E INTERVINIENTES EN LOS SISTEMAS DE PRÁCTICAS

Godino, Batanero y Font (2007) describen diferentes categorías en los objetos ligados a las prácticas matemáticas, que serán objetos institucionales si los sistemas de prácticas son compartidos en el seno de una institución o serán objetos personales si dichos sistemas de prácticas son realizados por una persona. Es decir, se tienen en cuenta las dimensiones social y personal del conocimiento y también el hecho de que una práctica personal pudiera ser o no adecuada desde el punto de vista de la institución. A continuación se describen las categorías de objetos propuestas:

- *Situaciones-problemas*: Son todas las aplicaciones extra-matemáticas, ejercicios, problemas o situaciones que inducen una actividad matemática. En nuestro caso el problema puede ser la búsqueda de una estrategia óptima en un juego, la toma de una decisión o bien problemas planteados por el profesor con relación a un recurso didáctico relacionado con la probabilidad condicional.
- *Lenguajes*: Los términos, expresiones, notaciones y gráficos que se utilizan para representar los datos del problema, las operaciones que hacemos con ellos, los objetos matemáticos que se utilizan y la solución encontrada. En la resolución de los problemas citados, se usa el lenguaje verbal, icónico y simbólico, así como el lenguaje gráfico (por ejemplo, el diagrama en árbol).
- *Conceptos-definición*: En las prácticas que llevan a cabo los estudiantes para resolver un problema matemático se usan implícita o explícitamente objetos matemáticos, de los cuáles el alumno ha de recordar o aplicar la definición. Por ejemplo, los estudiantes usarán implícitamente los objetos: aleatoriedad, espacio muestral, suceso, probabilidad, o independencia.
- *Proposiciones* o enunciados sobre relaciones o propiedades de los conceptos que igualmente se han de emplear al resolver problemas matemáticos. Por ejemplo, cuando los estudiantes tienen que recordar que la suma de probabilidades en el espacio muestral es igual a la unidad o la regla de la suma y el producto.
- *Procedimientos*: Serían los algoritmos, operaciones, técnicas de cálculo que los estudiantes han aprendido durante la enseñanza previa y que aplican al resolver el problema. En nuestro caso, los estudiantes usarán técnicas sencillas de cálculo de probabilidades, como técnicas combinatorias, uso de diagrama en árbol, etc.
- *Argumentos*: Serían los enunciados usados para validar o explicar las proposiciones y procedimientos o bien la solución de los problemas.

Godino, Batanero y Font (2007) indican que las seis categorías de objetos descritos están relacionados entre sí formando *configuraciones*, es decir redes de objetos intervinientes y emergentes de los sistemas de prácticas. Los autores clasifican estas configuraciones en *epistémicas* (cuando se trata de objetos institucionales) o *cognitivas* (si se refieren a objetos personales). La configuración epistémica es el conjunto de objetos matemáticos que intervienen en la resolución de las actividades. Dentro de la configuración se distingue la *previa* (los objetos que se supone el alumno conoce antes de trabajar en la unidad didáctica) y la *emergente* (lo que suponemos se va a aprender).

Figura 1.3.2. Objetos y procesos que intervienen en las prácticas matemáticas



Como vemos en la Figura 1.3.2, cada uno de estos objetos tiene asociado un proceso matemático: problematización, comunicación, definición, enunciación, argumentación y algoritmización (Godino, Batanero y Font, 2007). Además estos autores contemplan los objetos anteriores desde diferentes facetas duales:

- *Personal – institucional*. Si los sistemas de prácticas son compartidas en el seno de una institución, los objetos emergentes se consideran “objetos institucionales”, mientras que si estos sistemas son específicos de una persona se consideran como “objetos personales” (Godino y Batanero, 1994).
- *Ostensivo – no ostensivo*. Se entiende por ostensivo cualquier objeto que se puede mostrar a otro, es decir, materializado de alguna forma.
- *Expresión – contenido*. Antecedente y consecuente de cualquier función semiótica (que definimos en la siguiente sección).
- *Extensivo – intensivo*. Un objeto que interviene como un caso particular (un ejemplo específico de la media) y una clase más general, como “las medias de todas las muestras de una población”.
- *Unitario – sistémico*. En algunas circunstancias los objetos matemáticos se consideran unitarios (un dato) y en otros como compuestos de varios objetos (la distribución de datos).

También en la Figura 1.3.2 vemos que se consideran una serie de procesos asociados a estas dualidades: institucionalización – personalización; generalización – particularización; análisis/descomposición – síntesis/reificación; materialización/ concreción – idealización/abstracción; expresión/representación – significación.

1.3.3. RELACIONES ENTRE OBJETOS: FUNCIÓN SEMIÓTICA

Otro punto de interés en nuestro estudio es el de función semiótica, que sirve para resaltar los procesos de interpretación que se llevan a cabo en la actividad matemática y en los cuales a veces pueden aparecer desajustes (conflictos) de interpretación entre el alumno y el profesor.

Los autores del marco teórico describen la idea de función semiótica como correspondencia entre un antecedente y un consecuente, establecida por un sujeto (persona o institución). Por ejemplo, un Applet en Internet, puede considerarse como una función semiótica, donde el antecedente es el propio Applet, que representa visualmente un cierto juego (como el juego de Monty Hall) o bien un procedimiento o teorema matemático (por ejemplo, el teorema de Bayes). La correspondencia suele ser implícita, aunque en algunos casos, se incluyen instrucciones para la interpretación de los recursos. Tanto si hay instrucciones como si no, es posible que el estudiante que trabaja con el recurso no comprenda su funcionamiento o no interprete correctamente los resultados producidos. Si las interpretaciones hechas por el estudiante no son las esperadas por el profesor, diremos que hay un conflicto semiótico.

El antecedente (expresión) y el consecuente (significado) de una función semiótica no se restringen a conceptos, sino abarcan toda la anterior ontología de objetos matemáticos u organizaciones de estos en entidades más complejas, como sistemas conceptuales o teorías. Así en nuestro estudio, podemos descomponer el trabajo con tareas o recursos, generalmente en partes, tras las cuales aparecen diferentes objetos matemáticos implícitos. Se tratará de poner de manifiesto estos objetos, así como las dificultades que los estudiantes o profesores pueden tener en su uso.

1.3.4. CRITERIOS DE IDONEIDAD DIDÁCTICA

Otro concepto que usamos en nuestro trabajo es la noción de idoneidad didáctica, que se define como la articulación de seis componentes (Godino, Wilhelmi y Bencomo, 2005; Godino, Contreras y Font, 2006):

- *Idoneidad epistémica*: Representatividad de los significados institucionales implementados (o pretendidos) en una situación o recurso didáctico, respecto de un significado de referencia. Un alto grado indica que los significados de los objetos presentes en un recurso son adecuados desde el punto de vista matemático.
- *Idoneidad cognitiva*: Grado en que los significados pretendidos/implementados en una situación o recurso didáctico son asequibles a los alumnos, así como si los significados personales logrados por los alumnos son los significados pretendidos por el profesor.
- *Idoneidad interaccional*: Grado en que la organización de la enseñanza permite identificar conflictos semióticos y resolverlos durante el proceso de instrucción. Dependerá del discurso del profesor en el aula y de la forma de trabajo de los alumnos.
- *Idoneidad mediacional*: Se refiere a la disponibilidad y adecuación de los recursos necesarios para el desarrollo del proceso de enseñanza-aprendizaje. Dichos recursos pueden ser materiales tradicionales o tecnológicos, libros o guías de estudio, tiempo disponible o cualquier otro recurso.
- *Idoneidad emocional*: Interés y motivación del alumnado en el proceso de estudio.

1.3.5. NIVELES DE ANÁLISIS PROPUESTOS EN EL EOS

De acuerdo a Godino, Font y Wilhelmi (2008) una tarea importante para el profesor es valorar la propia práctica docente con la finalidad de favorecer el aprendizaje de los estudiantes. Para facilitar esta valoración los autores describen diversos niveles de análisis. En concreto, usaremos en nuestro trabajo los siguientes niveles de análisis:

1. *Sistemas de prácticas y objetos matemáticos (previos y emergentes)*. Este nivel de análisis se aplica durante la planificación de la enseñanza y pretende estudiar las prácticas matemáticas planificadas. Permite descomponer el proceso de estudio previsto en una secuencia de episodios y, para cada uno de ellos, describir las prácticas supuestas. Como consecuencia se identifica una configuración epistémica global.

2. *Procesos matemáticos y conflictos semióticos*. En toda práctica se identifica un *sujeto agente* (institución o persona). Este nivel de análisis se centra en los objetos y procesos que intervienen en la realización de las prácticas previstas, y también en los que emergen de ellas. La finalidad es describir la complejidad de las prácticas matemáticas como factor explicativo de los conflictos semióticos que se podrían producir en su realización.
3. *Idoneidad didáctica del proceso de estudio*. Se trata de analizar si se verifican los diferentes tipos de idoneidad descritos en el punto anterior. Para ello los autores proporcionan unas guías de descriptores para poder valorar cada uno de los tipos de idoneidad.

1.4. LA PROBABILIDAD CONDICIONAL Y LOS OBJETOS MATEMÁTICOS RELACIONADOS

En lo que sigue analizamos el objeto matemático "probabilidad condicional" y los objetos relacionados con él. Tomamos como base el trabajo de Díaz (2004) quien, para construir un cuestionario de evaluación sobre razonamiento condicional, lleva a cabo un análisis del contenido sobre probabilidad condicional en los libros de estadística usados en las universidades españolas en la licenciatura de psicología. Se amplía con algunos textos de estadística o probabilidad.

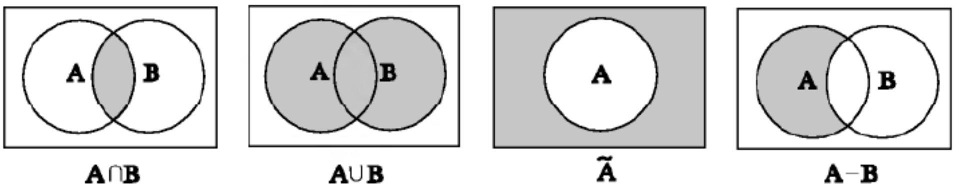
1.4.1. PROPIEDADES Y CONCEPTOS RELACIONADOS CON LA PROBABILIDAD CONDICIONAL

Álgebra de sucesos

Para introducir la probabilidad condicional, necesitamos primeramente algunas ideas previas, entre ellas los axiomas de probabilidad propuestos por Kolmogorov. Partiendo de un experimento aleatorio, definimos *espacio muestral* Ω como el conjunto de resultados posibles, es decir el conjunto de todos los sucesos aleatorios asociados a un experimento aleatorio.

Sea ρ una clase no vacía de conjuntos de Ω, que se denominarán sucesos aleatorios o simplemente sucesos. Se llama *suceso aleatorio A*, a cualquier subconjunto del espacio muestral. Se consideran las operaciones entre sucesos, dentro del mismo espacio muestral Ω: el suceso unión $A \cup B$, el suceso intersección $A \cap B$ y el suceso complementario $\bar{A}$, que se representan en la Figura 1.4.1. Estas operaciones son similares a las definidas entre conjuntos y tienen sus mismas propiedades, por lo que ρ tiene una estructura de Álgebra (de Boole o más generalizada).

Figura 1.4.1. Operaciones entre sucesos



Sí $A \cap B = \emptyset$; entonces decimos que son *incompatibles* o *mutuamente* excluyentes. Este concepto será importante en nuestro análisis pues los alumnos confunden sucesos excluyentes e incompatibles. Dados $\{A_i\}$ decimos que forman un sistema completo de sucesos si son mutuamente excluyentes y $\bigcup_{i=1}^{n} A_i = \Omega$.

Probabilidad. Axiomática de Kolmogorov

Dado un espacio muestral Ω, y un σ-álgebra de sucesos ρ sobre él, diremos que cualquier función P definida sobre una clase *P* es una probabilidad sobre ρ si cumple los tres axiomas de probabilidad:

1. P: $\rho \rightarrow [0,\ 1] \in \mathbb{R}$. La probabilidad es una función definida sobre ρ y que sólo toma valores positivos comprendidos entre 0 y 1.
2. P(Ω) = 1, La probabilidad del suceso seguro es 1.
3. Si *A*, *B* $\in \rho$; $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$, sí $\{A_i\}_{i=1}^{n} \in \rho$; $A_i \cap A_j = \emptyset$

   $\Rightarrow P(\cup A_i) = \sum_i P(A_i)$; es decir si un conjunto de eventos son mutuamente excluyentes o incompatibles, no ocurren simultáneamente, entonces la frecuencia relativa de su unión es la suma de las frecuencias relativas de cada uno.

Por lo tanto P es una función numérica sobre ρ es normada, no negativa y finitamente aditiva (medible) que toma sus valores en el intervalo [0,1]. Al valor de *P* para un suceso *A* se denominará *probabilidad de A* y se designará por $P(A)$. El par (ρ, P) se denomina *campo de probabilidad* y la terna (Ω, ρ, P) se denomina *espacio de probabilidad* o espacio probabilístico (Loeve, 1976).

Probabilidad condicional

Sean (Ω, ρ, P) un espacio probabilístico, sean dos sucesos $A, B \in \rho$ con probabilidades respectivas de ocurrencia $P(A)$ y $P(B)$. Dado el suceso *A* tal que $P(A) > 0$, cabe preguntarse si la probabilidad de que ocurra *B* seguirá siendo la misma o habrá sufrido alguna modificación. Para calcular la probabilidad del suceso *B*, condicionada por el suceso *A*, suponemos que el espacio muestral Ω está formado por *n* sucesos elementales de los cuales n_1 son favorables a la realización del suceso *A*, n_2 lo son a *B* y n_T lo son a la del suceso $A \cap B$.

Puesto que ponemos la condición del suceso *A* se produce una restricción del espacio muestral que ahora coincide con *A*. Los sucesos posibles serán los que forman parte de *B*. Los favorables son los que forman parte de la intersección. Por tanto, la probabilidad del suceso *B* condicionado por *A*, que representamos por *P(B/A)* es la razón entre las probabilidades $P(A \cap B)$ y $P(A)$.

$$P(B / A) = \frac{P(A \cap B)}{P(A)}; \text{ Siempre que } P(A) > 0.$$

Definición intuitiva de probabilidad condicional

Algunos libros definen intuitivamente la probabilidad condicional *P(A/B)* de un suceso *A* dado otro suceso *B* a partir de la frecuencia relativa. Por ejemplo, Peña (1986, p. 76) define la frecuencia relativa de *A* condicionada a la ocurrencia de *B*, en la forma siguiente: "*considerando únicamente los casos en los que aparece B y viendo en cuántos de esos casos ocurre el suceso A; es, por tanto, igual a la frecuencia de ocurrencia conjunta de A y B, partida por el número de veces que ha ocurrido B*". De esta definición deduce el concepto de probabilidad condicional, usando la siguiente notación:

$$P(B / A) = \frac{P(AB)}{P(A)},$$

donde *AB* representa el suceso ocurrencia conjunta de *A* y *B*, siempre suponiendo que *P(B)>0* (Peña, 1986, p. 76). En esta definición está implícita la restricción del espacio muestral. El autor no usa explícitamente el símbolo de la intersección de sucesos para la probabilidad conjunta.

Propiedades de la probabilidad condicional

Nortes (1993) indica específicamente y demuestra que una probabilidad condicional cumple los tres axiomas de la probabilidad:

1. La probabilidad condicional es siempre un número entre 0 y 1; $P\ (A/B) \leq 1$.
2. El axioma de la unión: $P(A \cup B/C) = P(A/C) + P(B/C) - P(A \cap B/C)$.
3. $P(E/A) = \dfrac{P(E \cap A)}{P(A)} = \dfrac{P(A)}{P(A)} = 1$. La probabilidad condicional del suceso seguro es igual a 1.

De estos axiomas principalmente podemos reducir las propiedades de la probabilidad condicional a las siguientes:

- En general *P(A/B)* es distinto a *P(B/A).* Uno de los errores frecuentemente descritos sobre la comprensión de la probabilidad condicional es que los sucesos *A* y *B* a veces se intercambian o no se tiene conciencia de que *P(A/B)* puede ser diferente de *P(B/A).*
- $P(\overline{A}/B) = 1 - P(A/B)$. La probabilidad condicional del suceso complementario a uno dado es igual a la unidad menos la probabilidad condicional del complementario.
- $P(\varnothing/A) = 1 - P(\Omega/A) = 0$; esto implica que la probabilidad conjunta de dos sucesos mutuamente excluyentes es igual a cero.
- $P(A \cup B/C) = P(A/C) + P(B/C) - P(A \cap B/C)$. Es decir la probabilidad de la unión de dos sucesos condicionada a otro es igual a la suma de las probabilidades condicionadas menos la probabilidad de la intersección condicionada.
- Sí $A \subset B \Rightarrow P(A/C) \leq P(B/C)$. Si un suceso está contenido en otro, la probabilidad del primero condicionada a otro suceso es siempre menor que la probabilidad del segundo condicionada al mismo. Si $A \subset B$ tenemos la siguiente propiedad (Poularikas, 1999): $P(A/B) = \dfrac{P(A)}{P(B)} \geq P(A)$.

- Si $A \subset B \Rightarrow P(A/B) = \dfrac{P(A)}{P(B)}$ y $P(B/A) = 1$. Si un suceso está contenido en otro, la probabilidad del primero condicionada al segundo es igual al cociente de las probabilidades respectivas. Si condicionamos el segundo al primero la probabilidad condicionada vale 1.
- $P(A \cap B) = P(A)P(B/A)$. Se obtiene fácilmente partiendo de la definición de la probabilidad condicionada $P(B/A)$.
- Para cualquier par de sucesos *A, B y C*, se cumple que:

$$P(A) = P(A \cap B) + P(A \cap \overline{B}) = P(B)P(A/B) + P(\overline{B})P(A/\overline{B}) .$$

Y en particular tenemos que:

$$P(A/C) = P(A \cap B/C) + P(A \cap \overline{B}/C) = \frac{P(A \cap B \cap C)}{P(C)} + \frac{P(A \cap \overline{B} \cap C)}{P(C)} .$$

Si *A* y *B* son mutuamente excluyentes. Tenemos las siguientes propiedades (Poularikas, 1999):

1. $P(A/B) = 0$;
2. $P(A \cup B/C) = P(A/C) + P(B/C)$;
3. $\dfrac{P(A/B)}{P(B)} = \dfrac{P(B/A)}{P(A)}$.

Peña (1986) hace hincapié en las siguientes propiedades:

- Es importante diferenciar entre $P(A \cap B)$ y *P(A/B).*
- Una probabilidad conjunta $P(A \cap B)$ es siempre menor que las probabilidades simples *P(A)* y *P(B).*
- Una probabilidad condicionada *P(A/B)* puede ser mayor, menor o igual que *P(A).*
- El espacio muestral en la probabilidad condicional *P(A/B)* queda restringido a *B.*

Respecto a esta última propiedad, Cuadras y cols. (1988) distinguen entre *P(A),* probabilidad libre o incondicionada, donde el espacio muestral es Ω, y *P(A/B)* probabilidad condicionada, donde el espacio muestral es *B. B* es el suceso condicionante o lo que Cuadras y cols. denominan “información”. Díaz (2004) recuerda que nunca podría aparecer en una fórmula una suma o diferencia de probabilidades condicionadas a diferentes sucesos, dado que en este caso las probabilidades estarían definidas sobre distintos espacios muestrales.

Aparte de las propiedades y los axiomas de probabilidad condicionada se han de tener en cuenta los siguientes conceptos relacionados.

Independencia

Para Borovnick, Bentz y Kapadia (1991) la independencia es el concepto que distingue la teoría de la probabilidad, de la teoría de la medida. Es parte de las definiciones fundamentales en esta rama de las matemáticas, pero no un axioma básico.

Ortiz (1999) indica que es una hipótesis clave en teoremas fundamentales, como el de Bernoulli, que establece un puente del enfoque estructural a la interpretación frecuencial y por ello la identificación de este concepto contribuyó a la aceptación rápida de la axiomática de Kolmogorov por la comunidad científica. Con el concepto de independencia de ensayos similares es posible deducir una medida de probabilidad en el espacio producto con solo conocer la distribución de probabilidad en los espacios muestrales de cada experimento simple. Por ello se convirtió en una pieza clave en la axiomática de Kolmogorov.

Dos sucesos A y B, tales que $P(B/A) = P(B)$ se conocen con el nombre de sucesos "independientes". Si por el contrario $P(B/A) \neq P(B)$ los sucesos los denominaremos "dependientes". Dicho en palabras, los sucesos son independientes si la probabilidad de uno de ellos no cambia al condicionarlo por el otro. Otra forma de definir la independencia es a partir de la regla del producto, aunque es menos intuitiva para los alumnos: A y B son independientes si y sólo si $P(A \cap B) = P(A) \times P(B)$.

Cuadras y cols. (1988) afirman que esta propiedad es condición necesaria y suficiente para que dos sucesos sean independientes. Esta regla del producto se puede extender a más sucesos independientes:

$$P\ (A_1 \cap A_2 \cap ... \cap A_k) = P(A_1) \times P(A_2) \times ...\ \times P(A_k)\ .$$

La independencia entre sucesos, aunque se puede prever en algunos casos como en lanzamientos de dados u otros juegos de azar, generalmente debe determinarse experimentalmente. Esto hace que, aunque la definición de independencia sea sencilla matemáticamente, los estudiantes tengan muchas dificultades al aplicarla para resolver problemas.

Cuadras y Cols. (1988) indican también algunas propiedades más de la independencia:

- *Si A* y *B* son dos sucesos independientes, también son independientes $A\ y\ \overline{B}$; $\overline{A}\ y\ B$; $\overline{A}\ y\ \overline{B}$.
- Es interesante observar que la relación de independencia es simétrica, puesto que: *P(B/A)=P(B)* implica que *P(A/B)=P(A).* En efecto, Si *P(B/A)= P(B),* se tendrá que $P(A/B)=P(A\cap B)/P(B)=P(A)\times P(B)/P(B)=P(A)$.

Montgomery y Runger (2002) destacan además que la independencia implica resultados relacionados como:

$$P(\overline{A}\cap\overline{B})=P(\overline{A})\times P(\overline{B}).$$

El concepto de independencia es una relación importante entre sucesos. A menudo, la independencia es asumida para ser la parte del experimento arbitrario que describe el sistema físico en el estudio. Hwei y Hsu (2003) añaden la siguiente propiedad: Si *A, B y C tal que* $P(A\cap B\cap C)=P(A)\times P(B)\times P(C)$ entonces:

- *A, B y C* son *independientes*;
- $P(A\cap B)=P(A)\times P(B)$;
- $P(A\cap C)=P(A)\times P(C)$;
- $P(B\cap C)=P(B)\times P(C)$.

Dependencia

Del concepto de independencia se deduce el concepto de sucesos dependientes. Cuadras y cols. (1988) lo definen de esta forma: *"Si* $P(A/B)\neq P(A)$, *la presencia de B altera la probabilidad de A y se dice que el suceso A es estocásticamente dependiente del suceso B. Si* $P(A/B)>P(A)$, *la presencia del suceso B favorece al suceso A. Si* $P(A/B)<P(A)$, *la presencia del suceso B desfavorece al suceso A".* Igualmente podemos decir que:

- *Cuando P(B/A)=P(B)* entonces la ocurrencia de *A* no tiene ningún efecto sobre la de *B*;
- Cuando *P(B/A)>P(B)* entonces el suceso *A* favorece al *B;*
- Cuando *P(B/A)< P(B)* entonces el suceso *A* desfavorece al *B;*
- Cuando *P(B/A)= P(B)* entonces la ocurrencia de *A* no tiene ningún efecto sobre *B.*

Como hemos indicado los alumnos confunden sucesos independientes y *mutuamente excluyentes.* También hemos visto que la probabilidad conjunta de dos sucesos mutuamente excluyentes es igual a cero. Por ello, dos sucesos mutuamente excluyentes son dependientes, ya que la ocurrencia de uno hace que el otro no pueda ocurrir.

Intercambiabilidad

En algunos de los textos de enfoque bayesiano se incluye este concepto, como hipótesis no tan restrictiva como la independencia. Por ejemplo, en Berry (1995, p. 134) se presenta este concepto en la forma siguiente: *"Dos experimentos son intercambiables si se cumple lo siguiente:*

1. *Los posibles resultados son los mismos en los dos;*
2. *La probabilidad de cada resultado de uno es la misma que en el otro;*
3. *La probabilidad condicional para el segundo experimento, dados los resultados del primero son las mismas que las probabilidades condicionales del primero, dado el segundo".*

Regla del producto

Una propiedad relacionada con las anteriores es la regla del producto, que permite calcular las probabilidades en experimentos compuestos. Esta regla es la siguiente:

$$P(A \cap B) = P(A) \times P(B / A).$$

Algunos de textos de estadística, como el de Nortes (1993), diferencian entre sucesos independientes y dependientes para la aplicación de la regla del producto, proporcionando explícitamente las fórmulas, o bien ejemplos de su aplicación, en los dos casos. Para dos sucesos independientes, la regla del producto se define en la fórmula siguiente:

$$P(A \cap B) = P(A) \times P(B) .$$

Botella y cols. (1993, p. 283) introducen el teorema del producto del siguiente modo: "*La probabilidad de verificación simultánea de dos sucesos independientes es igual al producto de sus probabilidades simples*". Para más de dos sucesos independientes, se iría generalizando la regla. Por ejemplo, para tres sucesos independientes la regla es la siguiente:

$$P(A \cap B \cap C) = P(A) \times P(B) \times P(C) .$$

Para n sucesos independientes la regla del producto es como sigue:

$$P\ (A_1 \cap A_2 \cap ... \cap A_k) = P(A_1) \times P(A_2) \times ...\ \times P(A_k).$$

Para dos sucesos dependientes, la probabilidad de que ocurran A y B (sucesivas o simultáneas) es igual a la probabilidad de que ocurra A por la probabilidad de que ocurra B condicionando de A, es decir, según expresa Nortes (1993, p. 219):

$$P(A \cap B) = P(A) \times P(B / A) .$$

En el caso de más de dos sucesos dependientes, la regla del producto sería:

$$P(A \cap B \cap C) = P(A) \times P(B / A) \times P(C / A \cap B) .$$

Teorema de la Probabilidad Total

Introducidas las nociones de dependencia e independencia y la regla del producto podemos pasar al teorema de la probabilidad total, que expresa la probabilidad de un suceso cuando se conocen las probabilidades de que este suceso sea debido a cada una de una serie de causas mutuamente incompatibles. Sea $\{A_i\}$ una sucesión de sucesos disjuntos dos a dos, y tales que $P(\bigcap_{i=1}^{n-1} A_i) > 0$ y su unión constituya una partición del espacio muestral (sistema completo de sucesos). Se verifica que:

$$P(A_1 \cap ... \cap A_n) \ = \ P(A_1)\ P(A_2 /\ A_1)\ P(A_3 /\ A_1 \cap\ A_2)\ ...\ P(A_n /\ A_1 \cap ... \cap A_{n\text{-}1})$$

$$P\left(\bigcup_{i=1}^{n} A_i \ / \ A\right) \ = \ \sum_{i=1}^{n} P(A_i \ / \ A) .$$

Teorema de Bayes

Sean n sucesos disjuntos $B_1, B_2, ..., B_n, A \in \ \Omega$ tales que $P(B_i) > 0, P(A) > 0;$ $i = 1,\ 2,\ ...,\ n$; tales que forman un sistema completo de sucesos. Se verifica que:

$$P(Bi\ /\ A)\ =\ \frac{P(A\ /\ Bi)P(Bi)}{\sum_i P(Bi)P(A\ /\ Bi)} = \frac{P(A\ /\ Bi)P(Bi)}{P(A)}.$$

Cuadras y cols. (1988, p. 122) definen los sucesos $\{B_1, B_2, ..., B_n\} \in \Omega$ con $P(B_i) > 0$, como una partición del espacio muestral *E*. A estos sucesos los denominan "*causas*" y al suceso *A* "*efecto*". *A* es un suceso cualquiera que si ocurre, lo hace conjuntamente con uno de los sucesos de la partición $\{B_1, B_2, ..., B_n\}$.

Las probabilidades $P(B_i) > 0;\ i = 1, ..., n$ ***s**e* denominan probabilidades a priori ya que son las que se asignan inicialmente a los sucesos B_i. Las probabilidades $P(A / B_i) > 0;\ i = 1, ..., n$; se denominan verosimilitudes del suceso *A* admitiendo la hipótesis B_i. Las verosimilitudes $P(A / B_i) > 0$ nos permiten modificar nuestro grado de creencia original $P(B_i)$ obteniendo la probabilidad a posteriori $P(B_i / A)$. Sean $B_1, B_2, ..., B_n$ y A k sucesos mutuamente excluyentes y exhaustivos, entonces:

$$P(A) = P(B_1 \cap A) + ... + P(B_k \cap A) = P(B_1)P(A / B_1) + ... + P(B_k)P(A / B_k).$$

1.4.2. PROBLEMAS Y PROCEDIMIENTOS

Díaz (2004, 2007) analiza los problemas relacionados con la probabilidad condicional en una muestra de libros de texto usados en la enseñanza de estadística en Psicología. Encuentra las siguientes categorías:

1. Calcular una probabilidad condicional, dentro de un experimento simple, como en el siguiente ejemplo, tomado de Gutiérrez, Martínez y Rodríguez (1993, p. 47): "*Se lanza un dado. Sabiendo que ha aparecido un número mayor o igual que 5, ¿Cuál es la probabilidad de que sea múltiplo de 3?*".
2. Calcular una probabilidad condicional en un experimento compuesto. Díaz (2004, 2007) diferencia en este caso el contexto de muestreo con reposición y sin reposición. Un ejemplo del caso de muestreo con reposición es el siguiente, tomado de Devore (1998, p. 119): "*Una urna contiene 5 bolas rojas y 8 verdes. Se extrae una bola y se reemplaza por dos del otro color. A continuación, se extrae una segunda bola. Se pide: (a) Probabilidad de que la segunda bola sea verde, (b) Probabilidad de que las dos bolas extraídas sean del mismo color*".

3. Calcular una probabilidad condicional a partir de probabilidades conjuntas y simples mediante la fórmula $P(A/B) = P(A \cap B)/P(B)$, siempre que *P(B) >0*. Un ejemplo es el siguiente problema resuelto de "*Sea A y B eventos con* $P(A) = 0,6$; $P(B) = 0,3$ *y* $P(A \cap B) = 0,2$. *Encuentre* $P(\bar{A})$, $P(\bar{B})$, $P(A/B)$, $P(B/A)$ *y* $P(A \cup B)$" (Lipschutz y Lars, 2001, p. 42).
4. Calcular una probabilidad compuesta haciendo uso de la regla del producto. Díaz (2004, 2007) diferencia los casos de sucesos dependientes e independientes. Un ejemplo, en el caso de suceso dependiente es el siguiente: "*En una bolsa se tienen 4 bolas blancas y 2 bolas negras; otra contiene 3 bolas blancas y 5 bolas negras. Si se saca una bola de cada bolsa, hallar la probabilidad de que ambas sean blancas*" (Spiegel, 2000, p. 139).
5. Determinar si dos sucesos son dependientes o independientes: Martin-Pliego y Ruiz-Maya (2006, p. 41) sugieren el siguiente problema: "*Un juego consiste en arrojar simultáneamente una moneda o un dado. Comprobar que los sucesos en los que está implicados la moneda y el dado son independientes*".
6. Diferenciar entre sucesos mutuamente excluyentes y sucesos independientes. Por ejemplo: "*¿Puede ser independientes dos sucesos mutuamente excluyentes que tienen probabilidad no nula?*" (Peña, 1986, p. 77).
7. Calcular la probabilidad total. Todos los problemas relacionados con el teorema de Bayes implican la resolución de un problema de probabilidad total. Además, algunos libros incluyen explícitamente este tipo de problemas. Por ejemplo el tomado de Gutiérrez, Martínez y Rodríguez (1993, p. 47): "*En una determinada población, el 40% son mujeres, el 30% son universitarios, y el 20% son mujeres universitarias. Si se selecciona al azar una mujer. ¿Cuál es la probabilidad de que sea universitaria?*".
8. Resolver problemas bayesianos. Son los problemas en que se da una partición del espacio muestral y la probabilidad inicial de los sucesos que la componen así como la condicional de un nuevo suceso que depende de los anteriores. Se pide la probabilidad final de los sucesos de la partición. "*Las probabilidades de que un artículo proceda de una fábrica A o de una B son 0,6 y 0,4, respectivamente. La fábrica A produce artículos defectuosos con probabilidad 0,01, y la B, con probabilidad 0,05. Se observa un artículo y resulta defectuoso. Calcula la probabilidad de que provenga de la fábrica A.*" (Arnold, 2004, p. 43).

9. Resolver problemas de probabilidad condicional y conjunta en experimentos compuestos de diferentes experimentos simples. Por ejemplo: *"En una universidad terminan la carrera el 5% de Arquitectura, el 10% de Ciencias y el 50% de Letras (sólo hay estas carreras). Eligiendo un estudiante al azar se pide: (a) Probabilidad de que sea de Arquitectura y haya terminado la carrera; (b) Nos dice que ha terminado la carrera. Probabilidad de que sea de Arquitectura"* (Cuadras y cols., 1988, p. 137).
10. Estimar probabilidades condicionales y conjuntas mediante simulación. Estos problemas sólo los hemos encontrado en los dos libros con enfoque basado en el ordenador. Por ejemplo: *"Considera el ejemplo de prueba sanguínea discutido en el capítulo. Supón que haces tres pruebas y son todas positivas. ¿Tienes una alta probabilidad de tener la enfermedad? Usa Minitab en los cálculos"* (Albert y Rossman, 2001, p. 34).

1.4.3. REPRESENTACIONES Y ARGUMENTOS

Los elementos del lenguaje, tales como términos, expresiones, notaciones, gráficos, etc. son necesarios para resolver los problemas matemáticos, para representar objetos abstractos, así como para argumentar su solución o para describirlos a otras personas. Respecto a la probabilidad condicional hemos encontrados los siguientes.

Términos y expresiones

Encontramos una gran variedad entre los que citamos los siguientes: probabilidad, probabilidad condicionada, diagrama de árbol, aleatorio, conjuntos vacío, total, unión, intersección, probabilidad de la unión, probabilidad de la intersección, función probabilística, variable aleatoria, equiprobabilidad, independencia, dependencia, espacio muestral, suceso, frecuencia relativa, ley de los grandes números, regla de Laplace, regla del producto, convergencia, simulación con uso del ordenador, etc.

Asimismo encontramos *términos asociados a los conceptos relacionados*: distribución de probabilidad, variables aleatorias discretas y continuas, propiedades estadísticas, cálculo de probabilidades de variables y estadísticos, ley de los grandes números. Hacemos notar que cada una de estas expresiones corresponde a conceptos o propiedades matemáticas que el estudiante debería haber aprendido antes de finalizar el tema, lo que muestra la complejidad del objeto de estudio.

Expresiones algebraicas

Expresiones como la siguiente, forman parte del conjunto de representaciones simbólicas relacionadas con la probabilidad y se usan sobre todo al dar *argumentos* de tipo deductivo:

$$P(G) = P((G \cap A) \cup (G \cap B)) = P(G \cap A) + P(G \cap B) = P(G/A)P(A) + P(G/B)P(B).$$

En ella tenemos una demostración de que la probabilidad de un suceso se puede descomponer como la suma de la descomposición de la probabilidad en otros sucesos compuestos. Otras expresiones algebraicas se usan *al argumentar* propiedades o teoremas relacionados: Por ejemplo, es frecuente encontrar expresiones algebraicas desarrollando la fórmula del teorema de Bayes.

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{P(B)} = \frac{P(B \mid A_i)P(A_i)}{\sum_{j=1}^{n} P(B \mid A_j)P(A_j)}.$$

Con ella podemos argumentar la descomposición de la probabilidad condicionada de un conjunto de sucesos mutuamente excluyentes y exhaustivos $\{A_1, A_2, ..., A_i, ..., A_n\}$, tales que la probabilidad de cada uno de ellos es distinta de cero y de un suceso B cualquiera del que se conocen las probabilidades condicionales $P(B/A_i)$, como el cociente de la descomposición de los sucesos simples en otros complejos.

Representaciones y argumentaciones gráficas

Los gráficos son elementos que siempre están presentes en la probabilidad condicional. Se reconoce en los gráficos un medio elocuente de comunicar ideas, para refutar teorías, de representar relaciones entre varias variables, de usarlos para distinguir con ellos el foco de atención, de cuestionar los datos, en sus propósitos, en la precisión de la información cualitativa extraída de ellos y la realización ocasional de funciones instrumentales.

A la hora de representar o argumentar con la probabilidad condicionada es común el uso de los diagramas de árbol. Según Fischbein (1987), el diagrama en árbol facilita la resolución de problemas de probabilidad, aunque algunos autores han mostrado que los alumnos tienen dificultades en construir un diagrama en árbol adecuado al problema. En la Figura 1.4.2 reproducimos un gráfico encontrado en Internet, en una web dedicada a la enseñanza de las tablas de contingencia y la utilidad de los diagramas de árbol: http://thales.cica.es/rd/Recursos/rd98/Matematicas/28/6.html.

Figura 1.4.2. Diagrama de árbol

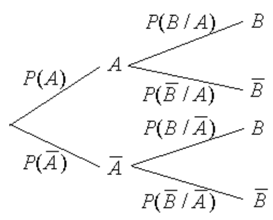
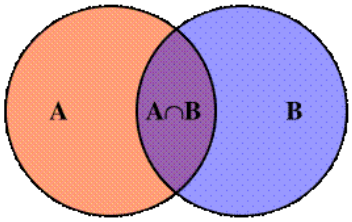


Figura 1.4.3. Representación de sucesos



Otro gráfico que se utiliza a la hora de representar la probabilidad condicionada es la representación de sucesos, mediante un diagrama de Venn (Figura 1.4.3), tomado de la página web http://www.monografias.com/trabajos55/estructuras-objetos-discretos/estructuras-objetos-discretos.shtml. Con este gráfico tenemos una visión espacial de las diferentes situaciones que se pueden presentar en el comportamiento de intersecciones, uniones y como se condicionan unos sucesos a otros.

Representaciones y argumentaciones basadas en el ordenador

Merece mención aparte *la simulación*, la cual permite a través del ordenador estudiar las propiedades de un fenómeno aleatorio reemplazándolo por otro isomorfo. Por medio de dispositivos como el modelo de urnas, es factible simular experimentos probabilísticos, tales como el lanzamiento de un dado, de una moneda, etc. El papel que juega la simulación en la enseñanza de la estadística y la probabilidad ha sido resaltado entre otros por Biehler (1997), Batanero (2003) y Fernández, Batanero, Contreras y Díaz (2009), avisando que la simulación permite poner en manos del estudiante un instrumento que posibilita la exploración y el descubrimiento de conceptos y principios que de otro modo serían mucho más abstractos. Sin embargo, si no se añaden las necesarias “guías prácticas” de la exploración del simulador, o en caso de haberlas se diseñan de manera parcial y difusa, provocará que los estudiantes tengan dificultades para observar el comportamiento del fenómeno (Godino, Contreras y Font, 2006).

1.5. LA PROBABILIDAD CONDICIONAL EN LA ENSEÑANZA NO UNIVERSITARIA

1.5.1. INTRODUCCIÓN

En esta sección describiremos algunos contenidos relacionados con la probabilidad condicional en el currículo español de Educación Secundaria Obligatoria y Bachillerato. Un análisis más completo de estos currículos se recoge en el Anexo 4. Posteriormente analizaremos las recomendaciones de los organismos internacionales sobre la enseñanza del tema, en particular las orientaciones del NCTM (2000) y el proyecto GAISE.

1.5.2. LA PROBABILIDAD CONDICIONAL EN LAS ORIENTACIONES CURRICULARES

La probabilidad en la Educación Primaria

El Ministerio de Educación (MEC, 2006) incluye los siguientes contenidos específicos, sobre azar y la probabilidad en la Educación Primaria:

- *Primer Ciclo*: Tratamiento de la información, azar y probabilidad: Carácter aleatorio de algunas experiencias. Distinción entre lo imposible, lo seguro y aquello que es posible pero no seguro, y utilización en el lenguaje habitual, de expresiones relacionadas con la probabilidad.
- *Segundo Ciclo*: Tratamiento de la información, azar y probabilidad: Valoración de los resultados de experiencias en las que interviene el azar, para apreciar que hay sucesos más o menos probables y la imposibilidad de predecir un resultado concreto. Introducción al lenguaje del azar. Lectura e interpretación de tablas de doble entrada de uso habitual en la vida cotidiana.
- *Tercer Ciclo*: Tratamiento de la información, azar y probabilidad: Presencia del azar en la vida cotidiana. Estimación del grado de probabilidad de un suceso.

Se señala que los juegos de azar proporcionan ejemplos que permitirán introducir las nociones de probabilidad e incertidumbre. La evaluación considerará además de los aspectos propios de la clasificación y representación de datos, la capacidad para deducir relaciones entre ellos y, sobre todo, la deducción de conclusiones y estimaciones a partir de los datos representados.

Para la probabilidad se pretende que el alumnado sea capaz de razonar sobre los posibles resultados de un experimento aleatorio sencillo. A la vez ha asignado probabilidades a sucesos equiprobables o no, utilizando distintas estrategias sobre técnicas de conteo. También se propicia la asignación frecuencial, a partir de experimentos organizados en la clase, que permiten enlazar estadística y probabilidad.

La Junta de Andalucía (Consejería de Educación, 2007a) remite a estos mismos contenidos, añadiendo que los juegos de azar proporcionan ejemplos que permitirán introducir las nociones de probabilidad e incertidumbre. Respecto a los criterios de evaluación repite los citados del ministerio.

La probabilidad en la Educación Secundaria Obligatoria

Respecto a la Enseñanza Secundaria Obligatoria el Decreto de Enseñanzas Mínimas de la Educación Secundaria (MEC, 2007a) incluye, entre otros, los siguientes contenidos dentro del Bloque 6, Estadística y probabilidad:

- *Primer Curso*: Formulación de conjeturas sobre el comportamiento de fenómenos aleatorios sencillos y diseño de experiencias para su comprobación.
- *Tercer Curso*: Sucesos y espacio muestral. Cálculo de probabilidades mediante la regla de Laplace. Formulación y comprobación de conjeturas sobre el comportamiento de fenómenos aleatorios sencillos. Cálculo de la probabilidad mediante la simulación o experimentación. Utilización de la probabilidad para tomar decisiones fundamentadas en diferentes contextos.
- *Cuarto curso. Opción A*: Experiencias compuestas. Utilización de tablas de contingencia y diagramas de árbol para el recuento de casos y la asignación de probabilidades.
- *Cuarto curso. Opción B*: Experiencias compuestas. Utilización de tablas de contingencia y diagramas de árbol para el recuento de casos y la asignación de probabilidades. Probabilidad condicionada.

Se incluyen, asimismo, los siguientes criterios de evaluación:

- *Primer curso*: Hacer predicciones sobre la posibilidad de que un suceso ocurra a partir de información previamente obtenida de forma empírica. Se trata de valorar la capacidad para diferenciar los fenómenos deterministas de los aleatorios y, en estos últimos, analizar las regularidades obtenidas al repetir un número significativo de veces una experiencia aleatoria y hacer predicciones razonables a partir de los mismos.
- *Tercer curso*: Predecir la posibilidad de que un suceso ocurra a partir de información previamente obtenida de forma empírica o como resultado del recuento de posibilidades, en casos sencillos. Se pretende medir la capacidad de identificar los sucesos elementales de un experimento aleatorio sencillo y otros sucesos asociados a dicho experimento. También la capacidad de determinar e interpretar la probabilidad de un suceso a partir de la experimentación o del cálculo (regla de Laplace), en casos sencillos. Por ello tienen especial interés las situaciones que exijan la toma de decisiones razonables a partir de los resultados de la experimentación, simulación o, en su caso, del recuento.

Respecto a la Consejería de Educación de la Junta de Andalucía (2007b), los temas de estadística se incluyen en Educación Secundaria en el Bloque 6: "Interpretación de fenómenos ambientales y sociales a través de las matemáticas". Se citan como contenidos de este Bloque los correspondientes a los Bloques 5, *Funciones y gráficas y 6, Estadística y probabilidad* del Decreto del Ministerio (MEC, 2006b), que se ponen en este currículo en relación directa.

La probabilidad en el Bachillerato

La enseñanza de las Matemáticas en el Bachillerato comprenden los cursos 1º y 2º, y el Decreto 1467/2007, de 2 de noviembre, por el que se establece la estructura del bachillerato y se fijan sus enseñanzas mínimas (MEC, 2007b) fija los siguientes contenidos en relación con el tema que nos interesa:

- *Matemáticas I, modalidad de Ciencias y Tecnología*: Distribuciones bidimensionales. Relaciones entre dos variables estadísticas. Regresión lineal. Estudio de la probabilidad compuesta, condicionada, total y a posteriori. Distribuciones binomial y normal como herramienta para asignar probabilidades a sucesos.

- *Matemáticas aplicadas a las ciencias sociales I, modalidad Humanidades y Ciencias Sociales*: Distribuciones bidimensionales. Interpretación de fenómenos sociales y económicos en los que intervienen dos variables a partir de la representación gráfica de una nube de puntos. Grado de relación entre dos variables estadísticas. Regresión lineal. Extrapolación de resultados. Asignación de probabilidades a sucesos. Distribuciones de probabilidad binomial y normal.
- *Matemáticas aplicadas a las ciencias sociales II, modalidad Humanidades y Ciencias Sociales*: Profundización en los conceptos de probabilidades a priori y a posteriori, probabilidad compuesta, condicionada y total. Teorema de Bayes. Implicaciones prácticas de los teoremas: Central del límite, de aproximación de la binomial a la normal y ley de los grandes números. Problemas relacionados con la elección de las muestras. Condiciones de representatividad. Parámetros de una población. Distribuciones de probabilidad de las medias y proporciones muestrales. Intervalo de confianza para el parámetro *p* de una distribución binomial y para la media de una distribución normal de desviación típica conocida. Contraste de hipótesis para la proporción de una distribución binomial y para la media o diferencias de medias de distribuciones normales con desviación típica conocida.

Otro aspecto a resaltar de los Decretos, son los criterios de evaluación que se contemplan, entre los que se pretende para el primer curso: comprobar la capacidad de apreciar el grado y tipo de relación existente entre dos variables, a partir de la información gráfica aportada por una nube de puntos; así como la competencia para extraer conclusiones apropiadas, asociando los parámetros relacionados con la correlación y la regresión con las situaciones y relaciones que miden. También se sugiere pretende evaluar si, mediante el uso de las tablas de las distribuciones normal y binomial, los alumnos son capaces de determinar la probabilidad de un suceso, analizar una situación y decidir la opción más adecuada.

Para el segundo curso, entre otros criterios de evaluación se establece: Valorar tanto la competencia para estimar y calcular probabilidades asociadas a diferentes tipos de sucesos como la riqueza de procedimientos a la hora de asignar probabilidades a priori y a posteriori, compuestas o condicionadas. Con este criterio se evalúa también la capacidad, en el ámbito de las Ciencias Sociales, para tomar decisiones sencillas de tipo probabilístico.

Se sugiere evaluar la capacidad para identificar si la población de estudio es normal. La competencia para determinar el tipo y tamaño muestral, establecer un intervalo de confianza para μ y p, según que la población sea Normal o Binomial, y determinar si la diferencia de medias o proporciones entre dos poblaciones o respecto de un valor determinado, es significativa.

Al desarrollar los núcleos de contenidos propuestos la Junta de Andalucía (2008) remite a estos contenidos, que deben desarrollarse teniendo en cuenta cuatro núcleos transversales:

1. La resolución de problemas.
2. Aprender de y con la historia de las matemáticas.
3. Introducción a los métodos y fundamentos matemáticos.
4. Modelización matemática, dando sugerencias de posibles contenidos históricos.

1.5.3. RECOMENDACIONES INTERNACIONALES

El interés por la enseñanza de la estadística y probabilidad se muestra en diferentes documentos y recomendaciones internacionales, en especial en los estándares americanos (NCTM 2000), y el proyecto GAISE (Franklin y cols., 2005). Estos documentos no sólo han tenido un impacto en ese país, sino que han influido en el cambio de otros currículos, que han seguido las directrices contenidas en ellos. En lo que sigue resumimos el contenido sobre probabilidad en estos documentos.

Estándares del NCTM

En los estándares curriculares (NCTM, 2000) la probabilidad se inicia desde el nivel K-2 (5 años) y continúa a lo largo de toda la escolaridad. A continuación describimos las competencias recogidas en estos estándares, con relación a este tema:

1. Respecto a la inferencia y la predicción basada en datos:

- Desde preescolar hasta el grado 2, todos los estudiantes deberían poder discutir eventos relacionados con experiencias de los estudiantes, como probable o improbable.
- En los grados 3-5 todos los estudiantes deberían poder: proponer y justificar conclusiones y predicciones basadas en datos y diseñar estudios para sacar conclusiones o predicciones.
- En los grados 6-8 todos los estudiantes deberían emplear observaciones sobre

diferencias entre dos o más muestras para hacer conjeturas sobre las poblaciones de las que se tomaron las muestras; hacer conjeturas sobre las posibles relaciones entre dos características de una muestra sobre la base de diagramas de dispersión; usar conjeturas para formular nuevas preguntas y planificar estudios nuevos para responderlas.

- En los grados 9-12 todos los estudiantes deberían: usar simulaciones para explorar la variabilidad de las muestras de una población conocida y construir distribuciones de muestreo; entender cómo las estadísticas de la muestra reflejan los valores de los parámetros de población y el uso de distribuciones de muestreo como la base para la inferencia; evaluar los informes publicados que se basan en datos mediante el examen del estudio del diseño, la adecuación de los análisis de datos y la validez de las conclusiones; comprender la concepción de las técnicas estadísticas que se utilizan en el lugar de trabajo.

2. Respecto a los conceptos básicos de probabilidad:

- En los grados 3-5 todos los estudiantes deberían describir sucesos probables o improbables y discutir el grado de probabilidad usando palabras como cierto, igualmente probable e imposible. Asimismo deben predecir la probabilidad de resultados de experimentos sencillos y poner a prueba las predicciones; entender que la medida de la probabilidad de un evento puede ser representado por un número de 0 a 1.
- En los grados 6-8 todos los estudiantes deberían comprender y utilizar terminología apropiada para describir sucesos complementarios y mutuamente excluyentes; comprender el uso de proporcionalidad y alcanzar un conocimiento básico de probabilidad para hacer y probar conjeturas sobre los resultados de experimentos y simulaciones. También se espera que sean capaces de calcular probabilidades para eventos simples y compuestos usando métodos como listados organizados, diagramas de árbol y modelos de área.
- En los grados 9-12 todos los estudiantes deberían entender los conceptos de espacio muestral y distribución de probabilidad y construir espacios muestrales y distribuciones en casos sencillos; usar simulaciones para construir distribuciones empíricas de probabilidad; calcular e interpretar el valor esperado de variables aleatorias en casos sencillos; entender los conceptos de probabilidad condicional y eventos independientes y calcular la probabilidad de un suceso compuesto.

Proyecto GAISE

Estas recomendaciones se recogen y amplían en el proyecto GAISE (Franklin y cols., 2005), o "Proyecto de los Lineamientos para la Evaluación e Instrucción en Educación Estadística", que está dirigido a dos grupos de estudiantes: para la educación K-12 y para grupos de estudiantes en cursos preuniversitarios. Para la educación K-12, se indica que cualquier curso de estadística y probabilidad debe tener como principal objetivo ayudar a los estudiantes a aprender los elementos básicos del pensamiento estadístico. Entre otros elementos destacan los siguientes relacionados con la enseñanza de la probabilidad:

- *La omnipresencia de la variabilidad.* Reconocer que la variabilidad es ubicua en muchos fenómenos cotidianos. La variabilidad es la esencia de la estadística como disciplina y no puede ser entendida sólo mediante estudio y lectura, sino que debe ser experimentada.
- *La cuantificación y explicación de la variabilidad.* Reconocer que la variabilidad puede ser medida y explicada tomando en consideración lo siguiente: a) aleatoriedad y distribuciones de las variables aleatorias; b) parámetros de tendencia central y de dispersión (tendencia y residuo); c) modelos matemáticos paramétricos; d) modelos de análisis exploratorio de datos.

En cuanto al curso Preuniversitario, en Estados Unidos la estadística se imparte a través de diversas disciplinas (matemáticas, sociales, ciencias, etc.). Los estudiantes que toman estos cursos provienen de diferentes especialidades de Bachillerato y tienen diferentes objetivos (por ejemplo, algunos esperan hacer sus propios análisis estadísticos en proyectos de investigación y otros sólo hacen el curso porque es obligatorio).

Algunos de esos cursos se imparten en clases extensas, cortas o seminarios; en laboratorios de computación y otros a distancia, sin tener contacto en persona con el profesor. La duración es variable y algunos profesores imparten cursos fuertemente enfocados a la enseñanza en la que los estudiantes deben tener una cultura estadística (saber interpretar, escribir y ser *consumidores* de datos). Otros cursos están sin embargo, más encaminados a la enseñanza de los estudiantes para convertirse en *productores* de análisis estadísticos.

Actualmente los objetivos tienden a enfocarse más en la comprensión de conceptos y a lograr la cultura estadística y pensamiento estadístico de los estudiantes, y menos en el aprendizaje de una serie de herramientas y procedimientos. El proyecto

GAISE propone reexaminar y revisar el curso de estadística con el fin de contribuir al logro de dichos objetivos en el aprendizaje para los estudiantes, haciendo las siguientes recomendaciones:

- Los estudiantes deben entender por qué:
    - La variabilidad es natural y también predictible y cuantificable.
    - La muestra aleatoria en encuestas y experimentos da resultados que pueden ser extendidos a la población de la cual fueron tomados.
    - La asignación aleatoria en experimentos comparativos permite extraer conclusiones de causa y efecto y asociación no es sinónimo de casualidad.
    - La significación estadística no necesariamente implica importancia práctica, especialmente para estudios con muestras grandes.
    - Encontrar diferencias o relación no significativas estadísticamente no necesariamente significa que haya diferencia o relación en la población, especialmente para estudios con muestras pequeñas.
- Los estudiantes deben reconocer:
    - Las fuentes usuales de sesgos en encuestas y experimentos.
    - Cómo determinar la población a la que los resultados de inferencia estadística pueden ser extendidos, dependiendo de cómo se recogieron los datos.
    - Cómo determinar cuándo una inferencia de causa y efecto puede ser deducida de una asociación, dependiendo de cómo se recogieron los datos (por ejemplo, el diseño de un estudio).
    - Que palabras, tales como "normal", "aleatorio" y "correlación" tienen significados específicos en estadística que pueden ser diferentes al uso común.
- Los estudiantes deben entender las ideas básicas de la inferencia estadística:
    - El concepto de distribución muestral y cómo aplicarlo en la inferencia estadística basada en muestra de datos (incluyendo la idea de error típico).
    - El concepto de significación estadística incluyendo nivel de significación y *p*-valor.
    - El concepto de intervalo de confianza, incluyendo la interpretación de nivel de confianza y margen de error.
- Finalmente, los estudiantes deben saber:
    - Cómo interpretar resultados estadísticos en el contexto.
    - Cómo criticar/interpretar noticias y artículos periodísticos que incluyan

información estadística identificando los errores en la presentación y defectos en los estudios o en los métodos utilizados para generar la información.

Vemos, en consecuencia, que el currículo norteamericano es muy avanzado en lo que se refiere a estadística y probabilidad. Este currículo ha tenido mucha influencia tanto en España como en otros países, cuyos diseños curriculares se guían en parte por los norteamericanos, aunque con menor profundidad de contenidos.

1.6. OBJETIVOS DEL TRABAJO

Finalizada la presentación de los fundamentos, estamos en condiciones de precisar los objetivos de nuestra investigación y la forma en que tratamos de llevarlos a cabo. La finalidad de esta investigación es recabar información sobre las necesidades formativas de los profesores en los niveles no universitarios, tanto respecto a su conocimiento matemático, como a su conocimiento didáctico respecto a la probabilidad condicional y analizar algunos recursos didácticos que podrían ser útiles para atender a estas necesidades.

Más concretamente, la investigación se compone de una serie de estudios independientes, aunque todos ellos relacionados con el tema del trabajo y apoyados, tanto en el marco teórico que se ha descrito en este capítulo, como en los antecedentes, revisados en el capítulo 2. Cada uno de estos estudios tiene sus propios objetivos, hipótesis y metodología que se describirán y discutirán con detalle en los diferentes capítulos de la tesis. Sin embargo, para orientar al lector, creemos necesario describir a continuación en forma sintética los objetivos generales y metodología de la investigación. Estos objetivos son los siguientes:

Objetivo 1. La investigación pretende proporcionar alguna información sobre el conocimiento matemático de los futuros profesores de educación primaria y secundaria, respecto a la probabilidad condicional.

Objetivo 2. Asimismo se desea evaluar algunos aspectos del conocimiento didáctico sobre probabilidad condicional en futuros profesores de educación primaria.

Estos dos objetivos, cuya finalidad será utilizar la información recogida para orientar la labor futura de los formadores de profesores, se abordan mediante dos estudios de evaluación. En el Estudio 1, descrito en el capítulo 3 presentamos una evaluación de algunos conocimientos matemáticos y didácticos de una muestra de 183 futuros profesores de educación primaria sobre la probabilidad simple, condicional y compuesta en una tarea sencilla, analizando sus respuestas a dicha tarea y los conflictos semióticos presentados en la solución de la misma. En la segunda parte de la tarea se plantean unas preguntas orientadas a evaluar puntos relevantes del conocimiento didáctico de dichos futuros profesores.

En el capítulo 4 presentamos otro estudio de evaluación (Estudio 2) de los conocimientos matemáticos y sesgos relacionados con la probabilidad condicional de futuros profesores de educación secundaria. A partir de las respuestas a un cuestionario compuesto por ítems de opción múltiple y preguntas de respuesta abierta, se analizan los sesgos de razonamiento y conocimiento matemático formal de una muestra de 196 futuros profesores. Se comparan los resultados con los obtenidos con el mismo cuestionario en estudiantes de psicología y también en los dos subgrupos que componen la muestra (estudiantes de la licenciatura de matemáticas y del máster de formación del profesorado de secundaria).

Objetivo 3. Se pretende analizar algunos recursos educativos que tengan potencialidad para ayudar a los futuros profesores o profesores en ejercicio respecto a las limitaciones observadas en su conocimiento común y especializado del contenido probabilidad condicional.

En el capítulo 5 se describen dos estudios relacionados. En el Estudio 3 se analizan algunos recursos disponibles en Internet útiles para la enseñanza de la probabilidad condicional en diferentes niveles educativos y la formación de los profesores. En el Estudio 4 se estudian algunas paradojas clásicas ligadas a los conceptos de probabilidad condicional e independencia. En ambos casos se analiza la actividad matemática ligada a los recursos, las posibles dificultades en el trabajo con ellos y su idoneidad didáctica.

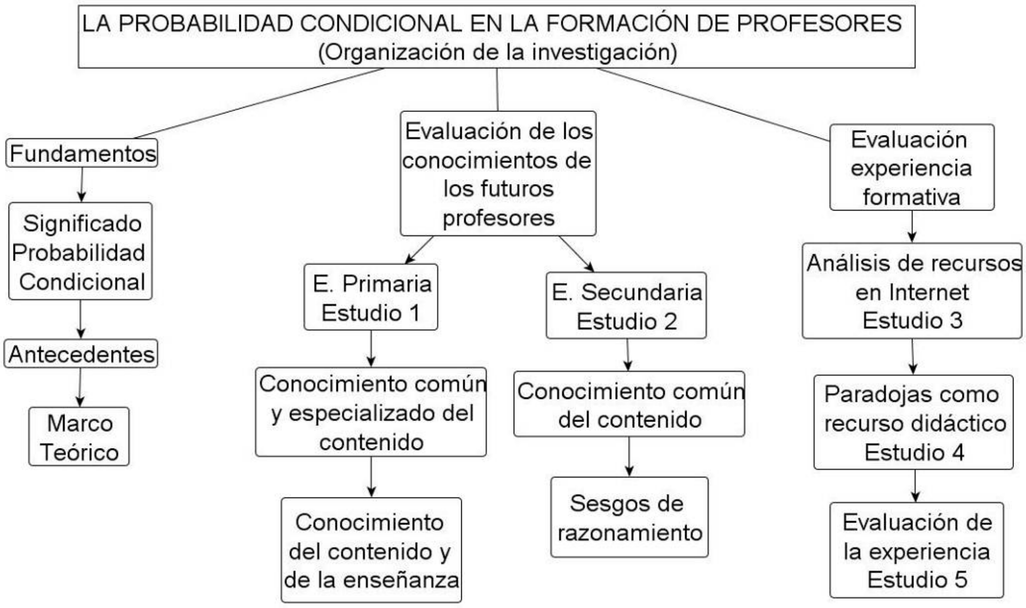
Objetivo 4. Se desea evaluar el cambio en el conocimiento matemático sobre probabilidad condicional de los profesores como consecuencia de una actividad formativa basada en uno de los recursos analizados en el objetivo 4.

En el capítulo 6 se presenta el Estudio 5, donde se evalúa una experiencia formativa basada en una paradoja clásica, dirigida a profesores de educación secundaria o superior en formación y ejercicio, analizando algunos conocimientos matemáticos de los participantes y describiendo su evolución. Se lleva a cabo asimismo una evaluación de componentes de su conocimiento didáctico.

1.7. ENFOQUE METODOLÓGICO Y FASES DE LA INVESTIGACION

Como ya se ha indicado al describir los objetivos, la investigación se ha organizado en una parte teórica (fundamento) y varios estudios empíricos. La parte teórica se compone a su vez (ver Figura 1.7.1) del marco teórico; del estudio de la probabilidad condicional como objeto matemático fundamental; del estudio de la probabilidad condicional en el currículo y de los antecedentes en los que nos basamos para nuestro estudio. Todos estos puntos se han descrito en los capítulos 1 y 2.

Figura 1.7.1. Fases de la investigación



La parte empírica se organiza en cinco estudios, cada uno de los cuales tiene distintos objetivos, hipótesis, poblaciones de estudio, materiales de recopilación de la información y análisis de datos. Todos ellos son de naturaleza, principalmente, descriptiva, exploratoria y cualitativa, excepto el Estudio 2, que mezcla la metodología cualitativa y confirmatoria, al comparar con estudio que nos antecede. En la Tabla 1.7.1 se resumen la metodología utilizada y de los resultados obtenidos en cada uno de los estudios, que se describen someramente a continuación y con más detalle en los capítulos 3 a 6.

- Estudio 1: *Evaluación de conocimientos de futuros profesores de educación primaria,* presentada en el capítulo 3. Se propone una tarea abierta con datos dados en una tabla de contingencia a una muestra de 183 futuros profesores. La primera parte de la tarea consta de tres preguntas en que se ha de calcular la probabilidad simple, condicional y compuesta. El análisis semiótico de las respuestas a estos problemas permite describir algunos conocimientos matemáticos de los futuros profesores, definiendo categorías de respuestas correctas e incorrectas y conflictos semióticos. En la segunda parte de la tarea dos preguntas permiten evaluar puntos relevantes del conocimiento didáctico. Como métodos estadísticos en esta fase se utilizan estadísticos descriptivos y el contraste Chi-cuadrado, así como el contraste t de diferencia de medias para muestras relacionadas.

Tabla 1.7.1. Métodos empleados en los diferentes estudios que componen el trabajo

Estudios	Descripción	Muestras	Material	Análisis	Resultados
Estudio 1	Evaluación de conocimientos de futuros profesores de educación primaria	183 futuros profesores de educación primaria	Tarea abierta con tres problemas y dos preguntas sobre el conocimiento didáctico	Análisis semiótico de respuestas, análisis descriptivos, test Chi-cuadrado y test t de diferencias	Categorías de respuestas en problemas Conflictos semióticos Conocimientos didácticos
Estudio 2	Evaluación de conocimientos de futuros profesores de educación secundaria	196 futuros profesores de educación secundaria de varias universidades españolas (95 alumnos de la licenciatura de Matemáticas y 101 alumnos del máster de secundaria)	Cuestionario de Razonamiento sobre Probabilidad Condicional	Análisis semiótico de respuestas, Análisis descriptivos, test Chi-cuadrado, test t de diferencias correlaciones, análisis factorial	Categorías de respuestas en problemas Conflictos semióticos
Estudio 3	Análisis de recursos en Internet		Recursos en Internet	Análisis semiótico de los recursos	Configuración de objetos, Conflictos semióticos potenciales Idoneidad didáctica
Estudio 4	Paradojas de probabilidad como recurso didáctico		Paradojas de probabilidad condicional e independencia	Análisis semiótico de los recursos	Configuración de objetos, Conflictos semióticos potenciales Idoneidad didáctica
Estudio 5	Evaluación de una experiencia de formación de profesores	166 profesores en activo y futuros profesores de secundaria de España, Portugal y México	Protocolos recogidos en talleres sobre ideas estocásticas fundamentales	Análisis semiótico de respuestas, análisis descriptivos, test Chi-cuadrado y test t de diferencias	Categorías de respuestas en problemas Conflictos semióticos Evolución en el taller Conocimientos didácticos

- Estudio 2. *Evaluación de conocimientos de futuros profesores de educación secundaria*, presentado en el capítulo 4. Se propone a una muestra de 196 futuros profesores de educación secundaria de varias universidades españolas (95 alumnos de la licenciatura de Matemáticas y 101 alumnos del máster de secundaria) una parte del cuestionario RPC (Díaz, 2007). Las respuestas de opción múltiple y las abiertas del cuestionario permiten evaluar por un lado los principales sesgos sobre razonamiento en probabilidad condicional que se describen en los antecedentes. Por otro el conocimiento del concepto, así como la capacidad de resolver los principales tipos de problemas identificados en el apartado 1.4, en nuestro análisis del significado de referencia de la probabilidad condicional en este trabajo. Como técnicas de análisis de datos se usan el análisis semiótico, estudio descriptivo y el test Chi-cuadrado, para analizar las diferencias con los resultados de Díaz (2007) y entre las dos muestras que componen el Estudio. El análisis de correlaciones y análisis factorial permite estudiar la relación entre sesgos en el razonamiento condicional y conocimiento formal del tema.
- *Análisis de recursos para la enseñanza de la probabilidad condicional*: Dividimos este análisis en dos estudios. Estudio 3: Análisis de recursos en Internet. Estudio 4: Paradojas como recurso didáctico en el estudio de la probabilidad condicional. Ambos estudios se presentan en el capítulo 5. En los dos casos se realiza el análisis semiótico de posibles soluciones a los problemas implicados en los recursos, dificultades potenciales de los estudiantes e idoneidad didáctica de los recursos.
- Estudio 5. *Evaluación de una experiencia de formación de profesores de educación secundaria*, descrito en el capítulo 6. Un taller formativo basado en una de las paradojas descritas en el capítulo 4 y propuesto por Batanero, Godino y Roa (2004) es experimentado en varios cursos dirigidos a profesores con un total de 166 profesores en formación o ejercicio. El análisis semiótico de las estrategias iniciales y finales en un juego permite evaluar el cambio de concepciones. Los contrastes Chi-cuadrado y *t* de comparación de dos muestras se utilizan para analizar los resultados sobre una parte de los conocimientos didácticos de los participantes.

CAPÍTULO 2

INVESTIGACIONES PREVIAS

2.1. INTRODUCCIÓN

Para fundamentar este trabajo, en este capítulo presentamos un resumen de las principales investigaciones relacionadas con nuestro objeto de investigación, que podemos clasificar en dos apartados.

En primer lugar realizamos una síntesis de las investigaciones sobre formación de profesores para enseñar probabilidad, aunque estas son todavía muy escasas. Si bien el campo de formación de profesores es uno de los que mayor atención ha recibido en las investigaciones en didáctica de la matemática, esta tendencia no se refleja en el caso específico de la probabilidad. A pesar de ello, se ha iniciado una línea de investigación consistente sobre la formación de profesores para enseñar estadística y probabilidad, impulsada por el ICMI Study 18, "Statistics Education in School Mathematics, Challenges for Teaching and Teacher Education" (Batanero, Burrill, Reading y Rossman, 2008; Batanero, Burrill y Reading, en prensa). Dentro de este bloque, analizamos las investigaciones realizadas sobre los aspectos afectivos que influyen en la labor del profesor (actitudes y creencias), los estudios de evaluación de los conocimientos probabilísticos de los profesores, y los relacionados con su conocimiento didáctico para enseñar probabilidad, haciendo también un pequeño resumen de algunos modelos específicos que describen los componentes del conocimiento profesional del profesor en el campo de la estadística.

El segundo bloque de nuestro estado de la cuestión está formado por investigaciones sobre la comprensión de la probabilidad condicional y temas relacionados con ella. Estas investigaciones son mucho más numerosas y se han realizado tanto en el campo de la Psicología, como en Educación Matemática. La síntesis que realizamos en la segunda parte del capítulo, extiende el artículo de Díaz y de la Fuente (2005), y el estado de la cuestión de Díaz (2007) y se amplía con otras investigaciones no tenidas en cuenta por estas autoras, así como los trabajos realizados posteriormente. En primer lugar revisamos los trabajos relacionados con la comprensión conceptual y procedimental, que son la mayoría, pues los trabajos sobre enseñanza del tema son bastante escasos. Incluimos también un apartado sobre experimentos de enseñanza del tema y otro sobre investigaciones centradas en el análisis de libros de texto y las conclusiones de este estado de la cuestión.

2.2. FORMACIÓN DE PROFESORES PARA ENSEÑAR PROBABILIDAD

2.2.1. INTRODUCCIÓN

Comenzamos el capítulo con el análisis de las investigaciones relacionadas con la formación de profesores para enseñar probabilidad. En los últimos años ha habido un gran número de investigaciones centradas en la formación inicial del profesor de matemáticas y su desarrollo profesional, como consecuencia de sus experiencias docentes (Batanero, Arteaga y Contreras, 2009). Toda esta investigación se recoge en libros y capítulos dedicados al tema (Brown y Borko, 1992; Thompson, 1992; Jaworski y Gellert, 2003; Llinares y Krainer, 2006; Ponte y Chapman, 2006; Hill, Sleep, Lewis y Ball, 2007; Franke y cols., 2007; Philipp, 2007; Sowder, 2007; Wood, 2008). También destacan en esta labor de recopilación el ICMI Study 15 (Even y Ball, 2008) y la revista *Journal of Mathematics Teacher Education*.

En España la investigación sobre el tema ha sido reforzada desde el Grupo de Investigación "Conocimiento y desarrollo profesional del profesor de matemáticas" de la Sociedad Española de Investigación en Educación Matemática (SEIEM), entre otros profesores Salvador Llinares, M. Victoria Sánchez, José Carrillo, Lorenzo Blanco y Pablo Flores. En este capítulo no trataremos de sintetizar todo este trabajo, sino sólo centrarnos en la parte del mismo relacionada con la formación de profesores para enseñar probabilidad. Esta formación está siendo objeto actual de debate, pues diferentes autores (Russell, 1990; Batanero, Godino y Roa, 2004; Franklin y Mewborn, 2006) coinciden en que muchos de los programas de formación en curso todavía no forman a los profesores adecuadamente para enseñar estadística y probabilidad, al no incluir un curso específico de didáctica del tema en su formación inicial.

La mayoría (aunque no todos) los profesores de educación secundaria y Bachillerato son licenciados en matemáticas, y estudiaron uno o más cursos de estadística. Pero, en general, la formación que tuvieron fue sólo teórica y no tienen experiencia con estudios de estadística aplicada, muestreo, o diseño de experimentos (Franklin y Mewborn, 2006). Aunque el enfoque que se recomienda hoy día para la enseñanza de la probabilidad es el frecuencial, y se sugiere que se organicen en la clase de matemáticas experimentos y simulaciones, pocos profesores tienen experiencia con este enfoque, que requiere un análisis cuidadoso de las situaciones y ser conscientes de la diferencia entre frecuencia relativa (que se refiere a los datos empíricos) y probabilidad (que es un objeto matemático teórico) (Chaput, Girard y Henry, 2008).

Por otro lado, son pocos los profesores que tuvieron un curso específico de didáctica de la estadística, pues, aunque muchos hicieron el Curso de Aptitud Pedagógica o están realizando el Master de Formación del Profesorado de Educación Secundaria, donde siguieron o siguen cursos de didáctica de la matemática, los principios generales que son válidos para la geometría, el álgebra u otras áreas de las matemáticas no siempre pueden ser aplicados a la estadística (Batanero, Godino y Roa, 2004). La situación de los maestros de primaria es que pocos han recibido una formación adecuada en el campo de la probabilidad, pues hasta ahora no se incluía apenas enseñanza de la probabilidad en la educación primaria. Por otro lado, en pocos cursos sobre didáctica de la matemática se incluyen los conocimientos didácticos que necesitan para enseñar probabilidad (Franklin y Mewborn, 2006).

Algunos autores describen ejemplos y experiencias de cursos específicamente dirigidos a capacitar a profesores de diferentes países, algunas de ellas basadas en modelos teóricos de cómo debería ser esta formación (por ejemplo, Kvatinsky y Even, 2002; Batanero, Godino y Roa, 2004). La evaluación del éxito de estos cursos está, en general, basada en muestras pequeñas o datos subjetivos, aunque proporcionan ejemplos e ideas para otros formadores de profesores y por ello es necesario ampliar la base empírica de tales estudios. A continuación, analizamos las investigaciones llevadas a cabo sobre los diferentes componentes que se necesitan en la formación de profesores para la enseñanza de la probabilidad.

2.2.2. ACTITUDES Y CREENCIAS

Un primer punto objeto de evaluación y formación son las creencias y actitudes de los profesores de matemáticas, que forman parte del dominio afectivo de las matemáticas. Mientras que las emociones son transitorias, las actitudes son sentimientos intensos, relativamente estables en el tiempo, que se desarrollan por repetición de respuestas emocionales y se automatizan con el tiempo. Las creencias, por su parte, tienen un carácter más cognitivo; son las concepciones que se tienen sobre una materia, sobre uno mismo como estudiante o sobre el contexto social en el que se realiza el aprendizaje (Estrada, 2007). El aspecto afectivo es muy importante en el caso de los profesores, ya que sus creencias incorrectas o actitudes negativas podrían condicionar la enseñanza y repercutir en las futuras actitudes de sus alumnos (Estrada y Batanero, 2008). Será por tanto necesaria una labor de motivación si queremos que la enseñanza de la estadística y probabilidad sea una realidad y no simplemente un deseo expresado en las orientaciones curriculares (Estrada, Batanero y Fortuny, 2004).

Son pocas las investigaciones específicamente centradas en las creencias o actitudes de los profesores hacia la probabilidad. Uno de los primeros autores que se refiere a este tema es Steinbring (1990), quien sugirió que los formadores de profesores tienen que hacer frente al problema de concienciar a los profesores de la naturaleza particular de lo estocástico y ver que es diferente de la matemática tradicional enseñada en la escuela. Al mismo tiempo, se tienen que producir las condiciones necesarias para la integración del pensamiento estocástico en la escuela, y ampliar los puntos de vista de los profesores sobre qué es la matemática.

El conocimiento estocástico es más complejo y sistémico y se basa mucho más en las actividades interpretativas que otras áreas de las matemáticas. Así, los conceptos están interrelacionados con el contexto en el que se aplican y están sujetos a controversias. Por ejemplo, para entender lo que es una variable aleatoria, una persona debe asumir un modelo para la aleatoriedad, un concepto ligado a muchas interpretaciones filosóficas (Steinbring, 1990).

Serradó, Azcárate y Cardeñoso (2006) realizaron una investigación utilizando cuestionarios y entrevistas para comprender las razones por las cuáles los profesores omiten la enseñanza de la probabilidad en la escuela y las fuentes de información que pueden influenciar esta decisión. Para ello realizaron un estudio de casos de cinco profesores, presentando los resultados obtenidos para dos de ellos, ninguno de los cuáles enseñaba probabilidad. El primero (A) era licenciado en matemáticas con especialidad de estadística y 5 años de experiencia; el segundo (B), también matemático, pero no especialista en estadística con 2 años de experiencia. Las fuentes de información consideradas por el profesor A, para llevar a cabo innovaciones, fueron el contacto con colegas que favorecían dichas innovaciones, y su omisión de la enseñanza de la probabilidad era debida a su creencia de que la probabilidad no tenía suficiente consistencia educativa en la enseñanza obligatoria ni garantizaba ningún propósito práctico para los estudiantes. El profesor A apoyaba estos argumentos en una visión formal de la matemática que había adquirido tanto en su formación como profesor, como en los libros de texto de secundaria. El profesor B por su parte rechazaba enseñar probabilidades pensando en las dificultades que el estudiante podría tener con el tema y suponiendo que la metodología usual para otros temas no sería útil para esta materia. La falta de apoyo de sus compañeros y la falta de fuente de información suponían otros obstáculos para incorporar la probabilidad en la enseñanza.

Además de las razones inferidas en la investigación de Serradó, Azcárate y Cardeñoso, los docentes también tienen teorías subjetivas sobre cuáles debieran ser los contenidos y los objetivos educativos relacionados con estos contenidos que les hacen retrasar o cambiar la enseñanza, respecto a lo indicado en las orientaciones curriculares. Eichler (2008) describe una investigación sobre las creencias de los profesores respecto a la enseñanza de la estadística, mostrando la influencia de estas creencias sobre su práctica docente y cómo esta práctica influye en los conocimientos que adquieren sus alumnos. Este autor diferencia entre el currículo oficial (marcado por las directrices curriculares), el pretendido por el profesor, el implementado en el aula y el aprendido por los estudiantes. Clasifica a los profesores en su investigación respecto a su visión estática frente a una visión más dinámica de las matemáticas y la orientación hacia la matemática formal versus aplicaciones matemáticas. Debido a estas creencias muestra que la enseñanza recibida por los estudiantes respecto a un mismo nivel educativo y contenidos, podrían diferir considerablemente.

Todos estos resultados alertan de la importancia de desarrollar actitudes y creencias positivas de los profesores hacia la probabilidad, tanto dentro de la matemática, como para la formación de sus estudiantes (Estrada y Batanero, 2008).

2.2.3. CONOCIMIENTOS PROBABILÍSTICOS

Ball, Lubienski y Mewborn (2001) indican que muchas actividades usuales que realiza el profesor en la clase de matemáticas dependen de sus conocimientos matemáticos. Sería por tanto necesario un sólido conocimiento de la probabilidad por parte de los profesores, pero la investigación que resumimos a continuación es muy escasa y se centra, en su mayor parte en profesores en formación. Los estudios de evaluación se han realizado mediante cuestionarios similares a los empleados en investigaciones con estudiantes, están centradas en el conocimiento común del contenido y tratan de determinar si el nivel de conocimiento de los futuros profesores es suficiente para resolver las tareas que tendrán que proponer a sus futuros alumnos.

Algunas de dichas investigaciones señalan la existencia de concepciones erróneas y dificultades sobre probabilidad en los futuros profesores. Por ejemplo, Azcárate (1995) en un estudio realizado con 57 futuros profesores de Educación Primaria, propuso una serie de enunciados en que se describían diferente fenómenos, preguntando a los participantes si consideraban que eran o no aleatorios, justificando su respuesta. La autora encontró que muy pocos sujetos en su estudio mostraban una idea clara sobre las características de los fenómenos aleatorios. Los participantes razonaron, en su mayoría, desde presupuestos causales; tuvieron una fuerte influencia de los aspectos contextuales y minusvaloraron el posible estudio matemático de los fenómenos aleatorios. Se detectó también falta de esquemas combinatorios y ausencia de instrumentos elementales para la asignación de probabilidades, cuantificando las expectativas de ocurrencia de un suceso desde criterios personales. Respecto a la idea de juego equitativo, Azcárate propone tres ítems basados en el lanzamiento de dos dados, preguntando si sería justo apostar a producto par, suma par y suma 5 o 6. Los participantes mostraron mucha dificultad para diferenciar los juegos equitativos de los no equitativos y basan su argumento sobre la equitatividad de un juego en la equiprobabilidad de los resultados, reglas aritméticas o argumentación combinatoria.

Begg y Edward (1999) proporcionaron un cuestionario y realizaron entrevistas a 22 profesores australianos en activo de escuela primaria, que apenas habían tenido enseñanza formal del tema durante su formación inicial. Entre otros aspectos evaluaron sus conocimientos sobre probabilidad con tres preguntas usando contextos bien conocidos de la vida cotidiana. La primera pregunta fue sobre el juego de la lotería, la segunda una actividad escolar típica sobre lanzamiento de monedas. La tercera pregunta, algo más formal fue una pregunta sobre la heurística de la representatividad en el contexto de nacimientos de niños en dos hospitales. Los autores llegaron a la conclusión de que dichos profesores tenían una escasa comprensión de la probabilidad. Sólo alrededor de dos tercios del profesorado de su muestra comprendía correctamente las ideas de aleatoriedad y de sucesos equiprobables y muy pocos aplicaron correctamente el concepto de independencia. El análisis de las justificaciones de las respuestas sugiere que los profesores empleaban la heurística de la representatividad (Tversky y Kahneman, 1982b) esperando que una muestra fuese semejante a la población y el outcome approach descrito por Konold (1989), es decir, interpretando un enunciado probabilístico en forma no probabilística. Los profesores también hicieron un uso excesivo de sus teorías previas al resolver los problemas.

Watson (2001) realizó una encuesta a 15 profesores de primaria (enseñanza elemental) y a 28 profesores de secundaria de Australia, utilizando las respuestas de los profesores para buscar patrones y para describir características generales de éstos. Respecto a los conocimientos de probabilidad de los docentes, informó de que los profesores de secundaria tenían un nivel de seguridad significativamente más alto que los de primaria a la hora de enseñar conceptos básicos de probabilidad. Estos hallazgos han sido corroborados en investigaciones con profesores de primaria (Pereira-Mendoza, 2002).

Ortiz, Nordin, Batanero, Serrano y Rodríguez (2006) evaluaron la capacidad de los futuros profesores de educación primaria para comparar probabilidades, utilizando tareas sencillas tomadas de un cuestionario de probabilidad que había sido utilizado previamente con niños de 11 a 14 años. Para ello analizaron las respuestas de 102 estudiantes de magisterio a siete problemas, estudiando los porcentajes de respuestas correctas y los argumentos proporcionados por los alumnos, y comparando sus resultados con los obtenidos por los alumnos participantes en la investigación de Cañizares (1997). Estos autores observaron una mejora del número de respuestas correctas respecto a los niños de la investigación de Cañizares en todos los problemas, aunque los porcentajes de errores son muy altos.

Por ejemplo, en un problema en que se pedía comparar la probabilidad de obtener una bola negra en dos urnas con diferente composición de bolas blancas y negras, donde la falta de proporcionalidad obligaba a usar la regla de Laplace, obtuvieron un 75% de errores en los futuros profesores de primaria. En otro problema similar en que se añade un factor subjetivo irrelevante (uno de los jugadores es más inteligente, por tanto podría tener más posibilidades de sacar la bola negra) se produce un aumento de la dificultad, debido al distractor de tipo subjetivo (70% de errores). Otro problema de comparación de probabilidades donde la composición de las dos urnas era proporcional, obtuvo 57% de errores. Aunque los futuros profesores de educación primaria parecían ser conscientes de esta falta de equivalencia entre las dos urnas, sin embargo se decantaron por la que tenía una mayor cantidad absoluta de casos favorables o por aquella en que la diferencia entre los casos favorables y desfavorables era mayor. Por otro lado en algunos problemas algunos futuros profesores obtuvieron respuestas correctas con un razonamiento inadecuado.

Batanero, Cañizares y Godino (2005) analizaron los resultados de la evaluación inicial en una muestra de 132 profesores en formación, utilizando un cuestionario con varios ítems que evaluaban la presencia de sesgos en el razonamiento probabilístico de los profesores. Los resultados de la evaluación mostraron que el 60% de la muestra razonaba de acuerdo a la heurística de la representatividad (Tversky y Kahneman, 1982b) esperando que una muestra fuese semejante a la población. Otro 60% de participantes mostró el sesgo de equiprobabilidad descrito en los experimentos de Lecoutre (1992), Lecoutre y Durand (1988), que consiste en la creencia de los sujetos en la equiprobabilidad de todos los sucesos asociados a cualquier experimento aleatorio, incluso en aquellos en que no es aplicable el principio de indiferencia o donde no hay una simetría física. Finalmente, el 23% de los profesores mostró el "outcome approach" (enfoque en los resultados) y el outcome approach descrito por Konold (1989), es decir, interpretando un enunciado probabilístico en forma no probabilística. Los autores realizaron actividades de simulación con dispositivos manipulativos, tablas de números aleatorios y el software Statgraphics, volviendo a evaluar a los participantes al finalizar la experiencia. Se observó una mejora general de las concepciones de los participantes, lo que sugiere el interés de la simulación como recurso didáctico.

En la investigación de Sánchez (1996) con 88 profesores de secundaria de Matemáticas, se observan muchas dificultades en la comprensión del concepto de independencia. El autor pregunta a los sujetos si, al sacar una carta de una baraja ordinaria, los sucesos "ser reina" y "ser una carta de corazones" son o no independientes. Sólo 39 profesores dieron una respuesta correcta al problema planteado y sólo 4 utilizaron la regla del producto y llegaron a la solución correcta. Las concepciones incorrectas de estos profesores se discuten con más detalle en el apartado 2 de este capítulo.

Carnell (1997) realizó un estudio sobre la comprensión de la probabilidad condicional en 13 profesores de enseñanza media en prácticas, donde todos demostraron tener concepciones erróneas relacionadas con la comprensión de la probabilidad condicional expuestas por Falk (1986), como por ejemplo, dificultad para definir el elemento condicionante a partir del enunciado verbal de un problema, dificultad en resolver un problema cuando el orden temporal del elemento condicionante no es el esperado, así como confundir la condicionalidad con la causalidad. Todos estos tipos de errores han aparecido también en otros estudiantes y se describen con mayor detalle en las siguientes secciones.

Otro concepto importante en estadística es el de variación, estudiado por Wild y Pfannkuch (1999), quienes la consideran el núcleo de su modelo sobre razonamiento estadístico. Aunque este concepto se asocia preferentemente a la estadística y no a la probabilidad, sin embargo, la comprensión de la idea de aleatoriedad y la de variación están ligadas, por lo que lo incluimos en este apartado. La investigación de Borim y Coutinho (2008) con nueve profesores brasileños de educación secundaria en activo muestra que el razonamiento predominante de dichos profesores sobre la variación era verbal, lo que les impedía enseñar a sus estudiantes el significado de medidas tales como la desviación típica, limitándose a la enseñanza de algoritmos. Ninguno fue capaz de integrar el razonamiento sobre la media, desviación respecto a la media, intervalo de k desviaciones típicas desde la media y estimación de la frecuencia en el intervalo, es decir, no comprendieron el significado de la desviación típica como medida de dispersión. La variación se estudia también en probabilidad (variabilidad aleatoria), de modo que las dificultades encontradas con la variación en probabilidad se presentarán previsiblemente también en el estudio de la probabilidad.

Respecto a la idea de distribución, otra idea central del razonamiento estadístico, que aparece tanto en el estudio de la estadística (distribución de datos) como de probabilidad (distribución de probabilidad), Canadá (2008) analizó el razonamiento de 55 estudiantes de secundaria y 58 futuros profesores de secundaria cuando comparan conjuntos de datos con la misma media y diferente dispersión. Aunque el grupo de futuros profesores tuvo mejor rendimiento que el de los estudiantes, todavía un 35% de los futuros profesores pensaba que dos conjuntos de datos con la misma media eran iguales, aunque la dispersión fuese muy diferente.

Estrada y Díaz (2007) realizaron un estudio exploratorio con 65 futuros profesores de educación primaria de la Universidad de Lleida, que estaban siguiendo un curso optativo de didáctica de la estadística. Les plantearon un problema con tres preguntas: (a) cálculo de una probabilidad simple, (b) cálculo de una probabilidad condicional y (c) cálculo de una probabilidad conjunta, dando los datos en una tabla de doble entrada. Analizan las respuestas abiertas y encuentran una gran proporción de errores, sobre todo en las preguntas sobre probabilidad condicional y conjunta. Las autoras explican sus resultados en términos de conflictos semióticos.

Por su parte Sáenz (2007) evalúa el conocimiento funcional de las matemáticas en una muestra de 140 futuros profesores de educación primaria, utilizando algunos ítems de las pruebas PISA. Encuentra un porcentaje medio de respuestas correctas de un 64%, indicando que los estudiantes de magisterio no superan significativamente los resultados de alumnos de 15 años. El rendimiento es menor a medida que aumenta la complejidad del ítem y sólo un 11% de participantes alcanza los ítems de mayor dificultad. El autor indica que es precisamente el área de probabilidad y estadística donde se obtienen los peores resultados.

En resumen, las escasas investigaciones sobre conocimientos probabilísticos de futuros profesores nos indican que este conocimiento es a veces escaso, sobre todo en los futuros profesores y más aún en los futuros profesores de educación primaria. Puesto que pocos de estos trabajos se han centrado específicamente en la probabilidad condicional, llevaremos a cabo estudios de evaluación de estos conocimientos, tanto con futuros profesores de educación primaria, como de educación secundaria. Para el estudio de evaluación con futuros profesores de educación primaria, añadiremos además algunas cuestiones para la evaluación del conocimiento didáctico de los profesores.

2.2.4. MODELOS SOBRE EL CONOCIMIENTO DEL PROFESOR

Los profesores también necesitan formación en el conocimiento didáctico relacionado con la educación estadística, al que no pueden transferirse algunos principios generales válidos para otras ramas de las matemáticas (Batanero, Godino y Roa, 2004). Si queremos impulsar el desarrollo de este conocimiento y su análisis requiere un modelo de sus componentes de dicho conocimiento. A continuación analizamos algunos modelos en didáctica de la matemática y la estadística sobre esta problemática.

Modelos en educación matemática

El interés por el conocimiento didáctico específico para enseñar matemáticas, es iniciado por Shulman (1986). Además del conocimiento matemático, reconoce la necesidad de que los profesores aprendan otros contenidos, tales como: (a) conocimientos pedagógicos generales, como los relacionados con la gestión y organización de la clase; (b) conocimiento del currículo, materiales de enseñanza y directrices curriculares; (c) conocimiento didáctico del contenido (o conocimiento pedagógico del contenido); (d) conocimiento de los alumnos y de sus características; (e) conocimiento del contexto educativo; y (f) conocimiento de los objetivos, valores y fines educativos y sus fundamentos filosóficos e históricos (Shulman, 1987).

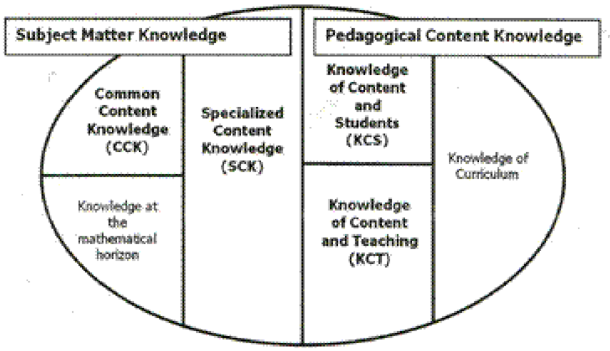
Shulman (1986) denomina *conocimiento pedagógico del contenido* al que va más allá del simple conocimiento de la materia y se refiere a la enseñanza de la misma, incluyendo sus diferentes representaciones y la forma de hacerla comprensible a los estudiantes. Sería "*la forma particular del conocimiento del contenido que incorpora el aspecto del contenido que guarda más relación con la enseñanza*" o bien "*esa amalgama especial de contenido y pedagogía que es el campo propio de los profesores, su forma especial de comprensión profesional*" (Shulman, 1986, p. 8-9).

Diversos autores se refieren a este conocimiento con diferentes nombres. Ball, Lubienski y Mewborn (2001) hablan del conocimiento matemático para la enseñanza, que se describe en Hill, Ball, y Schilling (2008) como "*el conocimiento matemático que utiliza el profesor en el aula para producir instrucción y crecimiento en el alumno*" (p. 374) y lo dividen en dos grandes categorías (ver Figura 2.2.1): Conocimiento del contenido y conocimiento pedagógico del contenido.

Dentro del conocimiento del contenido matemático distinguen entre "conocimiento común del contenido" (CCK), "conocimiento especializado del contenido" (SCK), y "conocimiento en el horizonte matemático". Mientras el "conocimiento común del contenido" es el puesto en juego para resolver problemas matemáticos por cualquier persona, el "conocimiento especializado del contenido" incluye aspectos que no necesariamente tiene una persona ordinaria, por ejemplo, elegir una secuencia de enseñanza o identificar las ideas matemáticas trabajadas en un problema. El "conocimiento en el horizonte matemático" aporta perspectiva a los profesores para su trabajo, por ejemplo, el conocimiento de la relación que tiene las matemáticas con otras materias, o la historia de las matemáticas.

Para el conocimiento pedagógico del contenido Hill, Ball, y Schilling (2008) proponen tener en cuenta, el conocimiento del contenido y los estudiantes (KCS), el conocimiento del contenido y la enseñanza (KCT), y el conocimiento del currículo. El conocimiento del contenido y los estudiantes (KCS) "es el conocimiento de cómo los estudiantes piensan, saben, o aprenden este contenido particular" (p. 375). Incluye el conocimiento de los errores y dificultades comunes, las concepciones erróneas, las estrategias utilizadas, el ser capaz de valorar la comprensión del alumno y saber cómo evoluciona su razonamiento matemático. Respecto al conocimiento del contenido y la enseñanza (KCT) resulta de la integración del contenido matemático con el conocimiento de la enseñanza de dicho contenido. Incluye saber construir, a partir del razonamiento de los estudiantes y las estrategias utilizadas por ellos, procesos pertinentes para tratar y corregir sus errores y concepciones erróneas. El conocimiento del currículo se refiere a las directrices curriculares, orientaciones, fines y motivaciones de las mismas, materiales curriculares y secuenciación del tema en los diferentes ciclos formativos. Este mismo modelo se puede tener en cuenta para el caso de la probabilidad considerando las características específicas de la materia.

Figura 2.2.1. Hill et al., 2007 (reproducida de Delaney et al., 2008, p. 174)



Modelos en educación estadística

Una de las primeras referencias al estudio del conocimiento del profesor para explicar estadística se hace en Godino, Batanero y Flores (1999), quienes describen los siguientes componentes básicos en la formación de profesores:

- *La reflexión epistemológica* sobre la naturaleza del conocimiento estocástico, su desarrollo y evolución.
- *Análisis de las transformaciones del conocimiento* para adaptarlos a los distintos niveles de enseñanza, incluyendo el nivel y forma particular en que un determinado concepto podría ser enseñado a una persona particular.
- *Estudio de las dificultades, errores y obstáculos* de los alumnos en el aprendizaje y sus estrategias en la resolución de problemas, actitudes, sentimientos y emociones.
- *Análisis del currículo, situaciones didácticas, metodología de enseñanza para temas específicos y recursos didácticos específicos:* libros, ejemplos, tareas, software y todo lo que contribuye a la enseñanza.

Estos componentes son ampliados en Batanero, Godino y Roa (2004), quienes incluyen las siguientes facetas en la formación de profesores para enseñar estadística, sugiriendo también la necesidad de utilizar situaciones de tipo constructivista en la formación de los profesores:

- Reflexión epistemológica sobre el significado de los conceptos que se enseñan (por ejemplo, sobre los significados diferenciados de la probabilidad). Esta reflexión incluiría conocimiento de tipo histórico, filosófico y cultural, así como las relaciones con otros dominios de las ciencias.
- Experiencias para adaptar el conocimiento estadístico a diferentes niveles de enseñanza y a la capacidad de los estudiantes, organizando e implementando proyectos estadísticos, simulaciones y gráficos, no sólo como ayudas metodológicas sino como formas útiles de aprender y comprender la estadística.
- Capacidad crítica para analizar libros de texto y materiales curriculares.
- Predicción de las dificultades, errores, estrategias y obstáculos de los estudiantes al resolver problemas, para así desarrollar y analizar ítems de evaluación e interpretar las respuestas de los estudiantes a los mismos.
- Experiencia con buenos ejemplos de situaciones didácticas, materiales y recursos.

Burgess (2008) parte de los modelos en educación matemática y propone otro que puede ser usado para estudiar el conocimiento del profesor y el conocimiento estadístico para la enseñanza. Para ello cruza los cuatro componentes definidos por Ball, Thames y Phelps (2005) con las categorías que definen los modos fundamentales de razonamiento estadístico de Wild y Pfannkuch (1999): reconocer la necesidad de datos, transnumeración (cambio de representaciones de esos datos para conseguir extraer información nueva de ellos), consideración de la variación (reconocer la variación en los datos, describir patrones acerca de la variación y tratar relacionar los datos con el contexto), razonamiento con modelos estadísticos (desde los simples, como tablas y gráficos, a los complejos) e integración de la estadística y el contexto (ver Figura 2.2.2). Además añade otros elementos del modelo de razonamiento estadístico de Wild y Pfannkuch: ciclo de investigación (problema, plan, datos, análisis y conclusión), el ciclo interrogativo (generar, buscar, interpretar, criticar y juzgar) y las disposiciones, tales como escepticismo e imaginación.

Figura 2.2.2. Modelo de conocimiento profesional de Burgess (2008, p. 3)

			Statistical knowledge for teaching			
			Content knowledge		Pedagogical content knowledge	
			Common knowledge of content (ckc)	Specialised knowledge of content (skc)	Knowledge of content and students (kcs)	Knowledge of content and teaching (kct)
Statistical thinking in empirical enquiry	Types of thinking	Need for data				
		Transnumeration				
		Variation				
		Reasoning with models				
		Integration of statistical and contextual				
	Investigative cycle					
	Interrogative cycle					
	Dispositions					

Debido a la relación entre estadística y tecnología Lee y Hollebrands (2008) presentan un marco para describir el conocimiento profesional para enseñar estadística con apoyo de la tecnología. Consideran cuatro componentes: (a) Concepciones de qué significa enseñar un tema particular integrando la tecnología en el proceso de aprendizaje; (b) Conocimiento de las estrategias de enseñanza y las representaciones para enseñar temas particulares con la tecnología; (c) Conocimiento sobre la comprensión, razonamiento y aprendizaje de los estudiantes con la tecnología y (d) Conocimiento del currículo y materiales curriculares que integran la tecnología en el aprendizaje y ponen el énfasis en el razonamiento estadístico.

Garfield y Ben-Zvi (2008) proponen otro modelo de conocimiento del profesor para enseñar estadística que incluye seis competencias que se han de desarrollar en la formación de profesores: (a) ideas estadísticas fundamentales; (b) uso de datos reales; (c) uso de actividades para el aula; integración de las herramientas tecnológicas; (d) implementación del discurso en el aula y (e) uso de métodos alternativos de evaluación. Estos componentes, sin embargo no se relacionan con los descritos en la literatura de formación de profesores; más bien serían aspectos a tener en cuenta en la formación estadística, tanto de alumnos como de profesores.

Un modelo integrador

Godino y colaboradores (Godino, Batanero, Roa y Wilhelmi, 2008) elaboran un modelo general para el conocimiento profesional del profesor de matemáticas con seis dimensiones: epistemológica (conocimiento del contenido matemático, las adaptaciones para ser enseñado, su desarrollo histórico y problemas filosóficos asociados), cognitiva (desarrollo del razonamiento en los estudiantes, dificultades de aprendizaje), afectiva (actitudes, afectos, emociones de los estudiantes respecto al tema), interacional (relaciones profesor – estudiantes, y entre los propios estudiantes y organización del discurso en la clase), mediacional (uso de recursos tecnológicos y el tiempo requerido para el estudio), ecológica (aspectos curriculares, socioculturales y relaciones con el entorno). Estos componentes, se relacionan con el concepto de idoneidad didáctica, concepto desarrollado también dentro del Enfoque Ontosemiótico y descrito en el capítulo 2 del presente trabajo. Si se quiere conseguir una idoneidad didáctica alta en los procesos de enseñanza y aprendizaje de las matemáticas, será necesaria la capacidad del profesor para valorar cada uno de los componentes de la idoneidad, lo que supone la adquisición de conocimientos específicos relacionados con cada uno de estos componentes.

Godino (2009) refina el modelo de niveles y facetas del conocimiento matemático didáctico del profesor que engloba los conocimientos citados anteriormente y propone, asimismo una guía para la formulación de cuestiones de evaluación de dicho conocimiento. El autor propone tener en cuenta las siguientes facetas para analizar los procesos de instrucción matemática:

- *Epistémica*: Conocimientos matemáticos relativos al contexto institucional en que se realiza el proceso de estudio y la distribución en el tiempo de los diversos componentes del contenido (problemas, lenguajes, procedimientos, definiciones, propiedades, argumentos).

- *Cognitiva*: Conocimientos personales de los estudiantes y progresión de los aprendizajes.
- *Afectiva*: Estados afectivos (actitudes, emociones, creencias, valores) de cada alumno con relación a los objetos matemáticos y al proceso de estudio seguido.
- *Mediacional*: Recursos tecnológicos y asignación del tiempo a las distintas acciones y procesos.
- *Interaccional*: Patrones de interacción entre el profesor y los estudiantes y su secuenciación orientada a la fijación y negociación de significados.
- *Ecológica*: Sistema de relaciones con el entorno social, político, económico, que soporta y condiciona el proceso de estudio.

La importancia dada al componente epistémico en el modelo de Godino (2009) hace que sea aplicable al caso particular de la estadística. Puesto que la estadística tiene sus propios problemas, lenguaje, conceptos, propiedades, argumentos y procedimientos, el profesor requiere un conocimiento específico de los mismos y de su interacción con el resto de componentes del modelo. Es decir, se requiere una didáctica específica de la estadística. En consonancia con el resto del marco teórico descrito en el capítulo 2 y puesto que el modelo reconoce la especificidad de la estadística, a través de la componente epistémica, lo utilizaremos en este trabajo. Asimismo, usaremos el modelo de Hill, Ball, y Schilling (2008) en los capítulos 3, 4 y 6 en la evaluación de los conocimientos de los profesores.

Godino (2009) también propone diferentes niveles de análisis didáctico, que pueden usarse, tanto en la formación de profesores, como en la evaluación de su conocimiento, o incluso que el propio profesor puede utilizar para mejorar su práctica docente:

1. *Prácticas matemáticas y didácticas*. En este primer nivel se analizan las acciones realizadas para resolver las tareas matemáticas propuestas para contextualizar los contenidos y promover el aprendizaje. También se describen las líneas generales de actuación del docente y discentes. Nosotros no usaremos este nivel.
2. *Configuraciones de objetos y procesos.* Se trata de describir los objetos y procesos matemáticos que intervienen en la realización de las prácticas, para mostrar la

complejidad de objetos y significados de las prácticas matemáticas. También sirve para explicar los conflictos en su realización y la progresión del aprendizaje. Nosotros usaremos este tipo de análisis en el estudio de las respuestas a diferentes tareas de evaluación (Estudios 1, 2 y 5) y de posibles soluciones a tareas planteadas en el trabajo con los recursos (Estudios 2 y 3).

3. *Normas y metanormas.* Identificación de las normas que condicionan y hacen posible el proceso de estudio, y que afectan a cada faceta y sus interacciones. Este nivel de análisis no se usa en nuestro trabajo.
4. *Idoneidad didáctica.* Identificación de potenciales mejoras del proceso de estudio que incrementen la idoneidad didáctica. En nuestro trabajo analizaremos la idoneidad didáctica de recursos en Internet (Estudio 3) y paradojas sobre probabilidad condicional e independencia (Estudio 4).

Además, para evaluar algunos componentes del conocimiento didáctico usaremos la metodología propuesta por Godino (2009) que consiste en dos pasos:

1. Elegir una tarea matemática cuya solución ponga en juego los principales aspectos del contenido, o de las competencias a desarrollar;
2. Formular consignas que cubran algunas facetas y niveles de análisis didáctico del modelo propuesto por el autor. Dicha consigna consistiría en: (a) resolver el problema (para evaluar el conocimiento común del contenido); (b) identificar los objetos y procesos matemáticos puesto en juego en la solución (para el conocimiento especializado del contenido); (c) describir los razonamientos que los alumnos han desarrollado al resolver la tarea propuesta o los principales conflictos en dicha solución (para el conocimiento del contenido y los estudiantes), y (d) análisis de las diferentes categorías de idoneidad didáctica (para evaluar los distintos componentes del conocimiento didáctico del profesor).

En nuestro trabajo utilizaremos consignas de tipo (a), (b) y (c) para evaluar el conocimiento común o especializado del contenido, el conocimiento del contenido y los estudiantes y el conocimiento del contenido y la enseñanza. En lo que sigue analizamos algunas investigaciones que han tratado de evaluar algunos de estos componentes.

2.2.5. CONOCIMIENTO DE LA PROBABILIDAD Y LA ENSEÑANZA

Los diferentes componentes del conocimiento didáctico se adquirirán preferentemente durante el ejercicio de la docencia (Ponte, 2008), pero, al tratarse de un tema nuevo en los currículos, ha habido, hasta el momento poca oportunidad para un desarrollo profesional de los profesores de educación primaria en probabilidad. Consecuentemente, las investigaciones relacionadas con el conocimiento didáctico de los profesores para enseñar probabilidad sugieren que este conocimiento es escaso.

Respecto al conocimiento especializado del contenido, algunas investigaciones examinan a profesores en ejercicio durante su enseñanza, observándolos a lo largo de un periodo de tiempo y deduciendo su conocimiento a partir de esta observación, complementada con entrevistas u otros medios. Un primer punto investigado es la capacidad de los profesores para reconocer qué conceptos pueden ser estudiados a partir de un recurso o problema dado. En la investigación de Chick y Pierce (2008) los profesores participantes no hicieron un uso adecuado de los datos y proyectos al planificar sus lecciones, pues fallaron en sacar a la luz los conceptos latentes, a pesar de la riqueza de conceptos de la situación didáctica planteada. Por el contrario se limitaron a pedir cálculos o nuevos gráficos, con pocas actividades de interpretación.

En la enseñanza de la probabilidad se recomienda actualmente el uso de investigaciones y simulaciones, de modo que se ayude a los estudiantes a mejorar su razonamiento probabilístico. Sin embargo son pocos los profesores que tienen experiencia previa con experimentos y simulaciones en probabilidad. Así, en el estudio de Stohl (2005) los profesores observados fallaron al implementar el enfoque experimental en la enseñanza de la probabilidad, porque las tareas que proponían a los estudiantes sólo utilizaban muestras pequeñas. Por este motivo, los estudiantes de estos profesores no pudieron apreciar la convergencia o el efecto del tamaño de la muestra sobre la misma, es decir no llegaron al punto central del enfoque frecuencial de la probabilidad. En su investigación, Watson (2001) encuentra pocos profesores de secundaria que usaran actividades basadas en la simulación y toma de muestras para reforzar la enseñanza de la probabilidad. Además, aunque los profesores de primaria utilizaban lecciones basadas en dichas actividades, no parecía existir un enfoque coherente hacia el estudio de los conceptos estadísticos.

Haller (1997) realizó una investigación con un grupo de cuatro profesores de educación secundaria que asistían a un curso de desarrollo profesional. Los profesores con un conocimiento del contenido más bajo, cometían errores conceptuales y de interpretación de las respuestas de sus alumnos durante sus lecciones. Estos profesores también dependían en gran medida de los libros de texto y no aprovechaban el contexto de la probabilidad para desarrollar relaciones entre las fracciones, los decimales y los porcentajes. Sin embargo, los profesores con un mayor conocimiento del contenido no cometían errores matemáticos, relacionaban los decimales, las fracciones y los porcentajes y añadían preguntas y actividades suplementarias, independientes del libro de texto. La experiencia docente no parecía tener tanta repercusión en la enseñanza de la probabilidad por parte de los profesores en esta investigación.

López (2006) analiza la forma en que los profesores diseñan y llevan a cabo unidades didácticas para la enseñanza de la probabilidad en la escuela primaria, y muestra la gran dificultad de estos profesores al enfrentarse a conceptos nuevos para ellos. En un estudio con profesores de enseñanza elemental en prácticas, Dugdale (2001) utilizó la simulación por ordenador para revelar el conocimiento didáctico. Observó que el uso de software permitió a los profesores diseñar un par de dados, con igual probabilidad de obtener pares que impares al multiplicar las cifras en cada tirada, simular un gran número de tiradas, calcular las frecuencias relativas, convenciéndose así de que el juego creado era representativo. Así mismo, destacó que los profesores en prácticas podían usar este software como una herramienta para promover el debate y la comprensión de la probabilidad desde un punto de vista que normalmente no permite un número limitado de pruebas con un dado físico, y que no se conformaban con observar las frecuencias relativas generadas por la simulación por ordenador, sino que pasaban a razonar sobre las probabilidades teóricas para así verificar los resultados.

Por otro lado, en algunos casos, los libros de texto y materiales curriculares preparados son insuficientes como soporte para el profesor. Ello es debido a que presentan una visión muy parcial de los conceptos (por ejemplo, sólo la aproximación clásica a la probabilidad o la inferencia). En otros casos las aplicaciones se limitan a juegos de azar, o no se basan en datos tomados de aplicaciones reales. También aparecen en ocasiones definiciones incorrectas o incompletas de los conceptos (Cardeñoso, Azcárate y Serradó, 2005).

2.2.6. CONOCIMIENTO DE LA PROBABILIDAD Y LOS ESTUDIANTES

Otras investigaciones se centran en el conocimiento del contenido y los estudiantes. Stohl (2005), examina cómo treinta y cinco profesores de educación secundaria interpretaban las interacciones de los alumnos con una herramienta de simulación. Las interpretaciones eran comparadas con un análisis sobre cómo los alumnos trabajaban con la herramienta de simulación, para así determinar las semejanzas y diferencias con las interpretaciones hechas por los profesores. Los resultados preliminares del análisis sobre el trabajo de los alumnos hecho por los profesores indicaban que estos últimos, a menudo, se centraban en criticar la ausencia de ideas formales sobre probabilidad pasando por alto detalles sobre las acciones y el lenguaje de los alumnos que indicaban el desarrollo de ideas probabilísticas. Como consecuencia, parece que los profesores no comprendían que el desarrollo de ideas sobre probabilidad es un proceso complejo y difícil de valorar.

Watson (2001) examinó el conocimiento de los profesores sobre las dificultades de sus alumnos con la probabilidad y la estadística. Cuando se les preguntó por las dificultades de los alumnos, sólo dos profesores de primaria que participaron en el estudio mencionaron haber encontrado dificultades, mientras que trece profesores de secundaria indicaron dificultades en aspectos procedimentales (calcular probabilidades, permutaciones, diagramas de árbol) o conceptuales (la probabilidad teórica, la inferencia o la probabilidad condicional). Aunque estos datos sugieren que los profesores de este estudio eran capaces de identificar los problemas de los alumnos, también muestran se centraban principalmente en aspectos de procedimiento y su enseñanza se basaba en un enfoque computacional. La autora también mostró que algunos profesores tenían una apreciación muy limitada de cómo sus alumnos podrían contestar preguntas tales como encontrar un error en un gráfico o identificar un error en un problema relacionado con el muestreo de una población.

2.2.7. EVALUACIÓN DE EXPERIENCIAS DE FORMACIÓN

La literatura sobre formación de profesores indica la complejidad de las competencias necesarias para enseñar (Brown y Borko, 1992) y la divergencia entre formación teórica y práctica para la docencia. Por ello, la formación de futuros profesores requiere de entornos y contextos en que se trabajen sobre problemas significativos relacionados con su desarrollo profesional así como de la reflexión sobre dichas actividades (Llinares y Krainer, 2006).

Para la estadística y probabilidad, algunas actividades sugeridas en este sentido son las siguientes:

- Promover el trabajo colaborativo entre los, pues a través del intercambio de ideas y materiales entre los profesores que tienen problemas y necesidades comunes emergen nuevas ideas para la introducción de actividades, prácticas o el desarrollo de competencias. El profesor de estadística y probabilidad ha de manejar dos tipos de incertidumbre en su trabajo diario: el primero de ellos se relaciona con el conocimiento estocástico. La incertidumbre aparece también en el trabajo de la clase debido a las interacciones inesperadas entre estudiantes y profesor (Groth, 2008). La discusión entre profesores, incluso a distancia por medio de la tecnología puede ayudar a manejar estos dos niveles de incertidumbre.
- Planificación de una lección para enseñar a los alumnos algún aspecto de la estadística. Esta planificación pone en juego su conocimiento estadístico y su conocimiento profesional (Chick y Pierce, 2008). También permite el análisis de las transformaciones del conocimiento para adaptarlos a los distintos niveles de enseñanza.
- Trabajo con proyectos o actividades abiertas. Puesto que los currículos y metodología nuevos indican que se debe implicar a los estudiantes con este tipo de actividades, es importante que los profesores trabajen en la clase con el mismo enfoque. Una forma de lograrlos es hacerles jugar sucesivamente el papel de alumno y profesor es decir "tener la práctica de aprendiz". Sobre todo cuando hay poco tiempo de enseñanza, se debiera implementar un ciclo formativo en que los profesores primero trabajaran con un proyecto estadístico o actividad abierta y luego lo analizaran desde el punto de vista didáctico como medio para enriquecer simultáneamente el conocimiento estadístico y profesional del profesor (Godino, Batanero, Roa y Wilhelmi, 2008). Nosotros implementaremos este enfoque en el taller analizado en el Estudio 5.
- Análisis de tareas o ítems de evaluación y respuestas de alumnos a las mismas. Con esta tarea se trata de analizar el contenido estadístico y los conocimientos requeridos para resolver alguna tarea de evaluación usada en alguna investigación sobre razonamiento estadístico. Este tipo de análisis permitió también revelar posibles

dificultades de los profesores con los conceptos estadísticos en la experiencia de Godino, Batanero, Roa y Wilhelmi (2008). Tanto en el Estudio 1 como en el Estudio 5 propondremos a los profesores este tipo de actividad.

- Resolución de situaciones problemáticas paradójicas y reflexión sobre su contenido. En el caso de formación de profesores de secundaria que tengan sólidos conocimientos matemáticos, se trata de hacer experimentar a los futuros profesores la dificultad que encuentra un alumno al resolver las tareas o problemas que se les plantea. Puesto que el futuro profesor en este caso tiene unos conocimientos muy sólidos de estadística, hemos de buscar algunos ejemplos de problemas, que, aparentemente sencillo, puedan, sin embargo tener soluciones contraintuitivas o sorprendentes (Batanero, Godino y Roa, 2004). Usaremos este tipo de actividad en el taller descrito en el Estudio 5.

La influencia de las nuevas tecnologías en la estadística y su enseñanza ha sido reconocida por la International Association for Statistics Education (IASE) en diversas conferencias: Round Table Conference, celebrada en Granada (Garfield y Burrill, 2007), conferencia sobre Educación Estadística e Internet y en los Congresos Internacionales sobre la Enseñanza de la Estadística. En ellas se discutió sobre el software disponible para la enseñanza, los cambios implicados en el contenido y metodología, y el efecto de la tecnología en el aprendizaje y las actitudes de los alumnos. Como se destaca en estas conferencias, se ha reducido el tiempo de cálculo, permitiendo trabajar en clase con aplicaciones reales y proyectos.

Las posibilidades de simulación y creación de micromundos estocásticos virtuales permiten explorar los conceptos de probabilidad e inferencia y sustituir las demostraciones formales por razonamientos más intuitivos. Se añaden a éstas también las funciones de tutor y ayuda al autoestudio, evaluación y ejercitación (Biehler, 1997, 2003; Ben-Zvi, 2000). Podemos añadir las experiencias didácticas basadas en el uso de recursos disponibles en Internet, incluyendo las relacionadas con la formación de profesores en estadística y la interacción a distancia entre alumnos o alumnos y profesor.

Un punto esencial para introducir la tecnología en la clase de estadística será la adecuada preparación de los profesores, tanto desde el punto de vista técnico como didáctico. Ello es necesario pues algunos futuros profesores tienen dificultades en la utilización adecuada del software para fomentar la comprensión de los alumnos y consideraban que la simulación es sólo útil después de estudiar la probabilidad de manera teórica.

En otros casos al trabajar con la tecnología pasan por alto las ideas previas correctas de los estudiantes, centrándose sólo en sus errores (Stohl, 2005).

2.3. INVESTIGACIÓN SOBRE LA PROBABILIDAD CONDICIONAL

2.3.1. INTRODUCCIÓN

El segundo foco de interés del estado de la cuestión son las investigaciones relacionadas con la probabilidad condicional, que hemos resumido en Díaz, Batanero y Contreras (2010). La mayoría se centran en la evaluación de la comprensión o de los sesgos de razonamiento de estudiantes de secundaria o universidad o incluso de profesionales. Hay también un número reducido de investigaciones sobre estrategias de resolución de problemas, análisis de libros de texto o experimentos de enseñanza.

2.3.2. COMPRENSIÓN CONCEPTUAL

Comenzamos esta sección con las investigaciones relacionadas con la comprensión conceptual, algunas de las cuáles se centran en analizar si los alumnos son capaces de aplicar la definición y propiedades de la probabilidad condicional en situaciones muy simples, como son el muestreo con o sin reposición.

2.3.2.1. COMPRENSIÓN DE LA PROBABILIDAD CONDICIONAL

Intuitivamente la probabilidad condicional $P(A/B)$ de un suceso A dado otro suceso B se define como la probabilidad de que ocurra el suceso A sabiendo que ha ocurrido el suceso B y formalmente se define mediante la siguiente expresión, con la condición de que $P(B)>0$.

$$P(A/B) = P(A\cap B)/P(B).$$

Algunas investigaciones analizan si los estudiantes comprenden estas definiciones y pueden aplicarlas en la resolución de problemas y toma de decisiones en la vida diaria.

En este contexto Maury (1986) analizó la comprensión intuitiva de la probabilidad condicional en 374 estudiantes del curso preuniversitario y últimos cursos de Bachillerato, usando un contexto de extracción de bolas en urnas y otro de ruletas. En el primer caso, la extracción se realiza con y sin reemplazamiento. Aunque el 60% de los alumnos resolvió correctamente los problemas de probabilidad simple, sólo el 25% resolvió adecuadamente los de probabilidad condicional La dificultad se debió principalmente a que los dos sucesos considerados no eran equiprobables. En otro experimento Maury (1985, 1986) planteó a 290 alumnos de entre 13 y 16 años el mismo problema usando sucesos equiprobables y la tasa de respuestas correctas subió al 70%.

Para Maury este mayor éxito indica el reconocimiento intuitivo de la independencia por parte de los alumnos. La autora también observó un éxito mayor cuando en la pregunta se listan todos los sucesos posibles del espacio muestral, que cuando no se listan, considerando, en consecuencia, que parte de la dificultad de los problemas se debe a que los alumnos no son capaces de identificar y restringir correctamente el espacio muestral en la probabilidad condicional. Como indica Saenz (1999) uno de los errores sistemáticos de los alumnos en el campo de la probabilidad es la incapacidad para hacer inventario completo de los sucesos asociados a un fenómeno aleatorio, debido a fallos en el esquema combinatorio.

Totohasina (1992) propuso problemas de probabilidad condicional a 67 alumnos del curso preuniversitario, que habían estudiado la probabilidad, pero no la probabilidad condicional. Aproximadamente el 60% de ellos lo resolvieron correctamente apoyándose en representaciones como el diagrama en árbol, la tabla de doble entrada o la representación rectangular. Algunos estudiantes no interpretaron probabilísticamente el enunciado, haciendo uso sólo de las frecuencias absolutas de los casos. Otros alumnos a veces no restringieron el espacio muestral al calcular las probabilidades condicionales o bien confundieron la probabilidad conjunta y la probabilidad condicional.

Díaz (2007) plantea una pregunta a una muestra amplia de estudiantes de psicología donde les pide explicar en sus propias palabras la diferencia entre una probabilidad simple y una probabilidad condicional. Sus resultados sugieren que una parte importante de los alumnos define correctamente las dos probabilidades o al menos una. Aun así la tercera parte de los alumnos no da respuesta o tiene imprecisiones en una o las dos definiciones. Usaremos esta pregunta en el Estudio 2 y compararemos nuestros resultados con los de la autora.

2.3.2.2. INTERCAMBIO DE SUCESOS EN LA PROBABILIDAD CONDICIONAL

Muchos estudiantes no discriminan adecuadamente entre las dos direcciones de la probabilidad condicional *P(A/B)* y *P(B/A),* error que Falk (1986) denomina *falacia de la condicional transpuesta* y que se ha observado principalmente en problemas de contextos médicos. En este contexto se confunde la probabilidad de tener una enfermedad cuando ha sido positiva una cierta prueba, con la probabilidad de un resultado positivo en la prueba, sabiendo que se tiene la enfermedad (Eddy, 1982). Por ejemplo, la probabilidad de tener una mamografía positiva si se tiene cáncer de pecho (que es muy alta), se confunde con la de tener cáncer de pecho si se tiene la mamografía positiva (que es muy pequeña). En todo caso, la prestigiosa revista Lancet publicó un estudio donde se advierte de los riesgos de un diagnóstico erróneo en la prueba de mamografía (Gotzsche y Olsen, 2000), de la alta frecuencia de falsos positivos y se indica que para mejorar su capacidad de decisión, las mujeres que toman la prueba debería ser informada de los riesgos potenciales en el caso de un falso positivo.

También Batanero, Estepa, Godino y Green (1996) encontraron esta misma confusión en la interpretación de tablas de contingencia donde se confunden las dos frecuencias condicionales relacionadas con una misma celda de datos. En su investigación, alrededor del 20% de los estudiantes del curso preuniversitario confundieron “porcentaje de fumadores que contraen cáncer de pulmón” con “porcentaje de personas con cáncer de pulmón que fuman”.

Esta confusión se extiende asimismo al contexto de la interpretación del nivel de significación α en los contrastes de hipótesis (Vallecillos, 1994). El nivel de significación α se define como la probabilidad condicional de obtener un resultado *R* en la región de rechazo cuando la hipótesis nula H_o es cierta, es decir $\alpha = P(R/H_o)$. Cuando un contraste de hipótesis resulta significativo (lo que quiere decir que *R* ha ocurrido) y alguien pregunta por la probabilidad de haber cometido un error (la probabilidad de que H_0 sea cierta) a menudo se contesta con α. En esta situación se estaría confundiendo $P(R/H_o)$ con $P(H_o/R)$. Pero la probabilidad $P(H_o/R)$ no tiene sentido en inferencia clásica, y sólo puede calcularse en inferencia bayesiana, con lo que los investigadores estarían mezclando la filosofía e interpretación de estas dos escuelas de inferencia, que son incompatibles (Díaz, 2007).

Una posible explicación dada por Falk (1986) a la confusión entre los dos sentidos de la probabilidad condicional es la imprecisión del lenguaje cotidiano. Cuando escribimos una probabilidad condicional usando la notación matemática es claro cuál es el suceso condicionante y cuál el condicionado, pero en el lenguaje ordinario la probabilidad condicional (tener cáncer si se es fumador) y su inversa (ser fumador si se tiene cáncer) no siempre se distinguen claramente entre sí. Lo mismo sucede con la probabilidad conjunta (ser fumador y tener cáncer).

2.3.3. CONFUSIÓN ENTRE PROBABILIDAD CONDICIONAL Y CONJUNTA. FALACIA DE LA CONJUNCIÓN.

También Pollatsek, Well, Konold y Hardiman (1987) coinciden en que muchas de las dificultades que las personas tienen con la comprensión de la probabilidad condicional pueden deberse a la redacción de los enunciados. Su hipótesis se basa en los resultados de Einhorn y Hogarth (1986). Estos autores sugieren que los enunciados que usan la conjunción "*y*" pueden llevar a confundir la probabilidad conjunta y la probabilidad condicional, como ocurrió en su investigación al preguntar a 24 estudiantes *"¿Cuál es la probabilidad de ir al supermercado y comprar café?"*. En la investigación de Ojeda (1995), las preguntas sobre intersección de sucesos fueron más difíciles que las de probabilidad condicional y la mitad de los sujetos del estudio interpretaron la intersección como condicionamiento.

Un error relacionado con la probabilidad conjunta es la *falacia de la conjunción* (Tversky y Kahneman, 1982b) o creencia de que es más probable la intersección de dos sucesos que la de uno de ellos por separado o la de su unión. Puesto que la probabilidad de la intersección se obtiene multiplicando dos probabilidades, que son números menores o iguales a la unidad, el resultado de esta multiplicación nunca puede ser mayor que la de uno de los factores. Sin embargo, muchos alumnos dan una mayor probabilidad a la intersección de dos sucesos, si uno de ellos es muy probable, como por ejemplo, consideran más probable "ser joven e ir a la discoteca" que simplemente "ser joven". El error es resultado de considerar a la conjunción "ser joven e ir a la discoteca" más representativa de la población que ser simplemente joven.

Díaz (2005) analiza la *falacia de la conjunción,* recordando que algunos problemas de probabilidad que aparecen en la vida diaria están relacionados con la pertenencia de un elemento a una categoría. Indica que para resolverlos muchas personas aplican la *heurística de representatividad.* Dicha heurística consiste en evaluar la probabilidad de un suceso como el grado de correspondencia o similitud entre una muestra y una población y se usa para predecir resultados. Generalmente, este modo de razonar produce buenas respuestas, ya que las muestras y los resultados más representativos tienen una mayor probabilidad de ocurrencia.

Una de las reglas básicas de la probabilidad es que cuanto más especificamos un suceso, menor es su probabilidad (Tversky y Kahneman, 1982a). Díaz indica que estas relaciones de orden entre las probabilidades no se transforman en otras correspondientes de representatividad, sino que, en ocasiones, la representatividad se incrementa cuando especificamos más nuestras condiciones. Al tratar de estimar la probabilidad de la intersección de dos sucesos a veces ocurre un conflicto en los casos en que la intersección de los dos sucesos es más representativa del caso cuya probabilidad se nos pide.

Díaz evalúa de la falacia de la conjunción, usando un cuestionario que incluye cuatro ítems con dos grupos de estudiantes de primer año de psicología, uno compuesto por 81 estudiantes y el segundo grupo por 76. La evaluación se realizó con finalidad diagnóstica, dentro de la asignatura Análisis de Datos en Psicología antes de comenzar el tema de probabilidad condicional, para conocer los posibles sesgos que presentaban los estudiantes y tenerlos en cuenta en la enseñanza del tema. Los ítems son versiones modificadas de los utilizados por Tversky y Kanhneman (1982a). En todos ellos se limitó a presentar un suceso simple y otro compuesto, para evitar la dispersión de la respuesta si se incluyen demasiados distractores. Se varió el formato (datos dados en forma de frecuencia o en forma probabilística) y contexto. A pesar de que los problemas se pasaron en dos grupos diferentes, se observan los mismos patrones de respuesta. Los resultados de este estudio indican diferencias estadísticamente significativas según el tipo de problema (formato), pero no del contexto del problema, curso, interacción o de la nota obtenida en el bachillerato. En consecuencia, podemos decir que el problema de la falacia de la conjunción se dio en forma similar en los dos grupos de estudiantes y que no depende de la calificación previa de los alumnos en Bachillerato. En contra de lo esperado, los resultados son algo mejores en los problemas presentados en términos de probabilidad, que en los de frecuencia.

Otros autores han tratado de analizar si la enseñanza mejora el razonamiento sobre las probabilidades compuestas y evita la falacia de la conjunción. Tversky y Kahneman

(1982b) comparan las respuestas de sujetos con diferentes conocimientos de probabilidad, encontrando tan sólo ligeras mejoras en los que han tenido mayor número de cursos de estadística, con la excepción de un grupo de sujetos con conocimientos estadísticos bastante avanzados. Crandall y Greenfield (1986) organizaron un enseñanza en que los estudiantes sistemáticamente usaron diagramas de Venn como ayuda para resolver problemas similares a los descritos anteriormente, encontrando una mejoría moderada en la capacidad de resolución de los problemas en el grupo experimental de sujetos que utilizó esta técnica.

2.3.4. RELACIÓN ENTRE INDEPENDENCIA Y PROBABILIDAD CONDICIONAL

Las dificultades de comprensión de la probabilidad condicional son debidas en ocasiones a la falta de percepción de la independencia, concepto ligado al de probabilidad condicional tanto en su definición, como por su significado intuitivo (Díaz, 2007). Dicha dificultad continúa en la Universidad, como muestra Sánchez (1996), quien pasó un cuestionario a 88 profesores de Matemáticas que participaban en México en un programa de actualización. Sólo 44 profesores hicieron intentos sistemáticos por resolver un problema de independencia. De ellos 39 dieron una respuesta, pero sólo 4 utilizaron la regla del producto y llegaron a la solución correcta. Entre las respuestas incorrectas encuentra dos tipos de razonamiento:

1. Creer que dos sucesos son independientes si y sólo si son excluyentes, error muy extendido, y descrito anteriormente por Kelly y Zwiers (1986), quienes suponen que es debido a la imprecisión del lenguaje ordinario, en que “independiente” puede significar, a veces, separado. Esta creencia es errónea pues dos sucesos excluyentes son justamente dependientes pues uno no puede ocurrir a la vez que el otro.
2. Creer que para que dos sucesos sean independientes cada uno de ellos ha de pertenecer a un experimento diferente y no al mismo experimento. También esto es falso, pues al tomar al azar una mujer de una población, los sucesos “ser menor de 20 años” y “ser sevillana” son independientes, aunque se refieran a un mismo experimento (tomar la mujer al azar de la población).

2.3.5. CONDICIONAMIENTO Y CAUSACIÓN

La causalidad es un concepto científico, filosófico y psicológico complejo, aunque las personas lo comprenden de forma intuitiva en base a su experiencia cotidiana con situaciones de relaciones de causa y efecto. La relación causal estricta ocurre cuando al variar un suceso *A* (por ejemplo, al cambiar la cantidad de agua en un recipiente) otro suceso *B* cambia (el peso del recipiente cambiará). Es difícil de hallar en el mundo real relaciones perfectas de causa y efecto y por ello los estadísticos hablan de relación de *causa débil* cuando al suceder *A* cambia la probabilidad de que ocurra *B* (Díaz, 2007). Es decir, cuando *P(B/A)* es diferente de *P(B).* Una relación de causalidad implica una dependencia de tipo estadístico entre los sucesos implicados. Por ejemplo, a mayor número de horas de estudio aumenta la probabilidad de una mejor nota, pero no es seguro, pues el alumno podría hacer un mal examen por un cansancio excesivo.

Desde el punto de vista de la probabilidad condicional, si un suceso *A* es la causa estricta de un suceso *B*, siempre que suceda *A*, sucederá *B*, por lo que *P(B/A)=1.* Sin embargo lo contrario no es cierto, ya que dos sucesos pueden ser dependientes, sin que uno de ellos sea causa del otro. Por ejemplo hay una relación inversa entre la tasa de natalidad de un país y la esperanza de vida, pero la relación no es de tipo causal. Desde el punto de vista psicológico, la persona que evalúa una probabilidad condicional *P(A/B)* puede percibir dos relaciones diferentes entre *A* y *B*:

- *Relación causal:* Si percibe que *B* (suceso condicionante) es causa de *A* (causa), estimando la probabilidad de un efecto *A* dado cierto conocimiento de la causa *B*. Por ejemplo, la probabilidad de que si una persona tiene pulmonía, ésta tenga fiebre es alta.
- *Relación diagnóstica:* Si percibe *A* como una causa de *B*, se realizaría un razonamiento diagnóstico, estimando la causa dado el conocimiento del efecto. En el ejemplo anterior, si una persona tiene fiebre, la probabilidad de que tenga pulmonía no es alta, pues hay muchas enfermedades que originan la fiebre.

Aunque matemáticamente los dos enunciados son equivalentes, desde un punto de vista psicológico hay una creencia que las relaciones causales son más fuertes que las relaciones diagnósticas (Tversky y Kahneman, 1982a) debido a la existencia de un sesgo causal cuando las personas se enfrentan con tareas relacionadas con la probabilidad condicional.

La relación de causalidad también se asocia, a menudo, con la secuencia temporal (Falk, 1986) y algunos estudiantes tienen problemas con la condicionalidad cuando se invierte el eje de tiempo en que los sucesos ocurren de una forma natural. Por ejemplo, si les preguntamos cuál es la probabilidad de que el abuelo de una persona tuviera los ojos azules. Esta creencia, que consiste en rechazar la evidencia ocurrida después del suceso que juzgamos, se conoce como *falacia del eje temporal* y consiste en suponer que el suceso condicionante ha de preceder temporalmente al condicionado.

Entender que la probabilidad de un suceso puede ser revisada a la luz de resultados posteriores es importante en la aplicación del Teorema de Bayes, donde la actualización de las probabilidades a la luz de los resultados juega un papel tan importante. Gras y Totohasina (1995) identifican tres tipos de concepciones erróneas sobre la probabilidad condicional en estudiantes:

- En la *concepción cronológica* los estudiantes interpretan la probabilidad condicional *P(A/B)* como una relación temporal, donde el evento condicionante *B* siempre precede al suceso *A.*
- En la *concepción causal*, los estudiantes interpretan la probabilidad condicional *P(A/B)* como una relación causal implícita, donde el suceso condicionante B es la causa y A la consecuencia.
- En la *concepción cardinal* los estudiantes interpretan la probabilidad condicional *P(A/B)* como la proporción $\frac{Card(A\cap B)}{Card(B)}$, que es correcta sólo en el caso de un espacio muestral finito equiprobable. Otros estudiantes interpretan *P(A/B)* como la proporción $\frac{Card(A)}{Card(B)}$, que es siempre falso.

2.3.6. RESOLUCIÓN DE PROBLEMAS

Otras investigaciones estudian las estrategias y errores de los estudiantes al resolver problemas. Las más importantes son las relacionadas con el Teorema de Bayes y las que analizan cómo influyen las variables del problema en la resolución de la tarea.

2.3.6.1. PROBLEMAS RELACIONADOS CON EL TEOREMA DE BAYES

Totohasina (1992) encuentra las siguientes estrategias espontáneas entre los alumnos que tratan de resolver un problema del Teorema de Bayes:

- Cambio de referencial: Consiste en restringir el espacio muestral para tratar el problema como si se tratase de un problema de probabilidad simple.
- Cálculo de cocientes de porcentajes, lo que implica, en la práctica la fórmula de Bayes, aunque deducida de una manera informal.

Totohasina, después de realizar un experimento de enseñanza de la probabilidad condicional, donde no se introduce formalmente el teorema de Bayes, aunque se plantean y resuelven problemas de probabilidad inversa basándose en árboles y tablas de doble entrada, propone a los alumnos un problema de Bayes y examina los tipos de diagramas en árbol utilizados por los alumnos. El autor supone que los alumnos pueden encontrarse con dificultades en función del tipo de representación elegida para resolver el problema, que les es dado en formato verbal. Al pasar, por ejemplo, a una tabla de doble entrada, se dificulta la percepción de la naturaleza secuencial de algunos problemas, porque lo que queda más visible es la intersección de los dos sucesos y puede llevar a los alumnos a confundir la probabilidad condicional y la conjunta.

De un total de 65 alumnos que participaron en la evaluación, 26 construyeron un diagrama en árbol, con lo que se resaltó claramente el aspecto secuencial de los experimentos (cálculo de la probabilidad directa); 9 de estos estudiantes llegaron a construir correctamente al árbol inverso, que muestra claramente la probabilidad inversa. Otros 7 alumnos usaron el diagrama en árbol, pero no contemplaron el aspecto secuencial y no llegaron a asignar correctamente las probabilidades. Sólo 9 alumnos llegaron a la solución correcta de los problemas. Los autores concluyen que el uso de un árbol es el recurso más efectivo para resolver problemas de probabilidad condicional, sobre todo cuando se refiere a un problema diacrónico (dirigido en el tiempo). Pero puede reforzar las concepciones causalistas o cronologistas en los alumnos que las manifiestan. Otra dificultad descrita por Totohasina es la necesidad de invertir condición y condicionado en los problemas tipo Bayes, ya que, como hemos señalado, los alumnos con frecuencia confunden el papel de estos dos sucesos en una probabilidad condicional y por tanto confunden una probabilidad condicional con su inversa.

Díaz y de la Fuente (2007) realizaron un estudio sobre resolución de problemas bayesianos por parte de estudiantes de Psicología, antes y después de la enseñanza de la probabilidad condicional. Su estudio se basa en otros estudios como el de Tversky y Kahneman (1982a) que estudian la denominada *falacia de las tasa base.* Esta falacia se refiere al hecho de ignorar la probabilidad a priori del suceso en la población en problemas que involucran la probabilidad inversa. Mientras que los estudios anteriores dan una explicación teórica para las dificultades de los estudiantes, la hipótesis de las autoras es que la resolución de problemas bayesianos involucra que los estudiantes deban recordar y aplicar varios conceptos y procedimientos probabilísticos. Las autoras tratan de identificar las dificultades de los estudiantes y de esta forma, diseñar procesos educativos que permitan trabajar estas tareas.

Las autoras proponen tres ítems a diferentes muestras de estudiantes de Psicología, dentro de la asignatura de Análisis de Datos antes de impartirles el tema de probabilidad condicional, aunque los estudiantes ya tenían nociones del tema, ya que lo habían estudiado en secundaria. Después de administrar el test a los estudiantes, se les enseñó probabilidad condicional, dedicándose tres clases teóricas y dos prácticas a este tema, que incluyó los siguientes contenidos: probabilidad condicional y conjunta, teorema de la probabilidad total y teorema de Bayes. Los estudiantes recibieron instrucción sobre el uso de diagramas de árbol y tablas de doble entrada para la resolución de los problemas y se les advirtió sobre la existencia de errores y sesgos como la falacia de la conjunción, la falacia del eje del tiempo y la falacia de las tasas base. Posteriormente se les volvió a pasar uno de ítems distribuidos aleatoriamente.

La gran proporción de respuestas en blanco en los ítems antes de la enseñanza sugiere que estos estudiantes habían olvidado el teorema de Bayes, estudiado en secundaria o que no reconocieron que tenían que aplicar este teorema. La falacia de las tasas base fue la principal respuesta y el formato frecuencial no ayudó a los alumnos a resolver el problema. Se obtuvieron mejores resultados en el ítem en el que aparecía la fórmula de Bayes, ya que esto ayudó a los alumnos a reconocer la respuesta correcta.

Para analizar los resultados después de la enseñanza las autoras decidieron cambiar los ítems para no forzar las respuestas de los alumnos. Distinguieron los siguientes pasos en las repuestas de los estudiantes, teniendo en cuenta el grado de corrección de su solución:

a) *Identificar los datos del problema*: El estudiante debe discriminar entre probabilidad simple y compuesta, para poder realizar correctamente las sucesivas particiones del espacio muestral, e identificar que datos se refieren a cada uno de los conceptos del enunciado del problema.
b) *Construir una representación adecuada*: El segundo paso es construir un diagrama de árbol adecuado para representar el experimento secuencial y la partición secuencial de la población. Esta representación debe servir al estudiante para reconocer el conjunto de sucesos posibles.
c) *Identificar la probabilidad condicional*: Para continuar, los estudiantes deben identificar qué probabilidad se pide en el problema y que ésta es una probabilidad condicional inversa. Las autoras encuentran errores del tipo de que los estudiantes confundan la fórmula la probabilidad condicional con su inversa, con una probabilidad simple o con una probabilidad conjunta.
d) *Calcular el denominador de la fórmula de Bayes*: Después de identificar el problema como el cálculo de una probabilidad condicional y recordar la fórmula de Bayes, el estudiante debe calcular el numerador y denominador, este último debe ser calculado con la regla de la probabilidad total, esto es, multiplicando las probabilidades del cada rama del árbol y sumando cada una de esas probabilidades conjuntas. El alumno debe entender que se trata de sucesos dependientes, para aplicar correctamente la regla del producto en este caso.
e) *Calcular la probabilidad inversa (teorema de Bayes)*: Finalmente, el estudiante debe sintetizar todos los pasos anteriores y calcular el numerador (probabilidad conjunta) y denominador (probabilidad total) para obtener la probabilidad inversa, es decir, aplicar el teorema de Bayes.

El número de respuestas correctas después de la instrucción sugiere la mayor dificultad del ítem cuando el problema se presentaba en términos de frecuencias, donde fue mayor el número de alumnos que no identificó correctamente los datos, o deja la respuesta en blanco. Para complementar el estudio, se realizó un análisis de los errores. La mayoría de los obstáculos fueron la identificación incorrecta de los datos, la realización incorrecta del diagrama de árbol o de una tabla de doble entrada, la partición incorrecta del espacio muestral, la confusión entre diversas probabilidades (simple, conjunta condicional) y entre una probabilidad condicional y su inversa y errores en la fórmula de Bayes, invirtiendo denominador y numerador u omitiendo algún término.

Observaron también fallos en el razonamiento proporcional, ya que algunos estudiantes no eran capaces de operar con fracciones o hallar el inverso de una fracción. Finalmente algunos alumnos operan conjuntamente probabilidades y valores esperados o confunden estos dos términos, obteniendo como consecuencia, valores mayores que la unidad para la probabilidad de un suceso, sin ser conscientes del error que esto supone.

2.3.6.2. INFLUENCIA DEL LENGUAJE Y EL FORMATO

Pollatesk y cols. (1987) analizan las variables que pueden influir en la resolución de los problemas de probabilidad condicional, entre otras, el formato en que se da el enunciado (probabilidad o porcentaje), y el contexto. El porcentaje de respuestas correctas fue similar para los problemas dados en probabilidades o dados en porcentajes, pero en el caso en que el factor causal se hacía presente, la versión porcentual resultó más sencilla y los alumnos dan mayor número de respuestas correctas que en la versión probabilística.

Gigerenzer (1994) sugiere que la dificultad en la resolución de problemas referidos al teorema de Bayes desaparece cuando las preguntas se plantean en términos de frecuencias. Llama a este formato *frecuencias naturales* porque se asemeja más a la forma en que recogemos información de las frecuencias de sucesos aleatorios en una situación de muestreo natural a lo largo de nuestra experiencia. Cuando la información se ofrece en términos de frecuencia, el cálculo de la probabilidad a posteriori es más natural, porque el sujeto no tiene que aplicar toda la complejidad del teorema de Bayes, sino sólo tener en cuenta los casos favorables y posibles, de modo que el problema se transforma en un problema simple de probabilidad. Sus resultados coinciden con los de Ojeda (1995) en sus investigaciones con estudiantes de secundaria.

Resultados similares se han obtenido al plantear otros problemas probabilísticos en términos frecuenciales, por lo que Gigerenzer y Hoffrage (1995) recomiendan, cuando sea posible, cambiar a frecuencias el formato de las preguntas. En el caso particular de la *falacia de la conjunción,* Fiedler (1988) encontró una reducción considerable del número de respuestas incorrectas al plantear las preguntas en formato de frecuencias.

Siguiendo esta línea de investigación, Lonjedo y Huerta estudian algunas variables que influyen sobre la resolución de problemas de probabilidad condicional por estudiantes con y sin nociones previas de los contenidos (Huerta y Lonjedo 2003; Lonjedo, 2003; Lonjedo y Huerta, 2005). Los autores analizan si la resolución de los problemas puede hacerse utilizando el razonamiento numérico y por tanto los estudiantes no necesitan utilizar las relaciones entre probabilidades para resolver el problema. Esto ocurre sobre todo cuando los problemas no son interpretados como probabilidades y a causa de ello no se usan las propiedades de la probabilidad para obtener la solución del problema. Es sólo al final del proceso de resolución del problema cuando los estudiantes responden a la pregunta del problema en términos de probabilidad.

Lonjedo y Huerta concluyen que hay factores no probabilísticos que afectan al éxito en la resolución del problema, como pueden ser la naturaleza de los datos de los problemas. En su estudio clasifican los problemas atendiendo a la naturaleza de los datos en el texto del problema, distinguiendo entre datos presentados en términos de probabilidad, datos presentados en frecuencias absolutas, datos presentados en términos de razón o datos expresados en combinación. En su estudio se les presentaban a los alumnos seis problemas de probabilidad condicional en el que la estructura de datos no varía, pero sí su naturaleza y el contexto. Los resultados mostraban que el éxito en la resolución de los problemas no dependía del uso correcto de unas determinadas fórmulas. Estudiantes capacitados para ello, no siempre usaban los datos interpretados como probabilidades. Por otra parte, estudiantes que no conocen dichas fórmulas resuelven los problemas, por lo que el éxito no depende de ellas sino del uso correcto del razonamiento aritmético aplicado a unos datos no interpretados como probabilidades, sino como razones o proporciones.

En Lonjedo y Huerta (2007) completan el estudio anterior y se centran no sólo en la influencia que tienen los datos numéricos sino también en la posible influencia de la expresión que refiere a una probabilidad condicional en el comportamiento de los estudiantes. Como resultado producen un esquema de análisis para el estudio de estos comportamientos, particularizándolo para tres problemas estructuralmente isomorfos. Observan como la presentación de los datos en los problemas de probabilidad condicional tienen cierta influencia en el éxito de los resolutores y en los procesos de razonamiento implicados. El incremento en el porcentaje de éxitos, se debe a la modificación de expresiones o términos que provocan ambigüedad en la condicionalidad y el formato de presentación de los datos. Cuando los datos están expresados en frecuencias absolutas y la probabilidad condicional en porcentajes el porcentaje de alumnos que tienen éxito crece.

2.3.6.3. USO DE REPRESENTACIONES EN LA SOLUCIÓN DE LOS PROBLEMAS

Las representaciones juegan un papel primordial en la resolución de problemas matemáticos. Los sistemas de representación son entidades abstractas compartidas que se usan para organizar la información mediante determinadas reglas sintácticas. Martignon y Wassner (2002) plantean el uso de diagrama en árbol, junto con las frecuencias naturales para enseñar la resolución de problemas que involucren el Teorema de Bayes. En la solución del mismo se comienza identificando los sucesos a que se refiere la pregunta del problema y se les denota. A continuación se les organiza para operar con ellos, por ejemplo, sustituyéndolos en una fórmula y finalmente se opera para obtener la solución.

Una representación adecuada de los problemas facilita el cálculo y produce soluciones acertadas a los problemas (Gigerenzer y Hoffrage, 1995). El éxito se debe a la estrecha relación entre el diagrama en árbol y la forma inductiva en que procesamos la información en las tareas bayesianas. Usando la regla matemática de Bayes realizaríamos las mismas operaciones, pero en un orden diferente, es decir, siguiendo un proceso deductivo, y aunque nuestra mente está equipada para ambos tipos de procesos, el primero es más intuitivo y natural.

2.3.6.4. PLANTEAMIENTO DE PROBLEMAS

La resolución de problemas tiene su complemento en el planteamiento de problemas, que permite investigar el pensamiento matemático avanzado de los alumnos. La mayoría de los estudios realizados sobre el planteamiento de problemas usan problemas aritméticos y corresponden a niveles no universitarios. Sin embargo Posadas (2008) y Penalva y Posadas (2009) analizan el planteamiento de problemas de probabilidad condicional por estudiantes universitarios.

Los autores indican que los procesos de planteamiento de problemas añaden a lo anterior un mayor nivel de abstracción que los de resolución y la necesidad de utilizar adecuadamente el lenguaje natural y formal. Se interesan por el trabajo que desarrollan grupos de estudiantes cuando resuelven y plantean problemas conjuntamente sobre esta temática, considerando también una serie de creencias erróneas sobre probabilidad. Se basan en investigaciones que indican que sólo una minoría de estudiantes universitarios analiza los fenómenos aleatorios desde un punto de vista formal y utiliza correctamente los procedimientos necesarios para el cálculo de la probabilidad de un suceso (Jones, Langrall y Money, 2007).

La investigación tuvo en cuenta dos etapas diferenciadas. En la primera se realizó un estudio exploratorio de las características de la actividad matemática que desarrollan los estudiantes cuando realizan tareas de resolución y planteamiento de problemas con contenidos de probabilidad, mientras que en la segunda etapa se enfoca en el estudio de los procesos de abstracción matemática que generan los estudiantes cuando resuelven dichas tareas. Analizan los recursos, heurísticos, tipo de razonamiento y demanda cognitiva de las tareas mediante estudios de casos en grupos de estudiantes que han resuelto las tareas propuestas y después han explicado su trabajo en una entrevista. Los autores identifican acciones epistémicas de reconocimiento y construcción que manifiestan los grupos de estudiantes cuando resuelven y plantean problemas relativos al teorema de la probabilidad total y al teorema de Bayes. La investigación permite también poner de manifiesto el aprendizaje del estudiante y el potencial de las tareas de plantear problemas para desencadenar procesos de abstracción.

2.3.7. ENSEÑANZA DE LA PROBABILIDAD CONDICIONAL

2.3.7.1. EXPERIMENTOS DESDE LA PSICOLOGÍA

Diferentes investigadores en psicología han tratado de explicar los errores y falacias en el razonamiento en situación de incertidumbre. Estas investigaciones tienen diversas implicaciones respecto a la forma en cómo debe llevarse a cabo la enseñanza del tema (Sedlmeier, 1999).

El programa “heurísticas y sesgos” interpreta las falacias como resultado de un proceso cognitivo que lleva a una conclusión incorrecta, bien por usar un modelo inapropiado de la situación o por falta de estructuras cognitivas específicas (Kahneman, Slovic y Tversky, 1982; Pérez-Echeverría, 1990). En este paradigma, las heurísticas y sesgos se suponen resistentes a la enseñanza. Por ejemplo, la instrucción específica no ha tenido efecto sobre la falacia de la conjunción en experimentos basados en este marco teórico y los sesgos y falacias descritos por los autores se han encontrado en sujetos con alta preparación matemática (Tversky y Kahneman, 1982b).

Otra perspectiva defendida por Nisbett y Ross (1980) es considerar que podemos adquirir un razonamiento estadístico correcto y que desarrollamos reglas abstractas intuitivas para resolver los problemas cotidianos, siempre que encontremos en las situaciones ciertas claves que nos permitan reconocerlas como aleatorias. La enseñanza podría mejorar nuestro razonamiento estadístico natural, que se adquiere por la experiencia repetida en resolución de problemas. Por tanto, estas reglas podrían enseñarse mediante ejercitación, aunque el aprendizaje en concreto de la probabilidad condicional requiere, además, enseñanza de la lógica condicional (Sedlemeier, 1999).

Gigerenzer (1994) sugiere que la dificultad en la resolución de problemas referidos al teorema de Bayes desaparece cuando las preguntas se plantean en términos de frecuencias. En sus investigaciones, cambia el enunciado, y observa un aumento notable de los razonamientos de tipo bayesianos. Gigerenzer sugiere que nuestra mente está mejor equipada para resolver problemas bayesianos cuando la información y las preguntas se dan en términos de frecuencias. Sin embargo, en un experimento organizado para hacer comprender a los sujetos que la probabilidad de la intersección ha de ser menor que la de cada suceso, Fiedler (1988) no consiguió una mejora de los resultados al incluir el formato frecuencial. Ojeda (1995) obtuvo el mismo resultado para el caso de problemas de Bayes, puesto que sólo el 45% de estudiantes de secundaria (255 alumnos en la muestra) fueron capaces de resolver problemas bayesianos dado en términos de frecuencias.

Martignon y Wassner (2002) plantean el uso de una representación especial en forma de diagrama en árbol, junto con las frecuencias naturales para enseñar la resolución de problemas que involucren el Teorema de Bayes. Estos autores indican que, mientras que los problemas de probabilidad que los estudiantes encuentran en la vida diaria son concretos y numéricos, los instrumentos de cálculo que les presentamos (por ejemplo el teorema de Bayes) son altamente formalizados. Los alumnos adquieren reglas nemotécnicas de aplicación que olvidan fácilmente. Los autores indican que una representación adecuada de los problemas facilita el cálculo y produce soluciones acertadas a los problemas tratados.

2.3.7.2. EXPERIMENTOS DE ENSEÑANZA EN DIDÁCTICA DE LA MATEMÁTICA

Estrada, Díaz y de la Fuente (2006) realizaron un estudio inicial sobre sesgos en el razonamiento sobre probabilidad condicional en alumnos universitarios. Las autoras presentan un estudio sobre la evaluación de los sesgos en el razonamiento condicional de una muestra de 159 estudiantes de Matemáticas, Magisterio y Psicología. La muestra participante en el trabajo está formada por 65 estudiantes de Magisterio que seguían una asignatura optativa de Estadística Aplicada a la Educación, 37 estudiantes de 5º curso de Matemáticas, que seguían un curso optativo de Didáctica de la Matemática y 57 estudiantes de Psicología que seguían un curso obligatorio de Análisis de Datos. Todos ellos habían estudiado probabilidad, aunque con diferente profundidad en las asignaturas citadas.

Los resultados muestran que la falacia de la conjunción no es tan frecuente como en la investigación de Tversky y Kahneman (1982a), y en su lugar aparece el sesgo de equiprobabilidad (Lecoutre y Durand, 1988), que consiste en considerar los dos sucesos equiprobables, en más de la mitad de los estudiantes de todos los grupos. No obstante, es claro que los estudiantes olvidan que el producto de probabilidades ha de ser menor que la probabilidad de cada suceso simple o no aplican esta propiedad, cuando se les presenta un problema en un contexto cotidiano.

Asimismo se encontró que los estudiantes confunden independencia y exclusión, cambiando los términos de la probabilidad condicional, y asignando a la probabilidad conjunta un valor mayor que a la probabilidad simple, es decir, violando las reglas lógicas del cálculo de probabilidades condicionales. Las autoras sugieren que la enseñanza de la probabilidad condicional requiere también que los estudiantes confronten sus propios razonamientos incorrectos, así como la integración con la enseñanza de la lógica condicional.

Díaz y de la Fuente (2006) se centran en la enseñanza del teorema de Bayes con apoyo tecnológico a 78 alumnos de psicología. La experiencia formó parte de un curso breve de introducción a la inferencia bayesiana apoyado en nuevas tecnologías. Los alumnos participantes se dividieron en cuatro pequeños grupos de entre 15 y 20 alumnos cada uno, repitiéndose la experiencia en cada uno de los grupos con la misma profesora. Se dedicaron un total de 3 horas a la enseñanza del Teorema de Bayes, de las cuales las dos primeras se llevaron a cabo en aula tradicional y la tercera en el laboratorio de informática donde los alumnos trabajaban en parejas resolviendo problemas con ayuda de subprogramas Excel preparados para la experiencia.

La evaluación se llevó a cabo mediante dos instrumentos; un cuestionario formado por 8 ítems que evalúa los conocimientos conceptuales relacionados con el teorema y dos problemas abiertos de aplicación. Los resultados indican que la mayoría de estudiantes adquirió una comprensión intuitiva del teorema de Bayes con pocas horas de trabajo, fueron capaces de identificar los datos en los problemas abiertos, organizar los datos en una tabla, realizar los cálculos e interpretarlos en el contexto de los problemas. La tasa de éxito fue mayor que la informada en otras investigaciones, incluso las basadas en formatos de frecuencias.

2.4. LA PROBABILIDAD CONDICIONAL EN LOS LIBROS DE TEXTO

Algunos autores analizan la presentación de la probabilidad condicional en los libros de texto, junto con otros conceptos que generalmente suelen aparecer ligados, tales como dependencia e independencia o probabilidades en experimentos compuestos.

En su revisión de libros de texto, Ortiz (1999) comprueba que muchos no incluyen el estudio de la probabilidad condicional y cuando lo hacen el estudio está muy restringido. En su análisis, Ortiz (1999) encontró dos variantes principales de la definición de probabilidad condicional en los libros de secundaria:

- Definición tipo 1: "*Si A y B son dos sucesos en un espacio muestral E, definimos la probabilidad de A condicionada por B, como el cociente entre la probabilidad de la intersección de A y B y la probabilidad de B*".
- Definición tipo 2: "*La probabilidad del suceso condicionante P(B) debe ser distinta de cero para poder definir la probabilidad condicional como*:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} \text{”.}$$

En otros casos, Ortiz encontró que se hacía mención de la probabilidad condicional a través de un ejemplo, aunque sin nombrarla explícitamente. Un hecho observado por el autor es la ambigüedad del lenguaje, lo que da lugar a confusión en los alumnos, ya que, en vez de indicar que la probabilidad del suceso *B* viene condicionada por el suceso *A* se dice que es el suceso *B* el que viene condicionado. Esto podría dar la falsa impresión de que estamos trabajando con una probabilidad simple o de que se quiere realizar otra operación entre sucesos. Esta ambigüedad llega a veces a presentar una definición incorrecta de “suceso condicionado” desde el punto de vista tradicional y señalada por autores como Falk (1986).

Pocos libros indican que el suceso condicionante tiene que tener una probabilidad no nula. No se destaca tampoco el papel de los dos sucesos que intervienen en la probabilidad condicional ni el diferente valor de la probabilidad condicional cuando intercambiamos estos sucesos entre sí.

Respecto a la independencia de sucesos, la definición que presentan los libros de texto es a través de la probabilidad condicionada, es decir que “dos sucesos son independientes cuando $P(B/A)=P(B)$” suele aparecer en gran cantidad de los textos. También aparece en la mayoría de los casos la definición de que dos sucesos son independientes cuando el resultado de uno de ellos no influye en la realización del otro. El teorema de la probabilidad total y las fórmulas de Bayes apenas se incluyen en los textos analizados.

Lonjedo y Huerta (2005) exploran los problemas escolares verbales de probabilidad condicional en los libros de texto, y observaron que las cantidades presentes en la mayor parte de estos problemas, no están expresadas en términos de probabilidad sino que presentan naturaleza diversa. Según los autores los problemas escolares de probabilidad condicional encontrados en los libros de texto sólo serán clasificados como problemas de cálculo de probabilidades si los datos del problema son interpretados como probabilidades y es necesaria las relaciones entre probabilidades para solucionarlos.

Batanero y Díaz (2005) analizan la presentación de la probabilidad condicional en los libros de texto recomendados en la asignatura de Análisis de Datos en Psicología. El método del estudio consistió en la elaboración de un listado con las 31 universidades españolas en las que se imparte la licenciatura de Psicología y la solicitación a los profesores correspondientes de la bibliografía recomendada. Las autoras consiguieron repuestas de 23 universidades. De un total de 79 libros diferentes recomendados sobre análisis de datos 20 eran citados por 4 o más universidades. Trece de ellos incluían el tema de probabilidad condicional y otros 5 libros de orientación bayesiana. En un primer momento se examinó las definiciones, propiedades y relaciones con otros conceptos presentados en la muestra de libros de texto. Se describió las siguientes categorías:

1. *Probabilidad condicional y propiedades*: La definición de probabilidad condicional se encontró en la mayor parte de los libros analizados. En unos casos se deduce el concepto de probabilidad condicional a partir de la definición de la frecuencia relativa de *A* condicionada a la ocurrencia de *B*, aunque en la mayoría la definición aparece como el cociente $P(A/B) = P(A \cap B)/P(B)$ donde $A \cap B$ representa el suceso ocurrencia conjunta de *A* y *B*, y siempre suponiendo que *P(B)>0*. En otros se demuestra que una probabilidad condicional cumple los axiomas de probabilidad o que la probabilidad condicional implica una reducción del espacio muestral.
2. *Dependencia, independencia e intercambiabilidad*: En la mayoría de los casos encuentran el término independencia deducida matemáticamente de la regla del producto: *A* y *B* son independientes si y sólo sí $P(A \cap B) = P(A)P(B)$. También aparece el termino *i*ndependencia relacionado con la *p*robabilidad condicional, ya que dos sucesos son independientes si la probabilidad de uno de ellos no cambia al condicionarlo por el otro. Un concepto directamente relacionado con los anteriores que se suele confundir con ellos es el de sucesos *mutuamente excluyentes*, en algunos libros se indica que, precisamente son dependientes, ya que la ocurrencia de uno hace que el otro no pueda ocurrir. En gran parte de los libros estudiados se deduce el concepto de sucesos dependientes a partir del concepto de independencia. En algunos de los textos de enfoque bayesiano se incluye el término intercambiabilidad, como hipótesis no tan restrictiva como la independencia. También se relaciona con la probabilidad condicional, ya que dos sucesos son independientes si la probabilidad de uno de ellos no cambia al condicionarlo por el otro.
3. *Regla del producto*: Este concepto aparece en los libros como una regla que permite calcular las probabilidades en experimentos compuestos. En algunos casos se proporciona la regla diferenciando entre sucesos independientes y dependientes para su aplicación. Pero en la mayoría de las ocasiones aparece expresada como una relación entre intersección de sucesos y el producto de ellos.
4. *Ley de la probabilidad total*: Este concepto suele estar asociado a la regla del producto, siempre partiendo de *n* sucesos incompatibles y mutuamente excluyentes. En otros casos no aparece explícitamente definida, pero se presenta implícitamente como parte del estudio del Teorema de Bayes.

Las autoras destacan que son pocos los libros de texto que usan formas de visualización (diagramas en árbol, tablas de doble entrada, diagramas de Venn) que ayuden al alumno a recordar estos conceptos. Por el contrario es usual presentar ejercicios resueltos, que ayudan a contextualizar los conceptos presentados. El recurso del ordenador sólo lo encontraron en los dos libros específicamente basados en este recurso. Los ejercicios observados en el estudio desvelan que son pocos los que se plantean en contexto de simulación o aplicando directamente la probabilidad condicional, siento también escasos los relacionados con la independencia. Las autoras los clasifican en varias categorías:

- Ejercicios que calculan la probabilidad condicional de un experimento simple;
- Ejercicios que calculan la probabilidad condicional en el contexto de muestreo con reposición;
- Ejercicios que calculan la probabilidad condicional en el contexto de muestreo sin reposición;
- Ejercicios que calculan la probabilidad condicional a partir de probabilidades conjuntas y simples;
- Ejercicios que calculan la probabilidad compuesta haciendo uso de la regla del producto en el caso de sucesos dependientes;
- Ejercicios que calculan la probabilidad compuesta haciendo uso de la regla del producto en el caso de sucesos independientes;
- Ejercicios para determinar si dos sucesos son dependientes o independientes;
- Ejercicios para diferenciar entre sucesos mutuamente excluyentes y sucesos independientes;
- Ejercicios para calcular la probabilidad total;
- Ejercicios para resolver problemas bayesianos;
- Ejercicios para resolver problemas de probabilidad condicional y conjunta en experimentos compuestos de diferentes experimentos simples;
- Ejercicios para estimar probabilidades condicionales y conjuntas mediante simulación.

Un punto a destacar es que los libros apenas tienen en cuenta los sesgos descritos en la literatura y es escaso también el uso de recursos didácticos y simulación. Hay que tener en cuenta también que, aproximadamente la tercera parte de los libros elegidos inicialmente, no incluían el tema de probabilidad condicional.

2.5. CONCLUSIONES DEL ESTUDIO DE LAS INVESTIGACIONES PREVIAS

En los apartados anteriores hemos realizado una revisión de las investigaciones sobre formación de profesores para enseñar probabilidad y sobre comprensión de la probabilidad condicional, conjunta, independencia y teorema de Bayes.

Respecto al tema de formación de profesores, nuestro análisis muestra que la investigación es muy escasa, y casi inexistente en lo que se refiere a la probabilidad condicional. Por otro lado la investigación centrada en los conocimientos matemáticos sobre probabilidad (en general) se ha enfocado predominantemente en futuros profesores de educación primaria. Sería necesaria completar esta investigación con estudios sobre el conocimiento matemático referido a la probabilidad condicional, tanto en profesores de educación primaria como secundaria.

Más escasa es todavía la investigación previa sobre conocimientos didácticos de profesores en formación o ejercicio de los dos niveles educativos en este tema e incluso respecto a la probabilidad en general o la descripción de la evolución de estos conocimientos como consecuencia de intervenciones educativas. Nuestro trabajo, por tanto aportara una primera información en un campo en que no se poseen apenas conocimientos que puedan guiar la acción formativa de los formadores de profesores.

Respecto a la investigación sobre comprensión de la probabilidad condicional, se han descrito diferentes errores y algunas de las razones para los mismos, así como algunas variables que facilitan la resolución de los problemas sobre estos conceptos.

La investigación revisada muestra que este no es un tema sencillo, que tiene una amplia variedad de matices y los alumnos lo asocian con la problemática de la causalidad y temporalidad, teniendo dificultad en la percepción de los experimentos compuestos en el caso de situaciones sincrónicas. Se confunde independencia y exclusión, se cambian los términos de la probabilidad condicional, se confunde ésta con la conjunta y se asigna a la probabilidad conjunta un valor mayor que a la probabilidad simple, violando las reglas lógicas del cálculo de probabilidades. La presentación en los libros de texto no siempre es completa y los problemas incluidos en el tema, en ocasiones son problemas aritméticos y no verdaderos problemas probabilísticos.

En consecuencia, pensamos que dada la importancia de la probabilidad condicional en estadística y en la vida cotidiana, las dificultades descritas en los alumnos y las carencias de los libros de texto, merece la pena continuar la investigación sobre el tema. Nuestros Estudios 3 y 4, incluidos en el capítulo 5, de recursos disponibles en Internet y paradojas clásicas para la enseñanza del tema se orientan a identificar aquellos que pudieran contribuir a la superación de algunos de los errores descritos. Este estudio lo usaremos como fundamento para el Estudio 5 descrito en el capítulo 6 sobre diseño de unos procesos de estudio de la probabilidad condicional y su didáctica, dirigidos a la formación de profesores.

CAPÍTULO 3
EVALUACIÓN DE CONOCIMIENTOS DE FUTUROS PROFESORES DE EDUCACIÓN PRIMARIA

3.1. INTRODUCCIÓN

Aunque la probabilidad condicional no figura explícitamente como contenido en la Educación Primaria, si aparece el tema de la lectura de tablas de doble entrada, en el cuál implícitamente se incluye la comprensión del cálculo elemental de la probabilidad simple, condicional y compuesta en situaciones sencillas. También aparece explícitamente el estudio de la probabilidad simple, incluyendo la comprensión del lenguaje ordinario relacionado con el azar y la probabilidad (que para un futuro maestro incluiría el comprender los enunciados del tipo “si” o condicional y del tipo “y” o conjunción) y el cálculo de probabilidades sencillas aplicando la regla de Laplace.

Por otro lado, la comprensión y el cálculo de probabilidades en la tabla de doble entrada es un recurso profesional para el futuro profesor, en temas como el diagnóstico de dificultades de aprendizaje o la evaluación de sus estudiantes. El estudio elemental de dicho tema se incluye, asimismo, en textos de probabilidad elemental dirigida a la formación de profesores, tales como, por ejemplo Godino, Batanero y Cañizares (1987), Azcárate y Cardeñoso (2001) o Batanero y Godino (2002).

Estas consideraciones nos han llevado a plantear un estudio de las competencias matemáticas y didácticas de los futuros profesores de Educación Primaria, en relación a la interpretación de probabilidades simples, compuestas y condicionales en tablas de contingencia muy elementales, que llevamos a cabo en este capítulo. En lo que sigue, se describen los objetivos e hipótesis del Estudio 1, la muestra de estudiantes, el contexto en que se toman los datos y se realiza un análisis a priori detallado de la tarea de evaluación propuesta. Seguidamente se presentan y discuten los resultados, separando los correspondientes al conocimiento matemático y al conocimiento didáctico.

3.2. OBJETIVOS E HIPÓTESIS DEL ESTUDIO 1

El Estudio 1 se orienta a evaluar los conocimientos matemáticos y didácticos de los futuros profesores de Educación Primaria en relación con algunos objetos matemáticos que aparecen implícitos en la interpretación de tablas de contingencia 2x2, es decir con dos filas y dos columnas. Más concretamente, podemos diferenciar dos objetivos en el estudio.

Objetivo 1. Analizar el conocimiento común de probabilidad que ponen de manifiesto los futuros profesores de educación primaria en el cálculo de probabilidades simples, compuestas y condicionales a partir de datos presentados en una tabla de contingencia 2x2.

Para lograr este objetivo plantearemos un problema elemental en el que tengan que calcular estas probabilidades una muestra de profesores en formación. Se realizará un análisis semiótico de las soluciones dadas, para categorizar las respuestas encontradas, identificando posibles conflictos semióticos. Se compararán los resultados con los obtenidos por Estrada y Díaz (2007) en una tarea semejante y una muestra reducida de futuros profesores de educación primaria de la Universidad de Lleida que habían seguido previamente un curso optativo de didáctica de la estadística.

Objetivo 2. Analizar algunos conocimientos didácticos de los futuros profesores de educación primaria sobre la tabla de contingencia 2x2 y el cálculo de probabilidades a partir de la misma.

Para lograr este objetivo, y una vez resuelto el problema anterior, se propone a los futuros profesores realizar un análisis didáctico de la tarea. Por un lado, se les pide analizar los contenidos matemáticos involucrados en la resolución del problema, para evaluar de este modo su conocimiento especializado del contenido matemático. Por otro, se les pregunta por las variables de tarea del problema para evaluar su conocimiento del contenido y la enseñanza.

Hipótesis

Respecto a este estudio, *esperamos que nuestros resultados muestren un conocimiento insuficiente de los futuros profesores de la muestra sobre la probabilidad condicional, tanto respecto al conocimiento matemático, como en relación a las componentes evaluadas del conocimiento didáctico.*

Esta hipótesis se basa en los resultados descritos sobre la comprensión de la probabilidad condicional en la sección 2.3, donde se muestran numerosos sesgos de razonamiento que es razonable que aparezcan también en los futuros profesores. Por otro lado, investigaciones previas como, por ejemplo, la de Estrada y Díaz (2007) con una tarea similar a la nuestra, aunque de menor tamaño, muestran un insuficiente conocimiento de los futuros profesores respecto a la probabilidad simple, condicional y conjunta. Resultados similares sobre la comprensión de la probabilidad se muestran en otras investigaciones como las de Azcárate (1995), Begg y Edward (1999), Watson (2001) y Ortiz, Nordin, Batanero, Serrano y Rodríguez (2006). Esperamos que estos resultados se repliquen en nuestro trabajo.

Respecto a los componentes del conocimiento didáctico, suponemos que será también insuficiente, aunque no tenemos antecedentes sobre este punto. Pero, sin embargo Ball, Lubienski y Mewborn (2001) sugieren que dicho conocimiento depende en gran medida del conocimiento matemático, por lo que, si este no es suficiente, tampoco es de esperar que lo sea el conocimiento didáctico.

La importancia de los citados objetivos es que permitirán obtener información sobre las necesidades formativas de los profesores, que pueda servir para preparar materiales didácticos y mejorar la acción educativa con estos futuros profesores.

A continuación se describe la muestra participante, la tarea propuesta y los resultados obtenidos, que se han publicado en Contreras, Estrada, Díaz y Batanero, (2010) y Contreras, Batanero, Díaz y Fernándes (2011).

3.3. MUESTRA Y CONTEXTO CURRICULAR

Los participantes fueron estudiantes de segundo curso de la Diplomatura de Magisterio, Educación Primaria, que cursaban la asignatura Currículo de Matemáticas en la Educación Primaria. Esta es una asignatura de carácter eminentemente práctico (dos tercios de las horas docentes se dedican a actividades prácticas) cuyo contenido exclusivo es la didáctica de la matemática. Los contenidos matemáticos que se imparten en esta asignatura son los mismos que se incluyen en el currículo de Educación Primaria en España, que abarca cuatro bloques temáticos: *Números y operaciones*, *La medida: estimación y cálculo de magnitudes*, *Geometría y Tratamiento de la información, azar y probabilidad* (MEC, 2006a).

Los datos se tomaron dentro de una evaluación parcial de la asignatura, que se relaciona con el cuarto bloque del currículo de matemáticas en educación primaria: *Tratamiento de la información, azar y probabilidad.* En la tarea propuesta también se contempla el tema de *Objetos y procesos matemáticos, Variables de tarea.* Los estudiantes realizaron la actividad individualmente y produjeron un informe escrito dentro del aula. Todos estos temas fueron objeto de las sesiones teóricas en la asignatura. Anteriormente los estudiantes habían realizado una actividad similar de análisis de variables de tarea y de los objetos y procesos matemáticos, en relación a un problema de proporcionalidad y otro problema de razonamiento algebraico elemental.

El año anterior al que se tomaron los datos, en la asignatura "*Matemáticas y su Didáctica*" (asignatura de 9 créditos ECTS, cada uno de los cuales equivale a 25 horas de trabajo del estudiante, incluido el trabajo dentro y fuera del aula) se dedicaron 2 créditos al bloque de *Tratamiento de la información, azar y probabilidad.* En este bloque los estudiantes revisaron y ampliaron su información sobre conceptos como: datos, variables estadísticas, distribuciones de frecuencias, gráficos, medidas de posición central y dispersión. También se revisaron y ampliaron las ideas de aleatoriedad, probabilidad y la asignación de probabilidades mediante regla de Laplace o estimación frecuencial de la probabilidad, incluyendo algunos ejercicios sencillos de probabilidad condicional y compuesta.

Hay que tener también en cuenta que los estudiantes proceden de diferentes especialidades de bachillerato. La mayoría de ellos, provienen del bachillerato de humanidades, en el que los contenidos matemáticos son menores y las matemáticas en los últimos cursos con opcionales. Algunos estudiantes provienen de bachillerato de ciencias sociales, y en este caso, habrán estudiado estadística durante un cuatrimestre completo, incluyendo probabilidad simple y condicional e incluso introducción a la inferencia, el curso anterior a su ingreso en la Facultad de Educación. No es habitual que el estudiante haya cursado el bachillerato de ciencias, pero en este caso, aunque tienen matemáticas el año anterior de ingresar en la Facultad, no se incluye la estadística dentro de esta materia. En todo caso, cualquiera que sea el bachillerato cursado, durante la Enseñanza Secundaria Obligatoria (12-16 años) todos los estudiantes habrán recibido una enseñanza mínima de probabilidad y probabilidad condicional.

3.4. ANÁLISIS A PRIORI DE LA TAREA

Figura 3.4.1. Tarea propuesta

La lectura de tablas de doble entrada es uno de los contenidos incluidos en Educación Primaria.

a. Resuelve el siguiente ejercicio.
b. ¿Qué contenidos matemáticos se trabaja? (clasifícalos según los objetos matemáticos estudiados en clase: problemas, lenguaje, conceptos, propiedades, procedimientos, argumentos).
c. Indica algunas variables que se pueda cambiar en esta tarea para variar la dificultad o para variar el contenido matemático.

Tarea. En un colegio se pregunta a los alumnos, obteniendo los siguientes resultados:

	Chicos	Chicas	Total
Le gusta el tenis	400	200	600
No le gusta	50	50	100
Total	450	250	700

Si elegimos al azar uno de estos alumnos:

a. ¿Cuál es la probabilidad de que le guste el tenis?
b. ¿Cuál es la probabilidad de que sea una chica y además le guste el tenis?
c. Sabiendo que el alumno elegido es chica, ¿Cuál es la probabilidad de que le guste el tenis?

En la Figura 3.4.1 se incluye el enunciado de la tarea propuesta, que consta de una primera parte de resolución de problemas y dos cuestiones sobre análisis didáctico. La primera pregunta está orientada a evaluar el conocimiento común del contenido de los futuros profesores, es decir, leer una tabla de doble entrada y resolver a partir de ella un problema de probabilidad simple, otro de probabilidad compuesta y otro de probabilidad condicional. El conocimiento especializado del contenido se evalúa en la segunda pregunta y en la tercera el conocimiento del contenido y la enseñanza.

3.4.1. CONFIGURACIÓN DE OBJETOS MATEMÁTICOS Y SOLUCIONES CORRECTAS AL PROBLEMA

En la primera parte de la tarea se muestra una tabla de doble entrada o de contingencia 2x2, que presenta en forma resumida la distribución de frecuencias de una muestra de datos, clasificada respecto a dos variables estadísticas, cada una de las cuáles tiene dos categorías. En la tarea propuesta hemos elegido esta forma elemental de tabla de doble entrada para mostrar que, incluso en este formato elemental, aparecen numerosas dificultades y conflictos semióticos, que seguramente serán más numerosos y serios cuando se pida a los futuros profesores interpretar una tabla más compleja.

La complejidad semiótica de la configuración de objetos matemáticos “tabla de contingencia” se pone de manifiesto al analizar su versión más simple (Tabla 3.4.2). Aunque todos los datos de las celdas *a, b, c, d*, mostradas en dicha tabla, se refieren a frecuencias absolutas, cada una de ella depende de una doble condición (valores de su fila y columna) y su significado no es equivalente desde un punto de vista psicológico (Estepa, 1993). La celda *a* representa la presencia conjunta del carácter *A* y *B* y la *d* su ausencia conjunta, mientras que en las otras dos celdas aparece sólo uno de los dos caracteres.

Tabla 3.4.2. Formato típico de la tabla de doble entrada 2x2

	B	*no B*	Total
A	*a*	*b*	*a+b*
no A	*c*	*d*	*c+d*
Total	*a+c*	*b+d*	*n=a+b+c+d*

Además, de acuerdo a Estrada y Díaz (2007), a partir de la frecuencia absoluta doble de una celda dada (por ejemplo *a*) se pueden deducir tres frecuencias relativas diferentes:

- Frecuencia relativa doble: $\frac{a}{a+b+c+d}$
- Frecuencia relativa respecto a su fila: $\frac{a}{a+b}$
- Frecuencia relativa respecto a su columna: $\frac{a}{a+c}$
- También es posible calcular dos tipos de frecuencias relativas marginales; respecto a las filas y respecto a las columnas: $\frac{a+b}{a+b+c+d}$ y $\frac{a+c}{a+b+c+d}$.

Suponiendo un contexto de extracción al azar de un elemento de la población o muestra dada y supuestos equiprobables todos los sujetos de la muestra, podemos aplicar la regla de Laplace como cociente entre casos favorables y posibles, y obtener una probabilidad diferente de cada una de las frecuencias relativas anteriores.

En el apartado (a) del problema se pide el cálculo de la probabilidad simple del suceso *A*={*le gusta el tenis*}, que viene dada como $P(A)=\frac{a+b}{a+b+c+d}$. El alumno ha de leer la tabla y localizar los casos favorables y desfavorables al suceso pedido, para aplicar la Regla de Laplace. Los casos posibles son el total de la muestra; los casos favorables el total de alumnos a los que gusta el tenis, es decir, una frecuencia absoluta marginal. La probabilidad de que le guste el tenis, *P(A),* se obtiene dividiendo $a+b=600$ (chicos y chicas a las que les gusta el tenis) entre el total $a+b+c+d=700$. Se obtiene una probabilidad de que guste el tenis $P(A)=600/700=0{,}857$.

En la Tabla 3.4.3 reproducimos un análisis semiótico simplificado de esta solución, en la cual se ponen en juego conceptos (experimento aleatorio, suceso, casos favorables y posibles, total, suma y división, frecuencia absoluta doble y marginal), procedimientos (suma, división, regla de Laplace), representaciones verbales numéricas y simbólicas, todo ello ligado con un argumento. Los alumnos deben realizar múltiples procesos de interpretación y representación; así como de particularización de los conceptos. La resolución del problema paso a paso supone un proceso de descomposición o análisis, mientras que la conclusión final alcanzada se logra mediante un proceso de síntesis.

Tabla 3.4.3. Análisis semiótico de la respuesta correcta en el apartado (a)

<table>
<tr><th>Expresión</th><th>Contenido</th></tr>
<tr><td>En un colegio se pregunta a los alumnos, obteniendo los siguientes resultados:
<table>
<tr><th></th><th>Chicos</th><th>Chicas</th><th>Total</th></tr>
<tr><td>Le gusta el tenis</td><td>400</td><td>200</td><td>600</td></tr>
<tr><td>No le gusta</td><td>50</td><td>50</td><td>100</td></tr>
<tr><td>Total</td><td>450</td><td>250</td><td>700</td></tr>
</table></td><td>- El alumno comienza con un proceso de significación para interpretar el texto y sus elementos. Lee la tabla (procedimiento y representación).
- Identifica las dos variables del enunciado (concepto) y establece una correspondencia (procedimiento) entre los códigos (chicos, chicas, le gusta el tenis, no le gusta; lenguaje) y los valores de la variable (concepto) en niños imaginarios (elemento fenomenológico).
- Identifica las frecuencias dobles, marginales y total en la tabla (conceptos y procedimiento) realizando un proceso de particularización.</td></tr>
<tr><td>Si elegimos al azar uno de estos alumnos: ¿Cuál es la probabilidad de que le guste el tenis?</td><td>- Ha de interpretar la pregunta en relación a un experimento aleatorio (tomar un alumno al azar) mediante procesos de significación y particularización del concepto.
- Identifica los posibles sucesos en el experimento (le gusta o el tenis); conceptos y particularización de los mismos.
- Ha de identificar lo que pide el problema: probabilidad simple (concepto) y particularizarlo a este caso.</td></tr>
<tr><td>$P(A)=\dfrac{b+d}{a+b+c+d}$ =600/700=0,857</td><td>- Identifica los casos favorables correspondientes a la frecuencia marginal (particularización de un concepto).
- Los representa mediante la expresión b+d (representación de una operación; concepto de suma).
- Identifica y representa los casos posibles a+b+c+d o total de la muestra.
- Aplica la regla de Laplace, dividiendo casos favorables entre casos posibles (procedimiento).</td></tr>
</table>

En el apartado (b) se pide la probabilidad de que ocurran simultáneamente *A* y *B*. El alumno ha de localizar en la tabla la frecuencia absoluta de que ocurran a la vez los dos sucesos, que serían los casos favorables. Al igual que el caso anterior, los casos posibles lo constituirían el total de la muestra. Aplicando simplemente la regla de Laplace se obtiene $P(A \cap B) = \dfrac{b}{a+b+c+d}$.

Es decir la probabilidad de ser chica y jugar al tenis, pedida en el apartado (b), se calcula dividiendo los casos favorables b =200 (chicas a las que les gusta el tenis) por los casos posibles $a+b+c+d$ =700 (alumnos en la muestra). Se obtiene por tanto una probabilidad igual a $P(A \cap B) = \frac{200}{700} = \frac{2}{7} = 0,28.$ En la Tabla 3.4.4 reproducimos un análisis semiótico simplificado de esta solución, en la cual se ponen en juego, además de los objetos matemáticos y procesos descritos para el primer apartado, el concepto de probabilidad compuesta, espacio muestral y sucesos en el experimento compuesto, así como frecuencia relativa doble. El alumno ha de comprender que los sucesos son dependientes, es decir, ha de leer la tabla y realizar un juicio de asociación en los datos.

Tabla 3.4.4. Análisis semiótico de la respuesta correcta en el apartado (b)

Expresión	Contenido
Si elegimos al azar uno de estos alumnos: ¿Cuál es la probabilidad de que sea chica y que le guste el tenis?	- Ha de interpretar la pregunta en relación a un experimento aleatorio (tomar un alumno al azar) mediante procesos de significación y particularización del concepto. - Identifica los posibles sucesos en el experimento compuesto (conceptos y particularización de los mismos). - Ha de identificar lo que pide el problema: probabilidad compuesta (concepto) y particularizarlo a este caso.
$P(A \cap B) = \frac{200}{700} = \frac{2}{7} = 0{,}28$	- Identifica los casos favorables *b*, correspondientes a la frecuencia relativa doble (concepto y particularización). - Identifica los casos posibles *a*+*b*+*c*+*d* o total de la muestra (concepto y particularización). - Aplica la regla de Laplace (procedimiento). - Utiliza la notación de la intersección de sucesos (representación y concepto). - Ha de realizar una lectura correcta de la tabla, para percibir que las variables están asociadas (juicio de asociación), pues en otro caso, se debería aplicar la regla del producto (procedimientos y discriminación de una propiedad).

El tercer apartado pide la probabilidad condicional de *B*, sabiendo que ha ocurrido *A*; $P(B/A)=\dfrac{b}{b+d}$; ya que ahora se pide la probabilidad de que le guste el tenis sabiendo de antemano que es una chica. Se obtiene dividiendo *b=200* (chicas a las que les gusta el tenis) entre *b+d=250* (número de chicas). El alumno ha de identificar los casos favorables y posibles en la tabla, y darse cuenta que el condicional supone una restricción del espacio muestral. Es decir los casos posibles ahora son sólo las chicas.

Tabla 3.4.5. Análisis semiótico de la respuesta correcta en el apartado (c)

Expresión	Contenido
Sabiendo que el alumno es chica, ¿cuál es la probabilidad de que le guste el tenis?	- Ha de interpretar la pregunta en relación a un experimento aleatorio (tomar un alumno al azar) mediante procesos de significación y particularización del concepto. - Identifica los posibles sucesos en el experimento (le gusta o el tenis); conceptos y particularización de los mismos. - Ha de identificar lo que pide el problema: probabilidad condicional (concepto) y particularizarlo a este caso. - Ha de identificar la condición y condicionado en la probabilidad condicional pedida (concepto y particularización).
$P(B/A) = \frac{200}{250} = \frac{4}{5} = 0{,}8$	- Identifica los casos favorables b, correspondientes a la frecuencia relativa doble (concepto y particularización). - Debe discriminar entre *P*(*A/B*) y *P*(*B/A*). - Identifica los casos posibles *b+d* o total de la chicas. (concepto y particularización). - Aplica la regla de Laplace (procedimiento y particularización).

Se obtiene por tanto una probabilidad igual a $P(B/A)=\frac{200}{250}=\frac{4}{5}=0{,}8$. En la Tabla 3.4.5 reproducimos el análisis semiótico. Observamos que a los objetos matemáticos y procesos anteriores se une la probabilidad condicional, condición y condicionado (conceptos), discriminación de una propiedad condicional y su transpuesta (propiedad) y la notación correspondiente.

3.4.2. POSIBLES CONFLICTOS SEMIÓTICOS

Aunque las tareas requeridas son muy sencillas y de hecho se reducen a la lectura correcta de la tabla y al cálculo de probabilidades simples utilizando la regla de Laplace, alrededor del 20% de los estudiantes del curso preuniversitario confundieron las frecuencias condicionales respecto a fila y columna en la investigación de Batanero, Estepa, Godino y Green (1996), cuando se les pidió calcular una frecuencia relativa condicional. También Estrada y Díaz (2007) describen diferentes conflictos en su muestra de futuros profesores en esta tarea.

Es de prever, por tanto, que los futuros profesores puedan tener dificultades para diferenciar los diferentes objetos matemáticos y para el cálculo de probabilidades a partir de la tabla. Según Falk (1986) muchos estudiantes no discriminan adecuadamente entre las dos direcciones de la probabilidad condicional *P(A/B)* y *P(B/A),* denominando a este error *falacia de la condicional transpuesta*. Nosotros interpretamos este error como un conflicto semiótico consistente en confundir los dos términos (condición y condicionado) o bien no diferenciarlos en un enunciado de probabilidad condicional.

La facilidad de resolución de tareas que implican probabilidades condicionales depende, según Pollatsek, Well, Konold y Hardiman (1987), de la redacción de los enunciados. Por ejemplo, Einhorn y Hogarth (1986) observaron que algunos alumnos interpretan incorrectamente la conjunción “y”, confundiendo la probabilidad conjunta y la probabilidad condicional. También la mitad de los sujetos del estudio en la investigación de Ojeda (1995) interpretaron la intersección como condicionamiento. Por tanto es de esperar que esta confusión aparezca en nuestro estudio.

Más recientemente la influencia del formato en que se presentan los datos, sobre la solución de los problemas simples de probabilidad condicional, se ha comprobado en la investigación de Huerta y Lonjedo (2005). En nuestra tarea, para evitar el efecto de esta variable, las tres preguntas se plantean con datos dados en un mismo tipo de formato.

3.4.3. VARIABLES DE TAREA EN EL PROBLEMA

En el último apartado de la tarea propuesta se pide a los estudiantes identificar las variables de tarea en el problema, que puedan colaborar a aumentar o disminuir la dificultad que tienen los estudiantes al resolverlo. En la bibliografía sobre resolución de problemas (Lester, 1983, Goldin y McClintock, 1984; Schoenfeld, 1985, Castro, 1991) se suele diferenciar varios tipos de variables. Nosotros consideraremos tres tipos:

- *Variables de la tarea*: Son las variables relacionadas con la naturaleza del problema y se pueden referir a su contenido, sintaxis, contexto o estructura (Castro 1991). En un mismo problema o tarea, ligeras variaciones en el enunciado pueden variar su dificultad, las estrategias con que los alumnos tratan de resolverlo o bien los contenidos matemáticos de la tarea.

En el problema planteado se podría variar el número de celdas en filas o columnas de la tabla, suprimir los totales, aumentar el tamaño de las frecuencias en las celdas o el total de la muestra, o el formato con que se presentan los datos; por ejemplo, se podrían presentar el enunciado verbalmente, o por medio de un listado o de forma gráfica. Se podría plantear preguntas más complejas (como calcular la probabilidad de elegir dos o más sujetos con o sin reemplazamiento en cada caso). Todos estos cambios afectarían la dificultad del problema. Se les podría preguntar por la proporción (en vez de por la probabilidad), por la asociación entre las variables o bien hacer un gráfico a partir de la tabla, con lo cual cambiaría el contenido matemático de la tarea.

- *Variables del sujeto*: los alumnos tienen diversas edades, nivel de formación, capacidades, estilos cognitivos, formación matemática previa, intereses, actitudes e historia. Las circunstancias sociales y familiares también pueden influir, por ejemplo, el apoyo de sus padres en el estudio o los medios que éstos le proporcionan.
- *Variables de la situación de resolución*. Dichas variables se refieren a las condiciones físicas y sociales, herramientas disponibles en el momento de resolver el problema o el modo en que se trabaja. Por ejemplo, es diferente si se trabaja sólo

o en grupo, si se dispone de un ejemplo resuelto o no, o de calculadora o del tiempo disponible para su resolución.

3.5. CONOCIMIENTOS MATEMÁTICOS EN RELACIÓN A LA PROBABILIDAD CONDICIONAL

Recogidas las respuestas, se procedió a leerlas y agrupar aquellas similares, mediante un proceso inductivo y cíclico, propio del análisis cualitativo de datos, hasta llegar a categorizarlas. Una vez identificadas las principales categorías de respuestas, realizamos el análisis semiótico de un ejemplo típico de cada una de ellas, para mostrar, por un lado, la complejidad de las respuestas correctas y los diferentes objetos matemáticos, que el estudiante ha de relacionar entre sí por medio de funciones semióticas. Asimismo mostramos los principales conflictos semióticos que dan lugar a respuestas parcialmente correctas o incorrectas. El análisis se completa con el estudio de las frecuencias de respuestas.

3.5.1. ANÁLISIS GLOBAL DE SOLUCIONES

En primer lugar, hemos clasificado simplemente las respuestas en correctas e incorrectas según la solución dada por los futuros profesores. Utilizando la nomenclatura de la Tabla 3.4.2, usaremos la notación presentada en las Tablas 3.4.3 a la 3.4.5 para asignar los diferentes sucesos, por lo que definimos como sucesos: A = “Le gusta el tenis”; $\bar{A}$ = “No le gusta el tenis”; B = “chicas”, $\bar{B}$= “chicos”.

Tabla 3.5.1. Frecuencias absolutas y porcentaje de respuestas correctas

Tarea	Respuestas correctas	Porcentaje
a. Cálculo de *P(A)*	120	65,5
b. Cálculo de $P(A\cap B)$	77	42,0
c. Cálculo de *P(A/B)*	80	43,7

Como podemos observar, el cálculo de probabilidades representa un problema importante para los futuros profesores, ya que, en ninguna de las tres preguntas planteadas se obtiene un porcentaje elevado de respuestas correctas. La mayoría de los estudiantes es incapaz de responder correctamente a dos de las tareas propuestas y en la tarea primera, referente al cálculo de la probabilidad simple, un tercio de los futuros profesores cometen algún tipo de error, aumentando el porcentaje en las tareas referentes al cálculo de probabilidades compuestas y condicionales. Estrada y Díaz (2006) obtuvieron un 75%, 55% y 53%, respectivamente, de respuestas correctas en el cálculo de probabilidad simple, compuesta y condicional en una muestra de 65 futuros profesores que habían seguido un curso optativo de estadística en la Universidad de Lleida. Sus resultados fueron mejores que los nuestros, aunque la tarea fue más difícil que la ahora planteada, pues incluía una desigualdad y una negación en el enunciado, que los alumnos tendrían que interpretar para definir el suceso *B*.

El problema se agrava con la gran cantidad de estudiantes que dejaron en blanco las tareas (aproximadamente un tercio), lo que indica una falta total de conocimientos básicos de probabilidad, destacando la tarea de probabilidad condicional (39,1%). En consecuencia, estos resultados refuerzan la necesidad de mejorar la formación estadística de los futuros profesores de educación primaria.

3.5.2. ANÁLISIS DETALLADO DE SOLUCIONES

En un segundo tipo de análisis, hemos clasificado los tipos de respuestas dadas por los futuros profesores no sólo en correctas e incorrectas, sino teniendo en cuenta el tipo de probabilidad calculada y los conflictos semióticos implicados, en el caso de respuesta incorrecta. Seguimos la metodología de análisis utilizada en investigaciones previas, como las de Olivo (2008) y Mayén (2009).

La importancia del estudio de conflictos semióticos la resalta Godino (2002, 2003), quien indica que en las prácticas matemáticas intervienen objetos matemáticos, algunos de los cuáles tienen un carácter ostensivo (símbolos, esquemas, gráficos, etc.), mientras que la mayoría son no ostensivos. Para resolver problemas matemáticos los símbolos (significantes) remiten a otros objetos (significados). Muchos autores en didáctica de las matemáticas destacan la importancia de las representaciones, aunque no siempre se reconoce la variedad de objetos que desempeñan el papel de representación y de objeto representado. Godino, Contreras y Font (2006) usan el término función semiótica como una "correspondencia entre conjuntos", con tres componentes:

- *Un plano de expresión*: objeto inicial, considerado frecuentemente como el signo;
- *Un plano de contenido*: objeto final, considerado como el significado del signo, esto es, lo representado, lo que se quiere decir, a lo que se refiere un interlocutor;

- *Un criterio o regla de correspondencia*, esto es un código interpretativo que relaciona los planos de expresión y contenido.

En ocasiones, el significado que el profesor o investigador quiere atribuir a una expresión matemática es interpretado en una forma matemáticamente incorrecta por el alumno y se produce el *conflicto semiótico.* Para Godino, Batanero y Font (2007) un *conflicto semiótico* es una interpretación incorrecta de expresiones matemáticas, que produce errores, que no son debidos a falta de conocimiento, sino a no haber relacionado adecuadamente los dos términos de una función semiótica. Godino, Wilhelmi y Font (2008), han propuesto diferentes niveles de análisis didácticos de las respuestas de alumnos o de las tareas e incluso de los procesos de enseñanza y aprendizaje de las matemáticas. En esta investigación aplicamos el que denominan Análisis Semiótico, que se centra en los objetos y procesos que intervienen en las prácticas matemáticas o emergen de ellos y permite describir la complejidad Ontosemiótica de dichas prácticas. A continuación se describen los resultados.

C1. Respuestas básicamente correctas

Serían los alumnos que calculan la probabilidad pedida en cada uno de los apartados. Por lo tanto han sido capaces, tanto de interpretar el enunciado del problema y los datos, como de leer correctamente la tabla de contingencia. Han identificado correctamente los casos favorables y posibles para cada apartado, aplicando correctamente la regla de Laplace. Serían las respuestas esperadas que se han comentado en el análisis a priori de la tarea.

C.1.1. *Expresión simbólica incorrecta de la probabilidad.* La solución es correcta, con un error de terminología, por aplicación de un símbolo que no es usual para denotar el concepto requerido. Aunque básicamente la respuesta es correcta, ésta denota un aprendizaje incompleto del lenguaje matemático. Un ejemplo en el primer apartado es el siguiente, en que el estudiante identifica correctamente los casos favorables y posibles y aplica la regla de Laplace, pero otorga la notación de probabilidad al resultado numérico, mostrando una terminología inadecuada*: "La probabilidad de que le guste el tenis es P*(*600/700*)*"* (alumno 70). No se realiza el análisis semiótico por ser básicamente igual a la solución correcta.

C.1.2. *Errores en operaciones*: Son aquellas respuestas que plantean bien el problema pero fallan en la operación. Un ejemplo (alumno 43), se reproduce a continuación y su análisis semiótico se presenta en la Tabla 3.5.2. A partir de este momento el análisis semiótico solamente se hará con la resolución, puesto que el enunciado del problema es el mismo para cada alumno, por lo que la actividad de lectura e interpretación de la tabla llevaría los mismos procesos y objetos que se describieron en la solución correcta.

"La probabilidad de que le guste el tenis es:

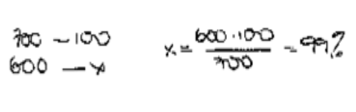

700 — 100
600 — x
$x = \frac{600 \cdot 100}{700} = 99\%$

Tabla 3.5.2. Análisis semiótico de la respuesta C1.2

Expresión	Contenido
La probabilidad de que le guste el tenis es: 70---100 600---x $x = \frac{600 \cdot 100}{700} = 99\%$	- Lenguaje (repite la pregunta del enunciado en forma afirmativa para indicar la solución), proceso de representación. - No es claro que el estudiante use la idea de experimento aleatorio, suceso y probabilidad. - Usaría la idea de proporción (concepto) y la propiedad de que el producto de los términos de una proporción ha de ser igual (conceptos) particularizando al ejemplo; representación de términos de la proporción, representa las operaciones e igualdad. - Identifica el número de personas a las que gusta el tenis, realizando una suma (procedimiento) y leyendo correctamente la tabla (procedimiento). - Identifica mediante otra suma $a+b+c+d$ el total de la muestra. - Aplica una regla de tres directa, por tanto el problema lo resuelve por un problema de proporción y no de probabilidad. - No divide correctamente, identificando un 1% como la diferencia de los casos favorables respecto a los posibles (600 de 700) (error de cálculo).

Observamos en el ejemplo que el alumno comienza realizando una regla de tres. Este procedimiento nos sugiere que el estudiante no interpreta el problema en forma probabilística, sino que lo trata como un problema de proporcionalidad. Es decir, no es claro que esté utilizando la idea de experimento aleatorio y probabilidad. Esta conjetura se ve reforzada por el hecho de que el estudiante nos dé el resultado en porcentaje y no como una probabilidad, no utilizando en ningún momento la notación probabilística. Por otro lado, obtiene una solución incorrecta por error en los cálculos.

C.1.3. *Calcula la probabilidad como complementaria a la del suceso contrario. Además comete un error en el cálculo del porcentaje.* Esta categoría es básicamente igual a la anterior, aunque el estudiante está aplicando correctamente una propiedad (que la suma de la probabilidad de un suceso y su contrario es igual a la unidad). En el siguiente ejemplo se observa que el alumno no sabe leer el dato de la tabla, no alcanzando un nivel de extracción de datos, que es la categoría más baja de lectura de tablas en la clasificación de Curcio (1989):

> "*100---700*
> *X=40---250*
> *40% de 250 son chicas, 5% no le gusta el tenis. La probabilidad de que sea chica y además le guste el tenis es del 35%. De 100 veces que tires 35 saldrá que es chica y le gusté el tenis*" (alumno 148).

Tabla 3.5.3. Análisis semiótico de la respuesta C1.3

Expresión	Contenido
100---700 *X=40---250* *40% 250 son chicas, 5% no le gusta el tenis*	- Descompone el problema en partes (proceso de análisis), fijando una variable (ser chica). - Aplica una regla de tres directa, para calcular el porcentaje de chicas en la muestra (procedimiento; observando que el producto de términos en la proporción es constante (propiedad y conceptos). - Resuelve el problema usando proporciones y no probabilidades. - Calcula el porcentaje de chicas correctamente; no es claro que esté interpretando el problema en términos probabilísticas. - Lee incorrectamente el porcentaje de chicas que no le gusta el tenis (*conflicto* en el proceso de interpretación de la tabla). - Redondea los cálculos (el resultado correcto es 35,7%).
La probabilidad de que sea chica y además le guste el tenis es del 35%. *De 100 veces que tires 35 saldrá que es chica y le gusté el tenis*"	- Aplica la probabilidad del suceso contrario (propiedad, particularización a este caso); usa los conceptos de suceso, suceso contrario y probabilidad. - Calcula el porcentaje de casos que cumplen *A* y *B* (chicas a las que no gusta el tenis, respecto al total); error en aplicación de procedimiento. - Calcula el porcentaje de chicas a las que gusta el tenis, restando los dos anteriores; obtiene un resultado incorrecto debido al error de cálculo.

A continuación se describen los diferentes tipos de respuestas incorrectas. Algunas de ellas repiten los errores descritos en los trabajos previos, incluidos los conflictos descritos por Estrada y Díaz (2007). Pero también se observan conflictos nuevos no descritos en las investigaciones anteriores.

C2. Confusión de la probabilidad que se pide en el problema

Una primera categoría de errores se produce debido a conflictos en el proceso de interpretación de la pregunta del problema, pues se confunde el tipo de probabilidad que se pide en cada caso. A continuación analizamos las variantes encontradas.

C2.1. *Confunde probabilidad simple y condicional*. Este conflicto ya fue descrito por Estrada y Díaz (2007). En el primer apartado del problema dan como respuesta una probabilidad condicional o bien en el tercer apartado dan como respuesta una probabilidad simple. En el ejemplo que reproducimos a continuación, y que se analiza en la Tabla 3.5.4, el estudiante obtiene dos probabilidades condicionales; la probabilidad de que, siendo chico le guste el tenis y la probabilidad condicional de que, siendo chica le guste el tenis. Siendo incapaz de leer en la tabla la frecuencia marginal, sustituye el problema original por dos problemas diferentes de cálculo de probabilidad condicional. Además del conflicto consistente en la confusión entre probabilidad simple y condicional, se observa la incapacidad de interpretar una tabla doble para encontrar una frecuencia marginal. El estudiante no alcanza el segundo nivel de lectura de datos definido por Curcio (1989), "leer entre los datos", pues no es capaz de encontrar datos más allá de los dados directamente en la tabla.

Primer apartado: "*Probabilidad de que le guste el tenis es:* $\begin{cases} Chico: 4/6 = 66,66\% \\ Chica: 2/6 = 33,33\% \end{cases}$" (alumno 36)

Tabla 3.5.4. Análisis semiótico de la respuesta C2.1

Expresión	Contenido
"Probabilidad de que le guste el tenis es: $\begin{cases} Chico: 4/6 = 66,66\% \\ Chica: 2/6 = 33,33\% \end{cases}$"	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - Proceso de representación, de los sucesos (chico/chica); representación incorrecta pues usa inadecuadamente el símbolo de igualdad. - *Conflicto* al dar una probabilidad mayor que 1 (no discrimina los conceptos probabilidad y porcentaje). - Identifica los casos favorables *a* y *b* correspondientes a las frecuencias marginales (conceptos y particularización). - Identifica los casos posibles $a+b$ o total marginal de los que les gusta el tenis (conceptos y particularización). - Aplica la regla de Laplace (procedimiento). - Intercambia el suceso que le guste el tenis por el suceso ser chico sabiendo que le gusta el tenis y por ser chica sabiendo que le gusta el tenis (*conflicto* al confundir probabilidad simple y compuesta). - No finaliza el problema pues no es capaz de aplicar la probabilidad total (propiedad y procedimiento).

C2.2. *Confunde probabilidad simple y compuesta.* En el primer apartado del problema da como respuesta una probabilidad compuesta o en el segundo dan como respuesta una probabilidad simple, confusión descrita por Estrada y Díaz (2007). En el ejemplo que reproducimos a continuación (Tabla 3.5.5), el alumno sustituye el cálculo de la probabilidad conjunta por dos problemas de cálculo de la probabilidad simple. Pero, en este caso, el alumno calcula dos probabilidades simples, la probabilidad de ser chica (aproximadamente 41,6% cuando en realidad es del 35,71%) y la probabilidad de que le guste el tenis (aproximadamente 80% cuando el resultado es 85,71%). Posiblemente la dificultad se deba a incapacidad para leer la tabla de doble entrada.

Segundo apartado: "*La probabilidad de que sea chica y le guste el tenis es 41,6% que es la* *probabilidad de ser chica y 80% que es la probabilidad de que le guste el tenis”* (alumno 92).

Tabla 3.5.5. Análisis semiótico de la respuesta C2.2

Expresión	Contenido
“La probabilidad de que sea chica y le guste el tenis es 41,6% que es la probabilidad de ser chica y 80% que es la probabilidad de que le guste el tenis”	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - *Conflicto* al identificar la probabilidad pedida; sustituye el cálculo de la probabilidad compuesta por cálculo de dos probabilidades simples. - Identifica los casos favorables $a+b$ y $c+d$ correspondientes a las frecuencias marginales (conceptos y particularización) para cada probabilidad simple. - Identifica los casos posibles $a+b$-c-d o total de la muestra (conceptos y particularización). - Aplica la regla de Laplace (procedimiento).

C2.3. *Confunde probabilidad condicional y compuesta.* En el segundo apartado del problema el estudiante da como respuesta una probabilidad condicional o en el tercer apartado da como respuesta una probabilidad compuesta, confusión encontrada por Ojeda (1995). Einhorn y Hogarth (1986) sugieren que los enunciados que usan la conjunción “*y*” pueden llevar a confundir la probabilidad conjunta y la probabilidad condicional. En el siguiente ejemplo, que analizamos en la Tabla 3.5.6, vemos como a la hora de calcular la probabilidad compuesta el alumno calcula una probabilidad condicionada, en este caso calcula la probabilidad de ser chica sabiendo que le gusta el tenis (200/600). El alumno ha hecho una restricción incorrecta de los casos posibles del experimento, considerando sólo los alumnos a los que gusta el tenis, en lugar de toda la muestra. Por ello, calcula una probabilidad condicional, en lugar de la conjunta.

“*Probabilidad de que sea chica y le guste el tenis es:* $2/6=33,33\%$ ” (alumno 143).

Tabla 3.5.6. Análisis semiótico de la respuesta C2.3.

Expresión	Contenido
“Probabilidad de que sea chica y le guste el tenis es: $2/6=33,33\%$ ”	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos) y los sucesos en el enunciado. - Identifica que la pregunta pide calcular una probabilidad pedida (proceso de interpretación) particularizando el concepto probabilidad al ejemplo. - Identifica los casos favorables al suceso compuesto “ser chica y gustarle el tenis”, *b*. - *Conflicto* al identificar como casos posibles *a-b* (frecuencia marginal), en lugar de $a+b+c+d$ (total de la muestra); restricción incorrecta del espacio muestral. - Aplica la regla de Laplace (procedimiento). - En lugar de calcular la probabilidad compuesta que le guste el tenis y sea chica calcula la condicionada o ser chica sabiendo que le gusta el tenis (conflicto al confundir una probabilidad condicional y conjunta).

C2.4. *Confunde una probabilidad condicional P(A/B) con su transpuesta P(B/A).* Serían los estudiantes que en el tercer apartado del problema dan como respuesta una probabilidad condicional errónea, al confundir el suceso condicionante y el condicionado. Estos estudiantes no discriminan adecuadamente entre las dos direcciones de la probabilidad condicional $P(A/B)$ y $P(B/A)$, error que Falk (1986) denomina *falacia de la condicional transpuesta*. En este ejemplo, que reproducimos a continuación, y analizamos en la Tabla 3.5.7, el alumno calcula la probabilidad de ser chica sabiendo que le gusta el tenis (200/600), por lo que está intercambiando el suceso condición que sea chica por el suceso que le guste el tenis.

Tercer apartado: *"Hay un 33% de probabilidad de que al elegir una chica le guste el tenis"* (alumno 71).

Tabla 3.5.7. Análisis semiótico de la respuesta C2.4.

Expresión	Contenido
"Hay un 33% de probabilidad de que al elegir una chica le guste el tenis"	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - *Conflicto* en el proceso de interpretación de la pregunta. - Intercambia el suceso que le guste el tenis siendo chica por el suceso ser chica sabiendo que le gusta el tenis (conflicto al confundir los sucesos de la probabilidad condicionada). - Identifica los casos favorables b (intersección de sucesos). - *Conflicto* al identifica los casos posibles $b+d$, tomando en vez de ello la frecuencia marginal $a+b$. - Aplica correctamente la regla de Laplace (procedimiento). - Expresa el resultado en porcentaje (proceso de representación y concepto).

C.2.5. *Confunde probabilidad simple, con probabilidad de suceso elemental.* No hemos encontrado descrito este conflicto en los trabajos previos. En lugar de calcular la probabilidad simple, $P(A)$, calcula la probabilidad de tomar al azar un sujeto entre todos los que componen la muestra. En el siguiente ejemplo, el estudiante considera todos los sujetos equiprobables, asignando a cada uno igual probabilidad (Tabla 3.5.8)

Primer apartado: *"La probabilidad de que le guste el tenis cogiendo un alumno al azar es de 1/700 ya que existen 700 alumnos"* (alumno 82).

Tabla 3.5.8. Análisis semiótico de la respuesta C2.5.

Unidad	Contenido
"La probabilidad de que le guste el tenis cogiendo un alumno al azar es de 1/700 ya que existen 700 alumnos"	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - *Conflicto* al interpretar la pregunta del enunciado, no identificando correctamente la probabilidad pedida (no identifica los casos favorables b, con el suceso "gustar el tenis"). - Considera todos los casos equiprobables; asigna igual probabilidad a cada individuo (propiedad). - Identifica los casos posibles $a+b+c+d$. - Aplica la regla de Laplace (procedimiento). - *Conflicto* al asumir equiprobabilidad.

C3. Confusión de los sucesos que intervienen en la probabilidad

Un segundo grupo de errores se producen por conflictos debidos a la identificación de los sucesos que intervienen en la probabilidad pedida. A continuación analizamos estos casos, que se deben a que el alumno no alcanza el nivel elemental de extracción de datos de la tabla de contingencia, en la terminología de Curcio (1989).

C3.1. *Confusión de un suceso y su complementario,* calculando correctamente la probabilidad pedida. En el caso que reproducimos a continuación, y analizamos en la Tabla 3.5.9, el alumno calcula $P(\bar{B}/A)$=50/250=20% en vez de $P(B/A)=$200/250=80%. La confusión se debe a que está calculando a partir del total de alumnas, aquellas a las que no les gusta el tenis en lugar de las que sí. Sin embargo, ha identificado correctamente que se trata de un problema de probabilidad conjunta y aplica la fórmula correctamente, salvo que confunde el suceso por su complementario.

Tercer apartado: "*La probabilidad de que le guste el tenis es* $\frac{50}{250} = 20\%$" (alumno 102).

Tabla 3.5.9. Análisis semiótico de la respuesta C3.1

Expresión	Contenido
"La probabilidad de que le guste el tenis es $\frac{50}{250}$=20% "	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - Identifica la probabilidad pedida (proceso de interpretación) particularizando al ejemplo. Aplica el concepto de probabilidad condicional (proceso de particularización). - *Conflicto* al confundir el suceso "le gusta el tenis" con su complementario, no alcanzando el nivel de extracción de datos al leer la tabla. - No identifica los casos favorables *b*, sino toma en su lugar *d* (casos favorables al complementario). - Identifica los casos posibles *b+d* que cumplen la condición ser chica (conceptos y particularización). Aplica la regla de Laplace (procedimiento).

C4. Confusión de otros objetos matemáticos

Los alumnos no sólo confunden entre si probabilidades o sucesos, sino otros objetos matemáticos, como analizamos a continuación.

C.4.1. *Confusión de probabilidades con casos posibles,* es decir de la probabilidad de un suceso con las frecuencias absolutas. Estos alumnos no discriminan propiedades elementales de la probabilidad, como que la probabilidad es un número incluido entre cero y uno. Este error también se encontró en algunos casos en el trabajo de Estrada y Díaz (2007). Un ejemplo (Tabla 3.5.10) referente al primer apartado sería : "$P(A) = 400 + 200 = 600$" (alumno 5). El mismo tipo de respuestas se ha encontrado en los otros apartados. Así, en el segundo apartado un alumno responde: "*La probabilidad de que le guste el tenis y sea chica es de 200*" (alumno 29). El alumno está dando el número de casos que cumple la condición. Ha podido leer correctamente la tabla, pero asigna el término probabilidad a la frecuencia absoluta confundiendo los dos conceptos.

Tabla 3.5.10. Análisis semiótico de la respuesta C4.1.

Expresión	Contenido
"*P(A)*=400+200= 600"	- Usa las ideas de experimento aleatorio y sucesos. - *Conflicto* al interpretar la pregunta sobre la probabilidad como pregunta sobre número de casos favorables. - Identifica los elementos que cumplen la condición $a+b$. - No identifica los casos posibles $a+b+c+d$. No usa la regla de Laplace. - *Conflicto* al confundir la probabilidad con la frecuencia absoluta. - Expresión simbólica inadecuada. - *Conflicto* al no discriminar un axioma de probabilidad.

C.4.2. *Obtención de probabilidades mayores que la unidad*. Los alumnos confunden los datos del problema, llegando en casos a obtener probabilidades mayores que la unidad, sin aplicar los axiomas de la probabilidad. En el ejemplo que reproducimos a continuación, y analizamos en la Tabla 3.5.11, el estudiante aplica la regla de Laplace, y posteriormente multiplica por 100, confundiendo probabilidad y porcentaje. Primer apartado: "*La probabilidad de que a un alumno le guste el tenis es de 600/7*" (Alumno 52). Esto indica que el alumno aprende la regla de Laplace de memoria y la aplica mecánicamente sin comprender su significado.

Tabla 3.5.11. Análisis semiótico de la respuesta C4.2.

Expresión	Contenido
"La probabilidad de que a un alumno le guste el tenis es de 600/7"	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - Identifica la probabilidad pedida (proceso de interpretación) particularizando al ejemplo. - Identifica los casos favorables $a+b$. - Identifica los casos posibles $a+b+c+d$. - Aplica la regla de Laplace (procedimiento). - Calcula el porcentaje ya que multiplica por cien la fracción 6/7. - *Conflicto* al confundir la fracción con el porcentaje ya que no aparece el término por ciento en ninguna parte.

C.4.3. *Confunde dependencia e independencia.* Un conflicto ya señalado por Estrada y Díaz (2007) es suponer independencia en los datos, aunque la dependencia sea patente en la tabla. Estos alumnos calculan la probabilidad de la intersección, como producto de las dos probabilidades $P(A)$ y $P(B)$, dando la solución $P(A \cap B) = \frac{250}{700}\frac{600}{700}$.

Esta respuesta implica que el estudiante recuerda la regla del producto de probabilidades, aunque no sabe discriminar su aplicación en los casos de dependencia e independencia de los datos. Un ejemplo en el segundo apartado (ver Tabla 3.5.12) es el siguiente:

"*La probabilidad de que sea chica y le guste el tenis es* $\frac{250}{700}\times\frac{600}{700}$ " (alumno 22).

Tabla 3.5.12. Análisis semiótico de la respuesta C4.3.

Expresión	Contenido
"La probabilidad de que sea chica y le guste el tenis es $\frac{250}{700}x\frac{600}{700}$ "	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos) - Identifica la probabilidad pedida (proceso de interpretación) particularizando al ejemplo. - Identifica los casos favorables $a+b$ y los casos posibles $a+b+c+d$ del suceso gustar el tenis. - Identifica los casos favorables $b+d$ y los casos posibles $a+b+c+d$ del suceso ser chica. - Aplica la regla de Laplace en cada una de las probabilidades (procedimiento). - Supone independencia a la hora de calcular la probabilidad de la intersección. - *Conflicto* al suponer que los sucesos ser chica y que le guste el tenis son independientes.

C5. Confusión de fórmulas de cálculo

Un tipo de confusión muy frecuente es la confusión entre procedimientos. El estudiante identifica la probabilidad que se pide calcular, pero, en lugar de asignar a esta probabilidad una fórmula correcta (un procedimiento adecuado de cálculo) asigna una incorrecta. Describimos a continuación los casos encontrados.

C.5.1. *Calcula* $P(A\cap B) = P(A\cap B)\, P\left(B/A\right)$. Estos alumnos son aquellos que a la hora de calcular la probabilidad conjunta utilizan un mal desarrollo de la descomposición de la probabilidad condicional, sustituyendo la probabilidad simple de A por la probabilidad compuesta. Reproducimos a continuación un ejemplo que analizamos en la Tabla 3.5.13. Conjeturamos que el alumno recuerda vagamente la fórmula de cálculo de la probabilidad compuesta en caso de sucesos dependientes (en este ejemplo los sucesos A y B lo son). Dicha fórmula, correctamente aplicada sería $P(A\cap B) = P(A)\, P(B/A)$. El conflicto aparece porque, en lugar de usar $P(A)$ como primer término de la fórmula utiliza $P(A\cap B)$, debido a una restricción incorrecta del espacio muestral en la probabilidad $P(A)$.

" $P(B) = \frac{200}{700}\frac{200}{250} = \frac{400}{175000}$ " (alumno 87).

Tabla 3.5.13. Análisis semiótico de la respuesta C5.1.

Expresión	Contenido
$P(B)=\frac{200}{700}\frac{200}{250}=\frac{400}{175000}$	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - Usa una notación incorrecta, pues expresa como probabilidad simple una probabilidad compuesta. - Identifica correctamente que los sucesos *A* y *B* son dependientes (particularización correcta de una propiedad) y trata de aplicar la regla del producto para sucesos dependientes. - Identifica la probabilidad pedida (proceso de interpretación) particularizando al ejemplo. - Identifica los casos posibles $a+b+c+d$ del suceso gustar el tenis y ser chica, pero tiene un conflicto al identificar los favorables (conceptos y procedimientos). - Identifica los casos favorables b y los casos posibles $b+d$ del suceso gustar el tenis siendo chica, es decir, para el cálculo de la probabilidad condicional. - Aplica la regla de Laplace en cada una de las probabilidades (procedimiento). - Multiplica las probabilidades. - *Conflicto* a la hora de utilizar la descomposición de la probabilidad conjunta como producto de la simple por la condicionada, el alumno conoce el procedimiento pero erra en los elementos ya que confunde la probabilidad simple por la compuesta.

C.5.2. *Calcula* $P(A\cap B)=N(A\cap B)/N(\overline{A}\cap B)$. Estos alumnos calculan la probabilidad conjunta dividiendo los casos favorables entre los que no lo son, como el ejemplo que reproducimos a continuación y analizamos en la Tabla 3.5.14. El alumno identifica correctamente la pregunta del problema y usa la notación adecuada a la probabilidad producto. También identifica correctamente los casos favorables, pero en lugar de dividir por los posibles, usa los desfavorables. Esto indica un aprendizaje memorístico de la regla de Laplace sin comprender su significado.

"$P(A\cap B)=\frac{casos\ \ favorables\ de\ A}{casos\ \ favorables\ de\ \overline{A}}=\frac{200}{400}$" (alumno 54).

Tabla 3.5.14. Análisis semiótico de la respuesta C5.2.

Expresión	Contenido
$P(A\cap B)=\frac{casos\ \ favorables\ de\ A}{casos\ \ favorables\ de\ \overline{A}}=\frac{200}{400}$	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - Identifica la probabilidad pedida (proceso de interpretación) particularizando al ejemplo. - Aplica una notación correcta de la probabilidad producto (proceso de representación). - Identifica los casos favorables al suceso producto b. - *Conflicto* al identificar los casos posibles del suceso gustar el tenis y ser chica. En su lugar Toma como tales los elementos complementarios a los favorables a. (casos desfavorables). - Aplica la regla de Laplace en cada una de las probabilidades (procedimiento).

C.5.3. *Calcular* $P(A \cap B) = P(A) \times P(A / B)$. El alumno recuerda que la intersección es el producto de la probabilidad simple de un suceso por la condicional, pero confunde los términos de la probabilidad condicional en la fórmula. En el apartado segundo, que reproducimos a continuación y se analiza en la Tabla 3.5.15, aparece el siguiente ejemplo:

"*La probabilidad de que sea chica y le guste el tenis es:*
$P(B) = P(Chica)P(chica / gusta\,tenis) = \frac{250}{700}\frac{200}{600} = \frac{5}{42}$" (alumno 93).

Tabla 3.5.15. Análisis semiótico de la respuesta C5.3.

Expresión	Contenido
"*La probabilidad de que sea chica y le guste el tenis es* $P(B)=P(Chica)P(chica/gusta\,tenis)=\frac{250}{700}\frac{200}{600}=\frac{5}{42}$	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - Identifica la probabilidad pedida (proceso de interpretación) particularizando al ejemplo. - Recuerda y trata de aplicar la regla del producto para sucesos dependientes; identifica correctamente que los sucesos son dependientes (propiedades y procedimientos). - Identifica los casos favorables *b* y los casos posibles *a*+*b* del suceso ser chica gustándole el tenis. - Identifica los casos favorables *b*+*d* y los casos posibles *a*+*b*+*c*+*d* del suceso ser chica. - Aplica la regla de Laplace en cada una de las probabilidades (procedimiento). - *Conflicto* al confundir una probabilidad condicional y su transpuesta.

C.5.4. *Calcular* $P(A \cap B) = P(A) \times P(A / B)$ y suponer que el porcentaje de chicos y chicas es 1/2. El alumno no lee los datos de la tabla sobre la proporción de chicos y chicas. Una vez calculada la probabilidad producto, la multiplica por la proporción supuesta de chicas (1/2) sin darse cuenta que en esta probabilidad producto ya se tiene en cuenta dicha proporción. Un ejemplo, en que además hay un uso incorrecto e la simbología matemática, lo reproducimos a continuación y analizamos en la Tabla 3.5.16 se da en el segundo apartado:

"*La probabilidad de ser chica y le guste el tenis es* $P = \frac{1}{2} \cup \frac{200}{600}$" (alumno 165).

Tabla 3.5.16. Análisis semiótico de la respuesta C5.4.

Expresión	Contenido
"La probabilidad de ser chica y le guste el tenis es $P = \frac{1}{2} \cup \frac{200}{600}$"	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - Identifica la probabilidad pedida (proceso de interpretación) particularizando al ejemplo. - Identifica los casos favorables *b* y los casos posibles *a*+*b* del suceso ser chica gustándole el tenis.

	- Aplica la regla de Laplace en la probabilidad producto (procedimiento). - Multiplica esta probabilidad por ½ (proporción supuesta de chicas) sin darse cuenta que la proporción no es esta y ya se ha tenido en cuenta en el producto. - *Conflicto* de notación (unión por intersección; además, no puede aplicar la unión al resultado numérico)

C.5.5. *Calcula* $P(A \cap B) = P(A/B) \times P(A \cap B)$, *confundiendo la probabilidad compuesta con la simple al utilizar la regla del producto*. Error que reproducimos y analizamos en la Tabla 3.5.17. A continuación un ejemplo del segundo apartado, donde el estudiante, trata de aplicar la regla del producto para el caso de sucesos dependientes:

"*En el colegio hay un total de 250 chicas por 700 alumnos, y a 200 de ellas les gusta el tenis. Por lo tanto, si elegimos al azar uno de los alumnos, la probabilidad de que sea chica y le guste el tenis es* $\frac{200}{250}\frac{200}{700} = \frac{5}{49}$" (Alumno 183).

Tabla 3.5.17. Análisis semiótico de la respuesta C5.5.

Expresión	Contenido
"En el colegio hay un total de 250 chicas por 700 alumnos, y a 200 de ellas les gusta el tenis. Por lo tanto, si elegimos al azar uno de los alumnos, la probabilidad de que sea chica y le guste el tenis es $\frac{200}{250}\frac{200}{700} = \frac{5}{49}$"	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - Identifica la probabilidad pedida (proceso de interpretación) particularizando al ejemplo. - Identifica que se trata de sucesos dependientes, trata de aplicar la regla del producto. - Identifica los casos favorables *b* y los casos posibles *b+d* del suceso gustar el tenis siendo chica. - Identifica los casos favorables *b* y los casos posibles *a+b+c+d* del suceso ser chica y le gusta el tenis. - Aplica la regla de Laplace en cada una de las probabilidades (procedimiento). - Supone independencia y multiplica las probabilidades de ambos sucesos. - *Conflicto* al calcular la probabilidad conjunta como producto de la probabilidad conjunta por la probabilidad condicionada de gustarte el tenis siendo chica.

C.5.6. *Calcula* $P(A \cap B) = P(A) + P(A \cap B)$, es decir *aplica la regla del producto incorrectamente y suma en lugar de multiplicar*. Esta confusión se reproduce a continuación, y se analiza en la Tabla 3.5.18. Un ejemplo del segundo apartado (alumno 171), que indica un aprendizaje memorístico de la regla del producto, sin comprensión de su significado es el siguiente:

"2,5/7+2/5=4,5/7 probabilidad de ser chica y además le guste el tenis
7—100
4,5—x
La probabilidad es del 64,3 y representado en fracciones 4,5 niños de cada 7".

Tabla 3.5.18. Análisis semiótico de la respuesta C5.6.

Expresión	Contenido
"2,5/7+2/5=4,5/7 probabilidad de ser chica y además le guste el tenis 7—100 4,5—x La probabilidad es del 64,3 y representado en fracciones 4,5 niños de cada 7"	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - Identifica la probabilidad pedida (proceso de interpretación) particularizando al ejemplo. - Identifica los casos favorables *b+d* y los casos posibles *a+b+c+d* del suceso ser chica. - Identifica los casos favorables *b* y los casos posibles *a+b+c+d* del suceso ser chica y le guste el tenis. - Aplica la regla de Laplace en cada una de las probabilidades (procedimiento). - Utiliza la descomposición de la probabilidad conjunta a partir de la probabilidad condicionada y la simple. - Suma las probabilidades simples y conjuntas para calcular la probabilidad conjunta. - Conflicto al confundir el operador de la descomposición ya que suma en vez de multiplicar.

C.5.7. *Calcular* $P(A/B)=P(A)/P(B)$. En este caso, el alumno recuerda que la probabilidad condicional viene dada por un cociente, pero hay un conflicto consistente en confundir el numerador de dicho cociente. En este ejemplo que reproducimos a continuación, y analizamos en la Tabla 3.5.19, el alumno a la hora de calcular la probabilidad condicional divide la frecuencia marginal de ser chica entre la frecuencia marginal de que le guste el tenis.

"La probabilidad de que le guste el tenis sabiendo que es chica es 250/600".

Tabla 3.5.19. Análisis semiótico de la respuesta C5.7.

Expresión	Contenido
"La probabilidad de que le guste el tenis sabiendo que es chica es 250/600"	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - Identifica la probabilidad pedida (proceso de interpretación) particularizando al ejemplo. - No identifica los casos favorables b, toma como casos favorables *b+d*, la frecuencia marginal del suceso ser chica. - No identifica los casos posibles *a+b+c+d*, toma como casos favorables *a+b*, la frecuencia marginal del suceso ser gustar el tenis. - Aplica la regla de Laplace en cada una de las probabilidades (procedimiento). - *Conflicto al tomar* como probabilidad conjunta el cociente entre las frecuencias marginales de cada uno de los sucesos.

C.5.8. *Calcula* $P(A/B)=\dfrac{N(A\cap B)}{N(\bar{A}\cap B)}$, es decir confunde la probabilidad condicional con el cociente entre el número de chicas a las que les gusta y no les gusta el tenis. Indica una comprensión incompleta, tanto del concepto de probabilidad condicional, como de la regla de Laplace y una restricción incorrecta del espacio muestral.

Reproducimos a continuación y analizamos en la Tabla 3.5.20 un ejemplo del tercer apartado:

"La probabilidad de que le guste el tenis siendo chica es 200/400" (alumno 76).

Tabla 3.5.20. Análisis semiótico de la respuesta C5.8.

Expresión	Contenido
"La probabilidad de que le guste el tenis siendo chica es 200/400"	- Usa las ideas de experimento aleatorio, caso favorable y posible (conceptos). - Identifica la probabilidad pedida (proceso de interpretación) particularizando al ejemplo. - Identifica los casos favorables *b*. - *Conflicto* al identificar los casos posibles, tomando la frecuencia de los que no le gusta el tenis y son chicos. - Aplica la regla de Laplace en cada una de las probabilidades (procedimiento). - Conflicto al tomar como probabilidad condicionada el cociente entre las chicas que les gusta el tenis y los chicos que les gusta el tenis, debido a confusión entre casos posibles y desfavorables.

C.5.9. *Calcula la probabilidad de un suceso más el de su complementario obteniendo la unidad.* Esta respuesta indica por un lado, incomprensión del concepto de probabilidad y sus propiedades, pues la probabilidad de una parte del espacio muestral no puede ser igual a la unidad. Y por otro lado, tampoco se comprende el segundo axioma, pues la probabilidad de un suceso y su complementario siempre será la unidad. Reproducimos a continuación (Tabla 3.5.21) un ejemplo del primer apartado:

"*La probabilidad de que le guste el tenis es* $\frac{400}{600} + \frac{200}{600} = \frac{600}{600}$" (alumno 164).

Tabla 3.5.21. Análisis semiótico de la respuesta C5.9.

Expresión	Contenido
"La probabilidad de que le guste el tenis es $\frac{400}{600} + \frac{200}{600} = \frac{600}{600}$"	- Usa las ideas de experimento aleatorio, caso favorable y posible, probabilidad (conceptos). - Identifica la probabilidad pedida (proceso de interpretación) particularizando al ejemplo. - Identifica los casos favorables y posibles como *a+b*. - *Conflicto* al calcular la probabilidad simple como la suma de la probabilidad condicionada de ser chico gustándole el tenis más la de ser chica gustándole el tenis; supone una aplicación incorrecta de la probabilidad total (propiedad y procedimiento) - Identifica los casos favorables a cada una de las probabilidades condicionales como *a* y *b* y los posibles como *a+b*. - Aplica la regla de Laplace en cada una de las probabilidades (procedimiento).

C.5.10. *Otras fórmulas sin sentido.* En ocasiones la fórmula no tiene sentido, como el siguiente ejemplo del segundo apartado, que no analizamos al ser difícil identificar el razonamiento: "*La probabilidad de que sea chica y le guste el tenis es 2/6-4/5*".

C6. No finaliza el problema

Incluimos en este apartado los casos en que el alumno comienza el problema pero no lo finaliza, lo cual puede deberse a varios motivos, que se describen en lo que sigue.

C.6.1. *Interpreta bien los sucesos pero no calcula las probabilidades*. En este caso el alumno dibuja un diagrama de árbol con todas las soluciones posibles pero no calcula ninguna probabilidad. Reproducimos a continuación, y analizamos en la Tabla 3.5.22, un ejemplo del primer apartado, donde debido a una incorrecta construcción del diagrama en árbol, el estudiante no llega a finalizar el problema. El problema surge porque la primera división en ramas del árbol no constituye una partición del espacio muestral (alumno 16).

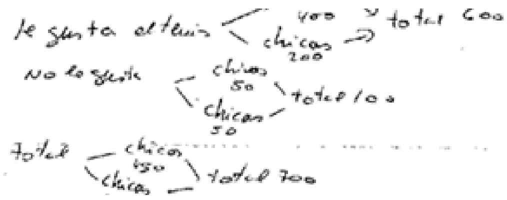
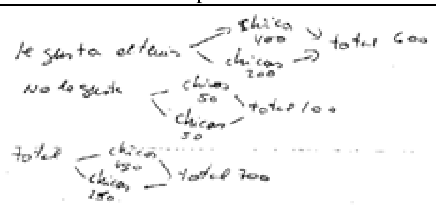


Tabla 3.5.22. Análisis semiótico de la respuesta C6.1.

Expresión	Contenido
	- Usa las ideas de experimento aleatorio, suceso (conceptos). - Identifica los sucesos en el enunciado (particulariza al ejemplo). - Calcula el número de casos favorables de los diferentes sucesos: gustar el tenis, no le gusta el tenis y sexo de los alumnos (conceptos y procedimiento). - Conflicto al realizar una representación incorrecta del diagrama en árbol, separando las variables A y B en diagramas diferentes, lo cual no es productivo para encontrar la solución; es decir, conflicto al realizar una partición del espacio muestral.

C.6.2. *No calcula las probabilidades sino que las compara con el 50%.* Reproducimos un ejemplo del primer apartado, en el que el estudiante, aparentemente no recuerda la regla de Laplace, aunque parece leer al menos las frecuencias marginales de la tabla. Lo analizamos en la Tabla 3.5.23:

> "*La probabilidad de que le guste el tenis es más de un 50% ya que el número total de los que le gustan el tenis es mayor de los que no les gustan*" (alumno 64).

Tabla 3.5.23. Análisis semiótico de la respuesta C6.2

Expresión	Contenido
"La probabilidad de que le guste el tenis es más de un 50% ya que el número total de los que le gustan el tenis es mayor de los que no les gustan"	- Usa las ideas de experimento aleatorio (conceptos). - Identifica los sucesos (concepto y procedimiento). - Compara los sucesos ser chica gustando el tenis y ser chico gustando el tenis; lee las frecuencias marginales (procedimientos). - Hay más chicos que les gusta el tenis que chicas que les gusta el tenis; resuelve que es más probable que le guste a los chicos. - No recuerda la definición de probabilidad o bien no recuerda la fórmula de Laplace.

C.6.3. *Incorrecta lectura de la tabla.* En el ejemplo, que reproducimos a continuación, y analizamos en la Tabla 3.5.24, el alumno 174 calcula la frecuencia de los que les gusta el tenis restando a los que les gustan los que no. Por tanto no es capaz de alcanzar el nivel de extracción de datos en la clasificación de Curcio (1989).

"$P(T)-P(\bar{T})=P(600)-P(100)=500$ *Por tanto la probabilidad de que le guste el tenis es de* $\frac{500}{700}$*; 500 alumnos de 700*".

Tabla 3.5.24. Análisis semiótico de la respuesta C6.3.

Expresión	Contenido
"$P(T)-P(\bar{T})$ =P(600)-P(100) = 500. La probabilidad de que le guste el tenis es de 500 alumnos de 700"	- Usa las ideas de experimento aleatorio y sucesos (conceptos) - Identifica los sucesos. - Conflicto en el cálculo de la frecuencia absoluta de los que les gusta el tenis. Recta a los que les gusta el tenis los 100 chicos que no les gusta. - Toma como casos favorables el resto anterior. - Identifica como casos posibles $a+b+c+d$. - *Conflicto* en la notación, al aplicar la probabilidad al valor de la frecuencia absoluta. - Expresa la probabilidad como una razón; conflicto en la definición de probabilidad.

3.5.3. RESULTADOS EN EL CÁLCULO DE PROBABILIDAD SIMPLE

Una vez categorizadas las respuestas en este apartado, exponemos las soluciones dadas por los futuros profesores a la tarea 1 en la Tabla 3.5.25. El 59% de los alumnos proporciona respuestas correctas en este apartado, porcentaje que es bajo, dada la simplicidad de la tarea y tratándose de futuros profesores. Díaz y Estrada (2006) obtienen 75% de respuestas correctas en un ítem muy similar a este con 65 futuros profesores de la Universidad de Lleida.

A pesar de la mayor facilidad del ítem (en nuestro caso se evita la negación en el enunciado), los resultados son peores, lo que explicamos porque los estudiantes de Díaz y Estrada seguían un curso optativo de Didáctica de la Estadística. Por su parte Díaz

(2007) obtiene, en el mismo ítem de Estrada, una proporción de aciertos del 84% en 590 estudiantes de psicología de cuatro universidades españolas. Aunque en la muestra de Díaz (2007) la mitad estudiaron el tema, esto es una indicación de la necesidad en incidir más en la formación de los profesores, que con una enseñanza adecuada pudieran salir con una mejor preparación de nuestras facultades.

Tabla 3.5.25. Soluciones dadas en la tarea 1

	Soluciones dadas por los alumnos	Frecuencia	Porcentaje
Básicamente correcto	Responde correctamente	108	59,0
	Plantea bien la formula pero tiene errores de cálculo	1	0,5
	Respuesta correcta sin escribir la formula	10	5,5
	Expresión simbólica incorrecta de la probabilidad	1	0,5
Confunde probabilidades	Confunde probabilidad simple y condicional	1	0,5
	Confunde probabilidad simple con probabilidad de suceso elemental	7	3,9
Confunde objetos	Confunde probabilidades con casos posibles	7	3,9
	Obtención de probabilidades mayores que la unidad	2	1,1
Confunde fórmulas	Calcula $P(A \cap B) = N(A \cap B) / N(\overline{A} \cap B)$	1	0,5
	Calcula la probabilidad de un suceso más el de su complementario obteniendo la unidad.	2	1,1
Otros errores	No calcula las probabilidades sino compara con el 50%	1	0,5
No finaliza el problema	Interpreta bien los sucesos pero no calcula las probabilidades	2	1,1
	Incorrecta lectura de la tabla	1	0,6
	Escribe un número sin sentido	5	2,7
	No contesta	34	18,6
	Total	183	100,0

La mayor frecuencia de errores en el cálculo de la probabilidad simple se deben a confusión de probabilidad con frecuencia absoluta, un error que, aunque con poca frecuencia de aparición, nos parece importante, dado que implica la no apreciación de los axiomas de probabilidad. También encontramos el error de dar la probabilidad de un elemento al azar de la población, que indica que el estudiante es incapaz de leer la tabla de doble entrada. Otros errores son dar la probabilidad mayor que la unidad, o errores de cálculo. Destacan también los casos en que el resultado es un número sin sentido y en un caso tenemos una confusión de la probabilidad simple con una condicional.

Se resumen las respuestas en la Tabla 3.5.26, donde el 23% de los alumnos no acaba el problema, el 65,6% lo hace casi correctamente y el resto tiene errores de varios tipos. También hemos estudiado el tipo de respuesta que daban los futuros profesores en función de si la solución estaba dada en porcentajes, fracción, fracción y porcentajes o un decimal entre cero y uno.

Tabla 3.5.26. Resumen de las respuestas de la tarea 1

Soluciones dadas por los alumnos	Frecuencia	Porcentaje
Básicamente correctas	120	65,6
Confunde probabilidades	8	4,4
Confunde objetos	9	4,9
Confunde formulas	3	1,6
Otros errores	1	0,5
No finaliza el problema	42	23,0
Total	183	100,0

Como vemos en la Tabla 3.5.27 el 14,2% de los alumnos responden a la tarea de probabilidad simple mediante un porcentaje. No queda claro si estos estudiantes han interpretado el problema en forma probabilística, aunque luego se podrá ver en la segunda parte cuando analizamos los objetos matemáticos identificados por los futuros profesores en el ejercicio. Sólo 8 alumnos en toda la muestra utilizan decimales, lo que indica la mayor dificultad de comprensión de este sistema numérico y la mayor dificultad para interpretar la probabilidad como un número entre 0 y 1 en lugar de como una fracción. El resto usa fracciones o una combinación de las anteriores notaciones.

Tabla 3.5.27. Formato tarea 1

Soluciones dadas por los alumnos	Frecuencia	Porcentaje
Contesta con un decimal entre 0 y 1	3	1,6
Porcentajes	26	14,2
Fracción y %	38	20,8
Fracción	67	36,6
Resuelve sin utilizar operaciones	10	5,5
Fracción y decimal	4	2,2
Decimal y %	1	0,5
No contesta	34	18,6
Total	183	100,0

3.5.4. RESULTADOS EN EL CÁLCULO DE PROBABILIDAD CONJUNTA

En la Tabla 3.5.28 se presenta los resultados en el cálculo de la probabilidad conjunta. En este apartado el porcentaje de respuestas correctas no llega al 40% de la muestra. Estrada y Díaz (2007) obtienen un 52% de respuestas correctas, también mucho menor que en la primera pregunta aunque mejores que los nuestros. Díaz (2007) obtiene 53% de respuestas correctas en alumnos de psicología.

Mayor variedad de errores aparecen en las respuestas, respecto al cálculo de la probabilidad simple, destacando la confusión de probabilidades condicionales y conjuntas ya encontrado por Einhorn y Hogarth (1986) y Ojeda (1995) y Estrada y Díaz (2007). Agrupando los resultados, el 24,9% de los estudiantes confunde unas probabilidades con otras. Aparece también un error no citado en la literatura consistente en calcular la probabilidad de obtener un sujeto del total de la población o bien con las condiciones dadas (chica a la que gusta el tenis). Es decir, dividir la unidad por el número de casos favorables para hallar la probabilidad, en lugar de aplicar la regla de Laplace. Destacamos también que 10 estudiantes confunden probabilidad simple y compuesta un error que no fue encontrado por Díaz y de la Fuente (2007). Díaz (2007) no informa sobre la naturaleza de los errores en su muestra de estudiantes de psicología.

Tabla 3.5.28. Soluciones dadas en la tarea 2

	Soluciones dadas por los alumnos	Frecuencia	Porcentaje
Básicamente correctos	Responde correctamente	69	37,8
	Respuesta correcta sin escribir la formula	5	2,8
	Plantea bien la formula pero tiene errores de cálculo	1	0,5
Confunde probabilidades	Calcula 1- *P*(suceso contrario) y tiene error de cálculo	2	1,1
	Confunde probabilidad simple y condicional	1	0,5
	Confunde probabilidad simple y compuesta	10	5,6
	Confunde probabilidad condicional y compuesta	25	13,8
	Confunde $P(A/B)$ con $P(B/A)$	1	0,5
	Confunde probabilidad simple con probabilidad de suceso elemental	6	3,4
Confunde objetos	Confunde probabilidades con casos posibles	5	2,8
	Suponer independencia en los datos	3	1,7
	Confunde unión e intersección y asume incompatibilidad	2	1,1
Confunde sucesos	Confunde un suceso y su complementario, correctamente la probabilidad pedida	2	1,1
Confunde fórmulas	Calcula $P(A/B)=\frac{N(A\cap B)}{N(\bar{A}\cap B)}$	2	1,1
	Calcula $P(A\cap B)=P(A)\times P(A/B)$	1	0,5
	Calcula $P(A\cap B)=P(A/B)\times P(A\cap B)$	1	0,5
	Calcula $P(A\cap B)=P(A)+P(A\cap B)$	1	0,5
	Calcula $P(A\cap B)=P(A)\times P(A/B)$ y supone que el porcentaje de chicos y chicas es ½.	1	0,5
	Fórmula sin sentido	1	0,5
Otros errores	Obtención de probabilidades mayores que la unidad	3	1,7
	No calcula las probabilidades sino compara con el 50%	1	0,5
No finaliza	Escribe un número sin sentido	6	3,4
	No contesta	33	18,1
	Total	183	100,0

Tabla 3.5.29. Resumen de las respuestas de la tarea 2

Soluciones dadas por los alumnos	Frecuencia	Porcentaje
Básicamente correctos	75	41,1
Confunde probabilidades	46	24,9
Confunde objetos	10	5,6
Confunde sucesos	2	1,1
Confunde formulas	7	3,6
Otros errores	4	2,2
No finaliza el problema	39	21,5
Total	183	100,0

Los resultados resumidos en la Tabla 3.5.29 indican que un 21,5% de los estudiantes no son capaces de finalizar el problema, un 25,1% confunde distintos tipos de probabilidad, el 41,1% resuelve el problema y el resto hace otros errores, como confundir objetos, fórmulas o los sucesos del enunciado. Respecto al tipo de formato que utilizan en la solución (Tabla 3.5.30) observamos de nuevo el uso exclusivo de la notación porcentual en un 14% y el escaso uso de la notación decimal. Sólo cuatro encuestados resolvieron el problema exponiendo el resultado como un decimal entre cero y uno.

Tabla 3.5.30. Formato tarea 2

Soluciones dadas por los alumnos	Frecuencia	Porcentaje
Contesta con un decimal entre 0 y 1	4	2,2
Porcentajes	26	14,2
Fracción y porcentajes	35	19,1
Fracción	68	37,2
Resuelve sin utilizar operaciones	12	6,6
Fracción y decimal	3	1,6
Decimal y porcentajes	1	0,5
No contesta	34	18,6
Total	183	100,0

3.5.5. RESULTADOS EN EL CÁLCULO DE PROBABILIDAD CONDICIONAL

Finalmente presentamos las respuestas en el cálculo de probabilidad condicional (Tabla 3.5.34), con un porcentaje de respuestas correctas similar al del segundo apartado y de nuevo inferior al obtenido por Díaz y Estrada (2006) quienes obtuvieron un 56% de respuestas correctas. Díaz obtiene un 52% de respuestas correctas en estudiantes de psicología. La principal confusión en este apartado se produce de nuevo entre probabilidad condicional y compuesta, coincidiendo con Einhorn y Hogarth (1986) y Ojeda (1995) y Estrada y Díaz (2007). Encontramos de nuevo algunos alumnos que confunden, la probabilidad pedida con probabilidad del suceso elemental, error que como hemos indicado no se ha descrito anteriormente. Aparecen diversos errores producidos al intentar aplicar las fórmulas directamente sin una comprensión adecuada de las mismas.

Los resultados se resumen en la Tabla 3.5.35, donde vemos que la cuarta parte de los estudiantes no es capaz de abordar este problema a pesar de su sencillez, el 16% confunde diversos tipos de probabilidad, y se producen con menor frecuencia otros errores, como confusión de probabilidad y casos posibles o confusión de sucesos.

Tabla 3.5.34. Soluciones dadas en la tarea 3

	Soluciones dadas por los alumnos	Frecuencia	Porcentaje
Básicamente correctas	Responde correctamente	77	42,1
	Plantea bien la formula pero tiene errores de cálculo	3	1,6
Confunde probabilidades	Confunde probabilidad simple y condicional	2	1,1
	Confunde probabilidad condicional y compuesta	14	7,7
	Confunde *P(A/B)* con *P(B/A)*	9	4,9
	Confunde probabilidad simple, con probabilidad de suceso elemental	5	2,8
Confunde sucesos	Confusión de un suceso y su complementario, calculando correctamente la probabilidad pedida	2	1,1
Confunde objetos	Confunde probabilidades con casos posibles	4	2,1
	Confunde unión e intersección, asumiendo incompatibilidad	1	0,5
Confunde fórmulas	Calcula *P(A/B)= P(A)/P(B)*	5	2,8
	Calcula $P(A/B)=\dfrac{N(A\cap B)}{N(\bar{A}\cap B)}$	1	0,5
	Calcula $P(A\cap B)=\dfrac{N(A\cap B)}{N(\bar{A}\cap B)}$	2	1,1
Otros errores	Obtención de probabilidades mayores que la unidad	2	1,1
	No calcula las probabilidades sino compara con el 50%	1	0,5
	contesta bien pero error en formula	1	0,5
	Resultado numérico correcto y formula sin sentido	1	0,5
	Escribe un numero sin sentido	6	3,4
No finaliza el problema	Contesta bien pero no escribe la formula	4	2,2
	no contesta	43	23,5
	Total	183	100,0

Tabla 3.5.35. Resumen de las respuestas de la tarea 3

Soluciones dadas por los alumnos	Frecuencia	Porcentaje
Básicamente correctos	80	43,7
Confunde probabilidades	30	16,5
Confunde objetos	5	2,6
Confunde sucesos	2	1,1
Confunde formulas	8	4,4
Otros errores	11	6,0
No finaliza el problema	47	25,7
Total	183	100,0

En la Tabla 3.5.36 se analiza la forma representar la solución que utilizan: sólo siete estudiantes usan en sus resultados números decimales.

Tabla 3.5.36. Formato tarea 3

Soluciones dadas por los alumnos	Frecuencia	Porcentaje
Contesta con un decimal entre 0 y 1	2	1,1
Porcentajes	23	12,6
Fracción y porcentaje	38	20,8
Fracción	62	33,9
Resuelve sin utilizar operaciones	9	4,9
Fracción y decimal	4	2,2
Decimal y porcentaje	1	0,5
No contesta	44	24,0
Total	183	100,0

3.5.6. RESUMEN DE RESPUESTAS

En la Tabla 3.5.37 hemos clasificado estos errores por los diferentes conflictos semióticos involucrados en las tareas pedidas. Las frecuencias corresponden a la suma de los conflictos identificados en el total de las cuatro tareas propuestas, teniendo también en cuenta que, a veces, una respuesta conlleva más de un conflicto semiótico. Por ejemplo, cuando en el cálculo de la probabilidad condicional *P(B/A)* el alumno da como respuesta la probabilidad conjunta $P(\overline{A} \cap B)$, un primer conflicto es confundir el condicionamiento con intersección, y un segundo confundir un suceso con su complementario. Al dividir el total de conflictos entre el número de estudiantes encontramos una media de 1,7 conflictos por estudiante en las tres tareas, es decir, en promedio, cada estudiante encuentra un conflicto en la mitad de las tareas propuestas.

Tabla 3.5.37. Frecuencia de tipos de conflictos semióticos identificados

	n=155
Confunde una probabilidad condicional y su inversa	45
Confunde probabilidad condicional y conjunta	77
Confunden un suceso y su complementario	23
Supone independencia	3
Confunde probabilidad simple con condicional o conjunta	11
Confunde unión e intersección	4
Probabilidad mayor que 1	18
Calcula casos posibles	15
Confunde la fórmula de la probabilidad condicional con *P(A)/ P(B)*	5
Errores cálculo o no identifica datos	30
No responde	45
Total conflictos en las cuatro tareas y total de muestra	262

Observamos que los errores consistentes en confundir una probabilidad condicional con una conjunta o confundir una probabilidad condicional con su inversa, citados en las investigaciones anteriores, son los más frecuentes, pero aun así, no explican por si mismos los altos porcentajes de errores (cercanos al 40%) en las tareas de cálculo de probabilidad condicional y compuesta. Hay un número apreciable de conflictos consistentes en confundir un suceso con su complementario y otros donde, aunque da correctamente la fórmula de cálculo, hay errores de cálculo, o no son capaces de identificar los datos en el enunciado. Con menor frecuencia se presentan otros conflictos referidos a la confusión de frecuencia y probabilidad o invertir la fórmula de cálculo de probabilidades, obteniendo valores mayores que uno. Finalmente algunos estudiantes no son capaces de dar una respuesta.

Tabla 3.5.38. Resumen de las soluciones dadas por los futuros profesores

	$P(A)$		$P(A \cap B)$		$P(A / B)$	
Soluciones dadas por los alumnos	Frec.	%	Frec.	%	Frec.	%
Básicamente correctos	120	65,5	75	41,1	80	43,7
Confunde probabilidades	8	4,4	46	24,9	30	16,4
Confunde objetos	9	4,9	10	5,6	5	2,6
Confunde sucesos	0	0	2	1,1	2	1,1
Confunde formulas	3	1,6	7	3,6	8	4,4
Otros errores	1	0,6	4	2,2	11	5,9
No finaliza el problema	42	23,0	39	21,5	47	25,9
Total	183	100,0	183	100,0	183	100,0

Para comparar más fácilmente los tres apartados, en la Tabla 3.5.38 se presentan un resumen conjunto. Como se observa el porcentaje de respuestas básicamente correctas, en el caso de probabilidad simple, es superior que el de las otras probabilidades pedidas, en un 24,6 % y un 21,8% respectivamente. El porcentaje de los que no finalizan es parecido, algo mayor en la probabilidad condicional. Es en este caso donde hay mayor confusión sobre la probabilidad pedida.

3.6. CONOCIMIENTO ESPECIALIZADO DEL CONTENIDO

3.6.1. INTRODUCCIÓN

En la segunda parte del ítem se pregunta por los objetos matemáticos utilizados para resolver los problemas propuestos de la primera parte, diferenciándolos en problemas, lenguaje, conceptos, propiedades, procedimientos y argumentos. En el apartado 3.4.1 se realizó un análisis semiótico de los objetos y procesos matemáticos implícitos en las soluciones correctas a cada uno de los tres apartados de la primera parte de la tarea (cálculo de la probabilidad simple, compuesta y condicional).

Tabla 3.6.1. Objetos matemáticos implícitos en la actividad planteada

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Calcular una probabilidad simple	- ¿Cuál es la probabilidad de que sea chica?
	- Calcular una probabilidad compuesta	- ¿Cuál es la probabilidad de que sea chica y que le guste el tenis?
	- Calcular una probabilidad condicional	- Sabiendo que el alumno es una chica, ¿cuál es la probabilidad de que le guste el tenis?
	- Formar el espacio muestral	- Calcular el conjunto de resultados posibles para la variable aleatoria y sus restricciones
Lenguaje	- Verbal	- Explicación de la situación
	- Simbólico	- Símbolos que representan los sucesos y resultados
	- Tabla	- Situación espacial de los sucesos
	- Gráfico	- Representación icónica del juego
Conceptos	- Probabilidad simple	- Proporción de casos favorables y posibles
	- Probabilidad compuesta	- Proporción de ocurrencia dos o más sucesos simultáneos respecto a los posibles
	- Probabilidad condicional	- Proporción de ocurrencia suceso respecto a la ocurrencia de otro
	- Experimento aleatorio	- Experimento que bajo las mismas condiciones iniciales, puede presentar resultados diferentes
	- Espacio muestral	- Conjunto de todos los posibles resultados individuales de un experimento aleatorio
	- Suceso	- Subconjunto de un espacio muestral, es decir, un conjunto de posibles resultados que se pueden dar en un experimento aleatorio
	- Suceso simple	- Subconjunto unitario del espacio muestral
	- Suceso compuesto	- Sucesos que engloban más de un elemento del espacio muestral
	- Casos favorables	- Número de sucesos que cumplen cierta condición
	- Casos posibles	- Total de elementos del espacio muestral
	- Fracción	- Expresión de una cantidad dividida entre otra
	- Razón	- Relación entre dos números semejantes
	- Proporción	- Igualdad entre dos razones
Propiedades	- La probabilidad está comprendida entre 0 y 1	- El cociente de un número menor positivo entre otro mayor positivo es un número entre 0 y 1
	- Probabilidad del suceso contrario	- Probabilidad del espacio muestral menos la del suceso
	- Restricción del espacio muestral en la probabilidad condicional	- Para el cálculo de la probabilidad condicional el espacio muestral queda reducido a los sucesos que cumplen la condición
	- Proporcionalidad	- Igualdad de razones
Procedimientos	- Operaciones aritméticas	- Suma, recta, multiplicación y división
	- Cálculo con fracciones	- División entre casos posibles y favorables
	- Interpretar una tabla	- Sacar conclusiones de los datos
Argumentos	- Argumentos deductivos	- Demostración de la solución

Un resumen de los objetos matemáticos identificados en dicho análisis semiótico se presenta en la Tabla 3.6.1. Los estudiantes habían recibido enseñanza, tanto sobre los tipos de objetos, como de los procesos matemáticos implicados, aunque en la pregunta sólo se les pidió describir y clasificar los objetos. Más concretamente, dentro del primer tema de la asignatura "*Currículo de matemáticas en educación primaria*", en la que se pasó el cuestionario como parte de la evaluación final, se había incluido los temas "*Tipos de objetos que intervienen en la actividad matemática*" y "*Procesos matemáticos*", como parte de la unidad primera: "*Perspectiva educativa de las matemáticas*".

Además de haberse dedicado una sesión teórica al tema, y haber proporcionado al estudiante, tanto material escrito como transparencias resumen del tema, se realizaron algunas prácticas similares a este ejercicio, en las que el estudiante primeramente tenía que resolver uno o varios problemas matemáticos y posteriormente realizar un listado y clasificación de los objetos matemáticos usados (además de otras actividades, como describir las estrategias usadas en la resolución de los problemas). La diferencia entre las actividades realizadas en clase respecto a la actual fue el contenido matemático (se resolvieron problemas aritméticos y algebraicos). No obstante, también se realizó una práctica sobre probabilidad, en la que los estudiantes primero resolvieron, ellos mismos, una serie de preguntas de probabilidad tomadas de test de Green (1983). Posteriormente se realizó un análisis didáctico de algunas de estas preguntas, evaluando respuestas al ítem de alumnos ficticios. Para las respuestas correctas se les pidió identificar el contenido matemático del ítem; para las incorrectas los posibles conflictos que causaban los errores de los estudiantes y sugerir actividades para ayudar a superarlos.

Mediante esta pregunta, siguiendo el modelo de Godino (2009), se evalúan algunos conocimientos especializados del contenido. Según Hill, Ball, y Schilling (2008) este es un conocimiento que va más allá del común e implica que los profesores sean capaces de enseñar dicho contenido. Sería un conocimiento específico del profesor que algunos sujetos adultos no tienen. Incluye, según los autores, la habilidad para analizar errores y evaluar ideas alternativas, dar explicaciones matemáticas y usar representaciones matemáticas, así como ser explícito acerca del lenguaje y la práctica matemática.

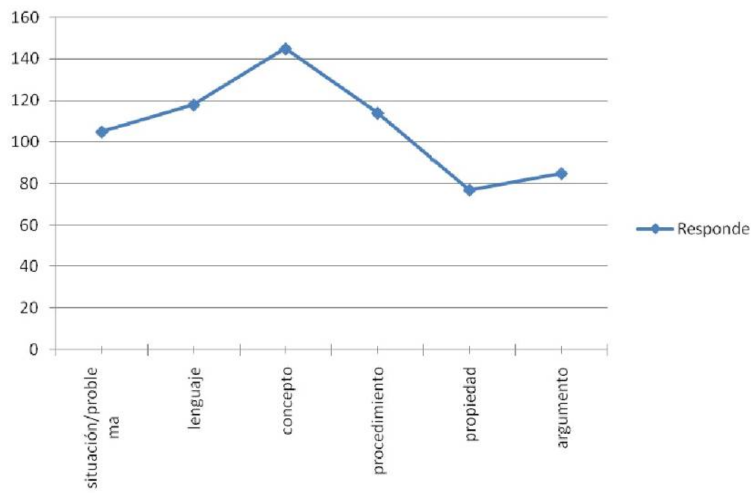
La preguntas propuesta es considerada por Godino (2009) pertinente para evaluar partes de este tipo de conocimiento, aunque, por supuesto, hay otras tareas posibles. El autor indica que responder a la misma supone pensar de manera sistemática en los diferentes procedimientos posibles de resolución, modalidades de expresión, conceptos y propiedades que se ponen en juego con distintos grados de formalidad en su formulación, así como pensar en las maneras de argumentar o justificar los procedimientos y propiedades empleadas.

Recogidos los protocolos se realizó un análisis cualitativo de las respuestas, anotando, cada tipo de objeto matemático (situaciones, lenguaje, conceptos, propiedades, argumentos, procedimientos), clasificando las respuestas en correctas e incorrectas dependiendo de si el objeto citado aparece o no implícitamente en los problemas planteados. Se codificaron el número de objetos correcta e incorrectamente identificados en la tarea por cada estudiante. Seguidamente se procedió a realizar un análisis estadístico descriptivo de estas variables. En la Tabla 3.6.2 se presenta en primer lugar el porcentaje de estudiantes que al menos cita un objeto en cada categoría.

Tabla 3.6.2. Porcentaje de respuestas y no respuesta, Identificación de objetos matemáticos

	Respuesta a las cuestiones de la tarea	
	Responde (%)	No responde (%)
Situación/problema	105 (57,4)	78 (42,6)
Lenguaje	118 (64,5)	65 (35,5)
Concepto	145 (79,2)	38 (20,8)
Procedimiento	114 (62,3)	69 (37,7)
Propiedad	77 (42,1)	106 (57,9)
Argumento	85 (46,5)	98 (53,5)

Figura 3.6.1 Respuesta a la identificación de objetos matemáticos



Observamos un alto porcentaje de respuestas en blanco, a pesar de que podemos suponer que el alumno hizo la tarea con interés, al ser parte de su evaluación final. Estos resultados indican que los futuros profesores tienen dificultad en analizar el contenido matemático de una tarea y no han comprendido la diferencia entre los diferentes tipos de objetos matemáticos. El mayor porcentaje de respuestas está relacionado con los conceptos, objeto que parece más fácil de diferenciar a los estudiantes, seguido del lenguaje, procedimientos, problemas, argumentos y propiedades. Estos datos se visualizan mejor en la Figura 3.6.1.

A continuación estudiamos los objetos correcta e incorrectamente reconocidos por los futuros profesores y haremos una clasificación de los errores cometidos.

3.6.2. SITUACIONES O PROBLEMAS

Los alumnos habían estudiado que la perspectiva histórica muestra claramente que las matemáticas son un conjunto de conocimientos en evolución continua y que en dicha evolución desempeña a menudo un papel de primer orden la necesidad de resolver determinados problemas prácticos (o internos a las propias matemáticas). Se organizaron con ellos discusiones y actividades prácticas sobre el papel de la resolución de problemas en la matemática y en el aprendizaje de esta materia. También se les había indicado que, a nivel escolar, se consideraba como "problema-situación", tanto los problemas, como los ejercicios propuestos a los estudiantes.

Tabla 3.6.3. Respuestas al objeto matemático situación/problema

	Responde (%)	No responde (%)
Total	105 (57,4)	78 (43,6)
Correcto	89 (84,8)	
Incorrecto	29 (15,2)	

En la Tabla 3.6.3 se clasifican las respuestas según se identifica correcta o incorrectamente la situación o problema. El objeto matemático "*situación o problema*" presenta dificultades a la hora de ser interpretado por los futuros profesores de primaria ya que apenas un 57,4% de ellos respondió la pregunta, por lo que hemos de suponer que el resto ni siquiera recordaba a qué nos referíamos con el término "problema". De los que respondieron, el 84.8% dio al menos un ejemplo correcto, es decir fue capaz de identificar al menos un problema matemático en la tarea propuesta; pero otro 15,2% proporcionó respuestas erróneas, es decir, identificó objetos que no correspondían a problemas matemáticos en la tarea. A continuación realizamos un análisis más detallado de las soluciones aportadas por los futuros profesores, clasificándolas en correctas o incorrectas, poniendo un ejemplo de cada una y analizando las razones por las cuáles consideramos la respuesta correcta o no.

Respuestas correctas

La mayoría de los participantes que proporcionaron respuesta en este apartado identificaron correctamente que en el enunciado se proponía un problema de cálculo de probabilidades, aunque en la mayoría de los casos no se indica si se trata de probabilidad simple, compuesta o condicional. Este sería un ejemplo de respuesta: "*Los problemas que se trabajan en este ejercicio son los de probabilidad*". Es decir, el nivel de análisis se mantiene muy incipiente.

En algunos casos, se aclara que el enunciado contiene tres problemas diferenciados de probabilidades, aunque no se indica el tipo de probabilidad que se pide calcular, como el siguiente: *"Se pretendían tres problemas: ¿Cuál es la probabilidad de que le guste el tenis? ¿De que sea chica y además le guste el tenis? y la probabilidad de que siendo chica le guste el tenis"*. Observamos que el estudiante se limita a reproducir el enunciado del problema. Pero, al menos, ha sido capaz de diferenciar que se plantean tres problemas.

Otros alumnos especifican cuáles son los datos y las preguntas del problema, por ejemplo: "*El problema presenta a unos alumnos (700), divididos entre chicos (450) y chicas (250) y dentro de estas divisiones se subdivide en si les gusta el tenis. Se pregunta por la probabilidad de que a un alumno le guste el tenis, que sea chica y le guste el tenis y que sabiendo que es chica le guste el tenis*". Es decir, además de aludir al problema, se identifican los datos dados y la pregunta propuesta en el mismo. El nivel de análisis de los problemas es similar en todas las respuestas, manteniéndose éste nivel muy incipiente, sin profundizar en el tipo de problema y mostrando un conocimiento especializado del contenido escaso para el campo de la probabilidad.

Respuestas incorrectas

Entre las situaciones o problemas identificados hemos considerado incorrectos aquellos en que los alumnos responden vagamente, sin indicar qué contenido matemático se trabajó, por ejemplo: "*A partir de los resultados obtenidos por azar, responder a las siguiente preguntas que aparecen debajo de los datos*". El estudiante identifica que se trata de un problema simplemente porque contiene datos y preguntas. Aunque esto ya indica un cierto conocimiento especializado del contenido (pues muestra que conoce los componentes que ha de tener un problema planteado a los alumnos), este conocimiento especializado no relaciona el problema con el campo de la probabilidad.

Entre las respuestas incorrectas también hemos encontrado alumnos que diferencian tres problemas en el enunciado "*El tema es un problema que encierra tres apartados, cada uno de los cuáles es un problema en sí*". Como en el caso anterior, el estudiante se fija en las características típicas de un problema, pero no profundiza en su análisis.

Algunos estudiantes confunden el enunciado del problema, con el problema en sí mismo, por ejemplo: "*El enunciado propuesto y la tabla de datos*". Es posible que la escasa capacidad de comunicación escrita de alguno de los estudiantes lleve a este tipo de respuestas. Otros futuros profesores confunden el problema con la estrategia de resolución: "*Realizar gráficos*"; además en este caso particular, la resolución de los problemas propuestos no requiere la realización de gráficos, por lo que es posible que este estudiante considere la tabla doble como un gráfico.

Hemos considerado incorrectos algunos estudiantes que indican que "*Es un problema de estadística*", porque posteriormente en el análisis del resto de objetos se proponen contenidos estadísticos y no probabilísticos; por ejemplo, se proponen como conceptos la media o la varianza, que no se requieren para resolver el problema planteado. Finalmente aparecen algunas respuestas no relacionadas con la matemática: "*Un problema que se le suele plantear en su vida*". En esta materia y ante una tarea que repercute en la evaluación final del alumno, algunos estudiantes que no han estudiado el tema prefieren dar cualquier respuesta que le viene a la cabeza antes que dejar la pregunta en blanco, con la expectativa de acertar con la misma.

3.6.3. LENGUAJE

A los alumnos se les había hablado del triple carácter de la matemática: como conjunto organizado de conceptos, como actividad y como lenguaje. Se les había mostrado múltiples ejemplos de lenguaje verbal, numérico, simbólico, oral, escrito, gráfico, tabular e incluso icónico. Además se discutió la doble función del lenguaje: (a) representacional: nos permite designar objetos abstractos que no podemos percibir; y (b) instrumental: como herramienta para hacer el trabajo matemático. Esperábamos, por tanto, que citaran algunos de estos tipos de lenguaje o funciones del lenguaje en forma general, o bien ejemplos de estas categorías en el enunciado del problema.

En la Tabla 3.6.4, vemos que el objeto matemático "*lenguaje*" presenta, al igual que las situaciones, dificultades a la hora de ser interpretado por los futuros profesores de primaria ya que apenas un 64,5% de ellos fue capaz de dar alguna respuesta a la cuestión. Dentro de los que la dieron el 68.7% fue capaz de dar al menos un ejemplo correcto de lenguaje utilizado en la tarea. A continuación analizamos las respuestas correctas e incorrectas en el apartado.

Tabla 3.6.4. Respuestas al objeto matemático lenguaje

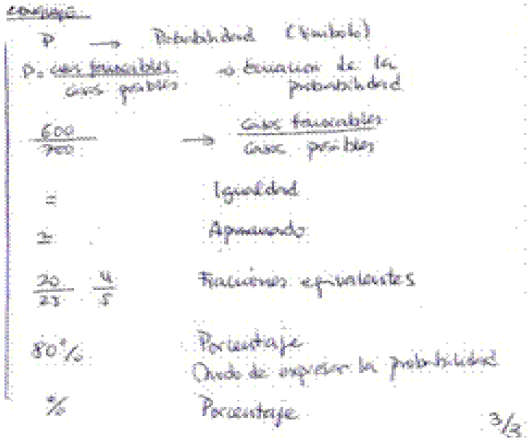
	Responde (%)	No responde (%)
Total	118 (64,5)	65 (35,5)
Correcto	81 (68,7)	
Incorrecto	37 (32,3)	

Respuestas correctas

Respecto a los tipos de lenguajes matemáticos que se encuentran implícitos en la tarea propuesta encontramos que los futuros profesores son capaces de identificar varios tipos de lenguajes, y añadir ejemplos tales como: "*lenguaje verbal textual: palabras de los enunciados*", "*lenguaje matemático: probabilidad*", "*numérico: datos de la tabla*", "*visual: la tabla*" y "*simbólico: en la resolución al operar con el producto y la división*". Observamos que en algunos de los ejemplos la expresión verbal no es muy alta, pero mejor que en el caso de los problemas; sirve de ejemplo el siguiente: "*Se trata del lenguaje matemático escrito en el que se trata la probabilidad*".

La mayoría de los futuros profesores son capaces de reconocer el lenguaje verbal matemático y la tabla de datos, como en las siguientes respuesta: "*El lenguaje que se* *utiliza es verbal matemático, ya que nos da unos datos escritos".* Pero también: *"El problema se representa a través de una tabla*"; "*Tabla en la que se representan los resultados obtenidos y las expresiones matemáticas utilizadas*" o "*Lenguaje matemático: cifras, tablas de frecuencias*". El lenguaje simbólico es también fácil de reconocer para los estudiantes, algunos de ellos son capaces de expresar el tipo de relaciones entre elementos e interpretarlas como en el ejemplo de la Figura 3.6.2.

Figura 3.6.2. Lenguajes correctamente interpretados



También encontramos alumnos que describen las funciones del lenguaje explicadas en la clase, aunque con poca capacidad de comunicación escrita, por ejemplo: "*En el lenguaje se utilizan tanto relacionales como instrumentales, ya que nos imaginamos los objetos abstractos. También instrumentales ya que se utiliza tablas*" o "*Se utiliza un lenguaje simbólico mediante el cual se representan unos datos que nos da información y hace que la entendemos mejor*". En algunos casos se incluyen ejemplos aislados de números como por ejemplo "*400, 200,…*", o términos matemático "*probabilidad, división, porcentaje",* sin indicar que se considera lenguaje porque representa un objeto matemático.

Respuestas incorrectas

Hay alumnos que reconocen el lenguaje pero lo interpretan como si fuese el problema, es decir confunden estos dos objetos matemáticos, como en la respuesta: "*Calcular la probabilidad*", e incluso confunden el lenguaje con los datos "*Le gusta el tenis, no le gusta el tenis, probabilidad de que a uno le guste el tenis, probabilidad de que salga una chica y le guste el tenis y probabilidad de que a una chica le guste el tenis*" o "*Interpretación de las tablas y de los datos representados en ella*" o con elementos del problema tales como "*Chicos /chicas, total, ¿cuál es la probabilidad?, gusto afirmativo o negativo, totalidad, resultados*"; "*El azar*" o "S*e realizan preguntas referentes a estos datos*". Los hemos considerado incorrectos porque, aunque se trata de ejemplos de lenguaje verbal, no hacen referencia a que son ejemplos de este tipo de lenguaje o de que estos términos sirven para representar otros objetos matemáticos.

La mayoría de las respuestas incorrectas se deben a que los alumnos responden con ideas generales, pero no aplican la categoría lenguaje a la situación analizada, por ejemplo: "*El lenguaje ha de ser claro y sencillo*" o "*El lenguaje que se utiliza es un lenguaje claro y conciso, para que el alumno no lo lleve a la confusión*".

3.6.4. CONCEPTOS

Los estudiantes han estudiado durante el curso la diferencia entre comprensión conceptual y procedimental y han analizado los contenidos conceptuales y procedimentales del currículo. Asimismo, se les ha mostrado ejemplos y han realizado actividades en las que debían identificar conceptos, entendiendo que un concepto es un objeto matemático para el que podemos dar una definición.

El objeto matemático "*concepto*" es el que presenta menos dificultades a la hora de ser interpretado por los futuros profesores de primaria ya que el 79,2% de ellos (como podemos ver en la Tabla 3.6.5) fue capaz de dar una respuesta a la cuestión de identificar algún tipo de concepto implícito en el contenido. El aspecto más significante de que este objeto matemático es más fácil de interpretar es que casi el 89% de los que contestaron fueron capaces de al menos contestar correctamente a alguno de los conceptos que aparecen implícitos en la tarea. A continuación analizamos las respuestas correctas e incorrectas.

Tabla 3.6.5. Respuestas al objeto matemático concepto

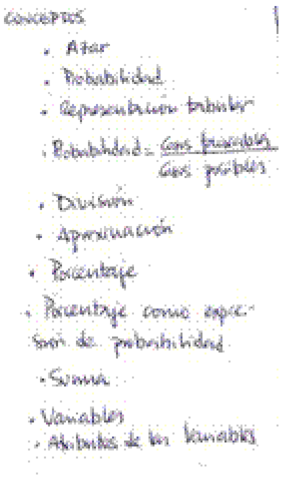
	Responde (%)	No responde (%)
Total	145 (79,2)	38 (20,8)
Correcto	129 (89)	
Incorrecto	16 (11)	

Respuestas correctas

El objeto matemático "*concepto*" es el más reconocido por parte de los estudiantes ya que estos objetos forman parte del vocabulario habitual de los alumnos y éstos son capaces de reconocerlos de forma más precisa que el resto de objetos matemáticos. Los conceptos más reconocidos son: "*azar, probabilidad, variables, porcentaje, división y suma*", los alumnos suelen dar varios ejemplos, como podemos ver en la Figura 3.6.3.

Otros conceptos reconocidos por los alumnos son: "*casos favorables, casos posibles, aproximación", "producto, suceso favorable, suceso no favorable", "números naturales, suma de números enteros, división y multiplicación de números naturales y tanto por ciento" o "Fracciones con igual o distinto denominador*". No se hizo referencia a la probabilidad compuesta o condicional.

Figura 3.6.3 Conceptos correctamente interpretados



Respuestas incorrectas

También encontramos errores en la interpretación de los conceptos, como algunos alumnos que confunden conceptos con procedimientos, por ejemplo: "*Para realizar estos problemas, los conceptos necesarios son saber construir tablas, colocar los datos en su lugar, conocer los logaritmos o estrategias para resolver un caso probable*" o también la respuesta "*Representación tabular*".

Como en los casos anteriores la expresión verbal de los estudiantes es muy pobre, por lo que en algunos casos no es claro si se están refiriendo al concepto o al procedimiento asociado: "*División para calcular la relación entre los casos favorables y posibles que nos ayuda a conocer las posibilidades que tengo entre 0 y 1*"; "*Multiplicación para pasar de conocer la probabilidad entre 0 y 1 a pasar a conocerla* *en porcentaje*"; "*Saber que la cantidad de numerador corresponde a los casos favorables y la denominador a los casos posibles*".

En otros casos se cita como concepto, la pregunta planteada en el problema: "*¿Cuál es la probabilidad de…?*". Algunos estudiantes como este, explicita conceptos como "*dato*" o "*tabla*", aunque no se refiere a ellos directamente: *"Los conceptos trabajados son la representación gráfica de los datos en forma de tabla".*

Hemos considerado también incorrectos los alumnos, que responden correctamente conceptos, pero no están relacionados con el problema tratado: "*Media, mediana, desviación típica, moda*". Otros errores que cometen los alumnos es confundir los conceptos con los elementos descritos en el problema: *"Chico, chica, le gusta el tenis, no le gusta"*, "*El número de chicas a las que si le gusta el tenis y a los que no*", "*Los conceptos son chicos y chicas que les gusta o no dicho deporte*" o "*Chicos, chicas, total*".

3.6.5. PROCEDIMIENTOS

Como se ha indicado, los estudiantes han estudiado la diferencia entre comprensión conceptual y procedimental y han analizado los contenidos conceptuales y procedimentales del currículo. Asimismo, se les ha mostrado ejemplos y han realizado actividades en que deben identificar procedimientos, entendiendo por ellos formas de actuar: operaciones, algoritmos, estrategias, etc.

Tabla 3.6.6. Respuestas al objeto matemático procedimiento

	Responde (%)	No responde (%)
Total	114 (62,3)	69 (27,7)
Correcto	90 (78,9)	
Incorrecto	24 (21,1)	

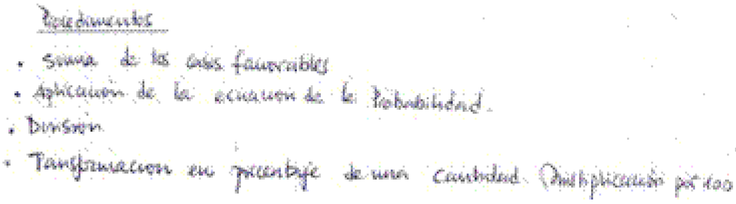
El "*procedimiento*" también presenta dificultades al ser interpretado por los futuros profesores de primaria, ya que apenas un 62,3% de ellos (Tabla 3.6.6) fue capaz de dar una respuesta (correcta o incorrecta) a la cuestión de identificar algún tipo de procedimiento implícito en el contenido de la tarea. De los alumnos que dieron una respuesta, casi un 78,9% fue capaz de responder correctamente a alguno de los procedimientos, de lo que se puede interpretar que los alumnos que contestaban eran aquellos que realmente estaban convencidos de reconocer el procedimiento en la tarea.

Respuestas correctas

Entre los procedimientos correctamente identificados se hace referencia a procedimientos efectivamente empleados para calcular las soluciones. La mayoría de las respuestas hacen referencia al modo de calcular las probabilidades como por ejemplo: *“Dividir casos posibles entre favorables y pasar de probabilidades a porcentajes”* en este caso utiliza Lagrange para calcular la probabilidad. También se hace referencia al cálculo de las probabilidades mediante tantos por ciento, o regla de tres: *“La recogida de datos para el análisis y la expresión de la probabilidad, ya sea en tanto por ciento o en un número de cada 10”, “Regla de tres”, “Aplicación de la regla de probabilidad”, “Transformación en porcentaje de la probabilidad multiplicando por 100”.*

Algunos estudiantes mencionan los procedimientos generales utilizados en la sesión como interpretar los datos del problema: “*Realización de una tabla o representación gráfica, división, multiplicación, comparación”*, pero la mayoría de las respuestas hacen referencia a procedimientos específicos tales como: “*Calculamos mediante la suma el número total de chicos y chicas para ver el total de la población”, “Suma de los casos favorables”* (ver Figura 3.6.4).

Figura 3.6.4. Procedimientos correctamente interpretados


Procedimientos
- suma de los casos favorables
- Aplicación de la ecuacion de la Probabilidad.
- Division
- Transformacion en porcentaje de una cantidad. Multiplicación por 100

Respuestas incorrectas

En algunos casos se confunde el procedimiento con el problema: “*Calcular el porcentaje de chicas respecto al total de los que les gusta el tenis de los que no”* o directamente responden con la cuestión planteada: “*Calcular la probabilidad de que le guste el tenis, calcular la probabilidad de que sea chica”.*

Pero también encontramos otros errores como confundir las soluciones con el procedimiento: “*Las soluciones obtenidas*”, “*La probabilidad del alumno al azar que le* *guste el tenis es 6/7*” o simplemente intentar definir el procedimiento de forma genérica: “*Pasos realizados en el ejercicio para obtener la solución correcta*”. También aparecen errores de confusión de los procedimientos con argumentos como en el siguiente ejemplo: “*El problema se resuelve por inducción*”.

3.6.6. PROPIEDADES

Las propiedades se presentaron a los futuros profesores como relaciones entre dos o más conceptos. En las prácticas realizadas a lo largo de la asignatura se analizaron propiedades utilizadas en la resolución de problemas. También al preparar su unidad didáctica se les pide especificar las propiedades que piensan enseñar.

A pesar de todo este trabajo previo, el objeto matemático “*propiedad*” es el que presenta mayores dificultades para los futuros profesores de primaria ya que apenas un 42,1% de ellos (Tabla 3.6.7) fue capaz de identificar algún tipo de propiedad implícita en el contenido de la tarea. Aparte del bajo porcentaje de respuestas, otro aspecto que demuestra que los alumnos no son capaces de identificar propiedades en la tarea es que el porcentaje que dio una respuesta correcta, dentro del grupo que respondieron, es de alrededor de un 39%, es decir un 61% de los futuros profesores interpretaron mal las propiedades que describían, lo que se puede interpretar que los alumnos que contestaban no reconocen propiedades en la tarea.

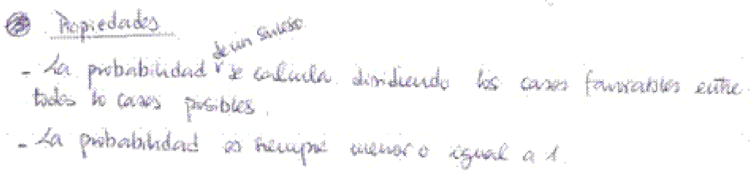
Tabla 3.6.7. Respuestas al objeto matemático propiedad

	Responde (%)	No responde (%)
Total	77 (42,1)	106 (57,9)
Correcto	30 (39)	
Incorrecto	47 (61)	

Respuestas correctas

Algunos de los alumnos identifican correctamente algunas propiedades relacionadas con la probabilidad, aunque la mayoría de ellas son generales y no están muy relacionadas con el tema de estudio, tales como: *“La probabilidad es mayor si a igual denominador el numerador en mayor y al contrario, a igualdad de numeradores es mayor el de menor denominador*”. Otros alumnos describen propiedades correctas desde el punto de vista del proceso: “*La probabilidad de un suceso es siempre menor o igual a 1”* o “*La probabilidad es el número de casos favorables entre posibles*”.

Figura 3.6.5. Propiedades correctamente interpretadas


Propiedades
- La probabilidad de un suceso se calcula dividiendo los casos favorables entre todos lo casos posibles.
- La probabilidad es siempre menor o igual a 1.

Respuestas incorrectas

Entre las propiedades incorrectamente identificadas hemos encontrado alumnos que confunden propiedad con proceso: *Algoritmo de la división; nº total de niñas/nº total gusta el tenis*", "*Suma de algoritmos de diferentes cantidades*", "*Suma, resta, división*", "*Reducción de un número*", "*Calcular el porcentaje, división de números enteros*" o "*Sumar llevándose*".

Incluso se confunde con algunos procedimientos mal descritos como por ejemplo: "*Observar la tabla, interpretar la tabla, centrarnos en lo que nos pide, indicar solución*". Otros alumnos confunden propiedades con conceptos: "*Muestra estadística*", "*Probabilidad, variables*", "*Fracciones, probabilidad*". Pero los resultados más frecuentes son aquellos en los que los alumnos confunden el problema con las propiedades: "*El resultado de una operación*", "*0,86; 0,29 y 0,33: Probabilidad que tengo de que me salga alguno de los sucesos*", "*Calcular el resultado*", "*Saber identificar claramente los contenidos de la tabla para así poder realizar correctamente el problema*" o "*Probabilidad de que le guste el tenis, de que sea chica y le guste el tenis y de que siendo chica le guste el tenis*".

3.6.7. ARGUMENTOS

En la enseñanza recibida se estudiaron los tipos generales de razonamiento matemático: deductivo e inductivo, se habló de la importancia de la justificación y se analizaron ejemplos de justificaciones, además de tener que identificar argumentos en las prácticas ya citadas. No obstante, el objeto matemático "*argumento*" presenta muchas dificultades, junto a las propiedades, para los futuros profesores de primaria ya que apenas un 46,4% de ellos (Tabla 3.6.8) fue capaz de identificar algún tipo de argumento implícito en el contenido de la tarea. Al contrario que ocurría con las propiedades una gran parte (61,2%,) de los que contestan lo hacen correctamente,.

Tabla 3.6.8. Respuestas al objeto matemático argumento

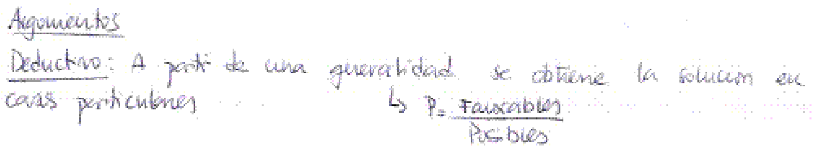
	Responde (%)	No responde (%)
Total	85 (46,4)	98 (53,6)
Correcto	52 (61,2)	
Incorrecto	33 (38,8)	

Respuestas correctas

Respecto a los argumentos correctamente identificados el citado con mayor frecuencia fue el argumento “*Deductivo*” o como indica este ejemplo: “*Deductivo, porque una vez que halamos resulto la suma de cantidades y hallado la probabilidad, tenemos que dar una explicación de por qué hemos sacado este resultado*”.

En una forma más imprecisa se indicaron como argumentos: “*Demostración de cómo realizar y justificar el resultado*” “*Razonar matemáticamente”* o “*A la hora de responder no sólo hay que dar la cifra sino también como ha llegado al resultado, aclarando como se sabe los conceptos*”.

Figura 3.6.6. Argumentos correctamente interpretados



Respuestas incorrectas

En relación a los argumentos incorrectos se mencionan algunos problemas, como en las siguientes respuestas: “*¿Cuál es la probabilidad?”* o “*Explicar porque la probabilidad es esa*”.

Otros confunden procedimientos con argumentos, por ejemplo: “*Como tienen que darse dos sucesos, las probabilidades se acumulan, la probabilidades de ser niña es la proporción de casos favorables (niñas) entre los casos totales (niños y niñas)”*, “*Dos fracciones con distinto denominador es mayor el de menor denominador, dos fracciones con igual denominador es mayor el de numerador mayor*”. Otros alumnos confunden argumentos con los resultados: “*Hay un 86% de posibilidades de que le guste el tenis, un 29% de que le guste el tenis y sea chica y un 33% de que eligiendo chica le guste el tenis*” o directamente intenta explicar el problema “*Te permite saber si hay una probabilidad de que ocurra*”.

3.6.8. SÍNTESIS DEL CONOCIMIENTO ESPECIALIZADO DEL CONTENIDO

En la Tabla 3.6.9 presentamos el número mínimo y máximo, media y desviación típica de objetos matemáticos correctamente identificados en cada categoría. Respecto a los correctamente identificados, los más frecuentes han sido los conceptos, con un promedio de al menos dos conceptos correctamente identificado por cada alumno, seguido de los procedimientos.

Tabla 3.6.9 Objetos matemáticos correctamente identificados en la tarea

Tipos	Mínimo	Máximo	Media	Desviación típica
Situaciones/ problemas	0	3	0,54	0,697
Lenguajes	0	8	1,37	1,911
Conceptos	0	9	2,22	2,012
Procedimientos	0	8	1,40	1,706
Propiedades	0	4	0,32	0,771
Argumentos	0	3	0,37	0,622

Pensamos que este resultado es debido a que en otras materias se hace mucho énfasis en lo conceptual y procedimental. También el lenguaje ha resultado sencillo, en comparación con las propiedades y argumentos. No obstante los resultados en general son pobres. Pensamos que una razón es que se dedica muy poco tiempo al estudio de la probabilidad en el curso de “Matemáticas y su didáctica” y en consecuencia los estudiantes tienen un escaso dominio del conocimiento común del contenido (como ya se vio en la primera parte del capítulo). Ello incide en que su conocimiento especializado del contenido sea asimismo muy escaso. La variabilidad entre estudiantes es muy alta; podemos observar como hay estudiantes que son capaces de identificar hasta 9 conceptos correctos y hasta 8 ejemplos de lenguaje o procedimientos correctos.

Tabla 3.6.10 Objetos matemáticos correctamente identificados en la tarea

Tipos	Responden (%)	No responden (%)	Media (De los que responden)
Situaciones/ problemas	89 (48,6)	94 (51,4)	1,28
Lenguajes	81 (44,3)	102 (55,7)	3,10
Conceptos	129 (70,5)	54 (29,5)	3,15
Procedimientos	90 (49,2)	93 (50,8)	2,85
Propiedades	30 (16,4)	153 (83,6)	1,95
Argumentos	52 (28,4)	131 (71,6)	1,30

Si se tiene en cuenta sólo los estudiantes que responden a cada tarea (Tabla 3.6.10), suben algo los valores medios. Sigue siendo el apartado conceptos donde se obtiene un mayor porcentaje de respuestas, indicando una mayor seguridad en este contenido. Consecuentemente la media sube, dando 3 conceptos correctos en promedio, en caso de responder. Donde mayor subida se observa es en el lenguaje, donde menos de la mitad responde. Es decir, la mitad de alumnos no identifica ningún tipo de lenguaje correcto en este apartado o bien la otra mitad menciona tres elementos diferenciados. En el resto de objetos la media de respuestas correctas sube pero en menor medida, por ejemplo en los procedimientos aumentan de 1,4 de media a 2,85. Los problemas duplican su media al igual que los argumentos, si se toman sólo las respuestas.

Respecto a los objetos incorrectamente identificados, el número de errores en los objetos identificados en cada categoría es menor que el de respuestas correctas, aunque si se tiene en cuenta sólo las respuestas, todas las medias suben. El objeto con mayor media de respuesta en este caso es el lenguaje con 2,61 de media seguido de los conceptos con 2,53. Este con mayor tasa de respuesta que los lenguajes ya que el objeto concepto es el que más veces erróneamente se responde, con un 70,5% de los alumnos dando algún ejemplo erróneo. Podemos ver los resultados completos en las Tablas 3.6.11 y 3.6.12.

Tabla 3.6.11. Objetos matemáticos incorrectamente identificados en la tarea

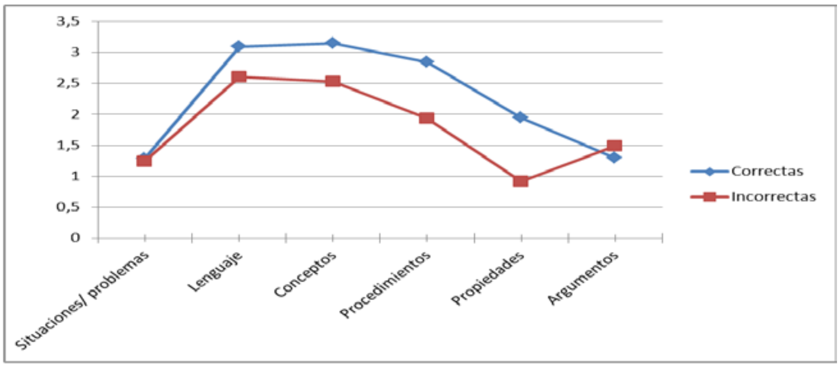
Tipos	Mínimo	Máximo	Media	Desviación típica
Situaciones/ problemas	0	3	0,17	0,469
Lenguajes	0	7	0,57	1,232
Conceptos	0	8	0,87	1,577
Procedimientos	0	5	0,38	0,889
Propiedades	0	3	0,26	0,623
Argumentos	0	3	0,26	0,623

Tabla 3.6.12. Objetos matemáticos incorrectamente identificados en la tarea (Sólo las respuestas)

Tipos	Responden (%)	No responden (%)	Media (de los que responden)
Situaciones/ problemas	25 (13,6)	158 (86,4)	1,24
Lenguajes	40 (21,9)	143 (78,1)	2,61
Conceptos	63 (34,4)	120 (65,6)	2,53
Procedimientos	36 (19,7)	147 (80,3)	1,93
Propiedades	52 (28,4)	131 (71,6)	0,92
Argumentos	32 (17,5)	151 (82,5)	1,49

En la Figura 3.6.7 podemos observar el comportamiento de las respuestas correctas e incorrectas para cada tipo de objeto matemático en referencia al número de respuestas media. El comportamiento es similar para los objetos matemáticos lenguaje, conceptos, procedimientos y propiedades ya que el número de respuestas (con una diferencia de escala de alrededor de 0,5) suele seguir el mismo patrón. En cambio esto no ocurre en el caso de los problemas y argumentos ya que en el primero de ellos el resultado es el mismo tanto para respuestas correctas como de incorrectas y en el segundo el número medio de respuestas incorrectas en mayor que el de correctas.

Figura 3.6.7. Número medio de objetos matemáticos correcta e incorrectamente identificados



Para comparar el número de objetos matemáticos correctamente e incorrectamente identificados en el total de la muestra se realizó pruebas *t* de Student de diferencia de medias para muestras relacionadas (comparación de Tablas 3.6.10 y 3.6.12). Como vemos en la Tabla 3.6.13, obtenemos diferencias estadísticamente significativas en los objetos matemáticos situaciones, lenguaje, conceptos y procedimientos. En todos estos casos el valor medio del número de objetos correctamente identificados es superior al incorrecto.

Tabla 3.6.13. Significación de diferencias entre objetos bien y mal identificados (n=183)

Contraste de medias para todos los alumnos	*p*-valor
Situaciones bien – situaciones mal	0,004
Lenguajes bien – lenguajes mal	0,007
Conceptos bien – conceptos mal	0,002
Procedimientos bien – procedimientos mal	0,003
Propiedades bien – propiedades mal	0,076
Argumentos bien – Argumentos mal	0,066

No se comportan de igual manera las propiedades y los argumentos en el que el p-valor es mayor que el nivel de significación establecido (alfa= 0,05) por lo que no existen diferencias significativas entre ambos conjuntos.

Hemos realizado otro estudio de comparación de medias para muestras independientes teniendo en cuenta sólo las respuestas a los objetos matemáticos. Para deducir el tipo de contraste *t* que se debe usar, se comienza utilizando el test de Levene, test de hipótesis utilizado para comprobar la igualdad de varianzas, cuyos resultados los podemos ver en la Tabla 3.6.14. En ella observamos que a un nivel de significación del 5% podemos aceptar que no hay diferencias significativas entre las varianzas de las muestras de las situaciones/problemas, de los conceptos y de las propiedades y en menor medida de los lenguajes ya que el *p-valor* es mayor que el nivel de significación. No ocurre lo mismo para los procedimientos y para los argumentos en los que podemos afirmar que existen diferencias significativas entre las muestras.

Tabla 3.6.14. Contrastes de varianzas objetos bien y mal identificados (Alumnos que responden)

Contraste de varianzas para respuestas	*p*-valor
Situaciones bien – situaciones mal	0,895
Lenguajes bien – lenguajes mal	0,116
Conceptos bien – conceptos mal	0,968
Procedimientos bien – procedimientos mal	0,032
Propiedades bien – propiedades mal	0,887
Argumentos bien – Argumentos mal	0,004

Posteriormente se ha realizado un estudio de comparación de medias de las muestras asumiendo igualdad de varianzas en el los casos anteriormente aceptados y en el resto asumiendo diferencias entre ellas. Los resultados, que podemos ver en la Tabla 3.6.15, muestran que existen diferencias significativas entre las dos muestras en los procedimientos, no ocurriendo así en los otros objetos en los que vemos que el *p*-valor es superior al nivel de significación por lo que aceptaríamos la hipótesis de igualdad de medias.

Tabla 3.6.15. Contrastes de medias objetos bien y mal identificados (Alumnos que responden)

Contraste de medias para respuestas	*p*-valor
Situaciones bien – situaciones mal	0,967
Lenguajes bien – lenguajes mal	0,356
Conceptos bien – conceptos mal	0,602
Procedimientos bien – procedimientos mal	0,002
Propiedades bien – propiedades mal	0,552
Argumentos bien – Argumentos mal	0,222

3.7. CONOCIMIENTO DEL CONTENIDO Y LA ENSEÑANZA

3.7.1. INTRODUCCIÓN

La importancia que se da a la resolución de problemas en los currículos actuales es el resultado de un punto de vista sobre las matemáticas que considera la resolución de problemas como la esencia de la actividad matemática. Como indica Puig (2008), varios factores nacionales e internacionales coincidieron para hacer que la resolución de problemas se colocara como uno de los centros fundamentales de la reforma del currículo de matemáticas.

Las nuevas orientaciones curriculares (MEC 2006a y b) conceden un papel predominante a la resolución de problemas en el currículo escolar, especialmente la Junta de Andalucía (Consejería de Educación 2007), que incluye la resolución de problemas como bloque transversal a los diferentes bloques de contenido a lo largo de toda la Educación Primaria.

Godino, Batanero y Font (2007) indican que la resolución de problemas ha sido una de las áreas de investigación de mayor impacto en la didáctica de las matemáticas. Los investigadores interesados en entender la interacción entre el estudiante y la tarea de resolución de problemas han analizado las tareas presentadas, las características de los estudiantes, de la situación de evaluación, la enseñanza recibida y otros puntos, tratando de ver cuáles de ellos influyen tanto en el éxito del alumno al resolver el problema como en su aprendizaje. Será entonces importante que los estudiantes puedan analizar las tareas matemáticas y determinar cambios en las mismas para aumentar o disminuir su dificultad o adaptarla a sus alumnos, según su diversidad.

Este conocimiento se podría incluir en el “conocimiento del contenido y la enseñanza”. Con objeto de realizar una evaluación, aunque sea somera de dicho conocimiento, en la tercera parte de la tarea se pide a los futuros profesores indicar algunas variables que se pueda cambiar en esta tarea para variar la dificultad o para variar el contenido matemático. A continuación se analizan las respuestas obtenidas en esta parte.

3.7.2. IDENTIFICACIÓN DE VARIABLES DE TAREA

Para Lester (1983), hay una multitud de variables que inciden en la resolución de problemas, especialmente en matemáticas. Algunas categorías de variables que se usan en las investigaciones en resolución de problemas matemáticos son las siguientes: variables de tarea, relacionados con la naturaleza del problema; variables del sujeto, o características de la persona que resuelve el problema; variables de la situación durante la resolución de problema. A continuación las analizamos en el problema planteado.

1. *Variables del problema:* En un mismo problema o tarea, ligeros cambios en el enunciado, pueden variar su dificultad, las estrategias con que los alumnos tratan de resolverlo o bien los contenidos matemáticos de la tarea.

En el problema planteado a los estudiantes podemos identificar fácilmente algunas variables de tarea que se refieren, al contenido matemático de las mismas, el contexto o el formato de presentación. Respecto al contenido matemático, se podría variar el número de celdas en filas o columnas de la tabla, con lo cual se pasaría de una variable dicotómica a otra politómica; también se podría cambiar la escala de medida de una o las dos variables, para usar variables ordinales o numéricas. Si se suprime los totales de las filas y/o columnas, los alumnos tienen que calcular frecuencias marginales, que ahora vienen dadas. Si se aumenta o disminuye el tamaño de las frecuencias en las celdas o total de la muestra, se aumenta la dificultad. También cambiaría el contenido matemático si en vez de dar los datos en frecuencias absolutas se diesen en frecuencias relativas dobles o frecuencias relativas condicionales.

Se podría plantear preguntas más complejas; por ejemplo, calcular la probabilidad de coger dos o más sujetos con o sin reemplazamiento en cada caso. Todos estos cambios afectarían a la dificultad del problema. Se les podría preguntar por la proporción (en vez de por la probabilidad) o bien hacer un gráfico a partir de la tabla.

Respecto al formato con que se presentan los datos; se podría presentar el enunciado verbalmente, o por medio de un listado o gráfico. La tabla de doble entrada se da elaborada, pero si los datos se diesen a partir de un listado, los alumnos tendrían que clasificarlos y formar la tabla, aumentando la dificultad.

2. *Variables del sujeto:* Los alumnos tienen diversas capacidades, intereses, actitudes e historia. Las circunstancias sociales y familiares también pueden influir, por ejemplo, el apoyo de sus padres en el estudio o los medios que éstos le proporcionan.

En este ejemplo, los estudiantes podrían mencionar que el problema se ha propuesto a futuros maestros, pero la dificultad podría cambiar si se propone a niños (sería más difícil) o si se propone a estudiantes de matemáticas u otras especialidades que hayan seguido un curso de probabilidad (caso en que el problema sería más sencillo).

3. *Variables de la situación:* De resolución, serían los cambios en la situación en que se resuelve el problema, que pueda afectar a las estrategias, soluciones o dificultad el mismo. Por ejemplo, las herramientas disponibles mientras se resuelve, si se trabaja sólo o en grupo, etc.

En este ejemplo, cada estudiante ha de resolver el problema por sí mismo y sin ayuda de libros de consulta o calculadora. La dificultad sería menor si el problema se hubiera podido resolver en grupos o si el alumno hubiese podido consultar fuentes o usar ordenadores o calculadoras. Asimismo, si hubiese tenido más tiempo para resolverlo podría mejorarse el porcentaje de soluciones correctas.

Los diferentes autores, que han estudiado la formación de profesores respecto a resolución de problemas, indican que los profesores debieran saber identificar aquellas variables cuyo control se puede considerar como un recurso del profesor, es decir sobre las que podemos actuar y que producen un cambio significativo en lo que el alumno aprende; son las llamadas variables didácticas. Generalmente son variables de tarea o de la situación; pero también a veces se puede actuar sobre las variables del sujeto, por ejemplo, tratando de aumentar el interés o mejorar la actitud de los alumnos. En lo que sigue analizamos las respuestas de los estudiantes, estudiando el número de variables correcta e incorrectamente identificadas en el problema, comenzando por las variables de tarea.

Variables de tarea correctas

En primer lugar presentaremos algunas respuestas correctas sobre las variables de tarea que podrían cambiarse en el problema. Podemos diferenciarlas en varios grupos ya que en algunos casos las variables se refieren al contenido matemático, otras al formato de presentación y otras al formato en que se dan los datos.

Respecto al formato de la tarea, encontramos futuros profesores que cambiarían los datos del ejercicio por otros tomados dentro de la clase: *"Cambiar la obtención de datos, tomándolas dentro del curso"*. En otro ejemplo: *"se podría hacer que los alumnos tuvieran que realizar la tabla partiendo de las respuestas de los alumnos de la clase"* el sujeto indica posibilidad de hacer una encuesta y de que los alumnos preparen las tablas por ellos mismos. Por lo tanto, además del formato, esta variable afecta al contenido del problema, pues se añaden dos procedimientos (realizar una encuesta y confeccionar la tabla).

Otros participantes sugieren cambiar el formato de presentación de los datos y utilizar gráficos estadísticos para representarlos, por ejemplo, indicando incluso gráficos específicos: *"Elaborar gráficos de barras o diagramas de sectores para ver las habilidades de representación gráfica de los alumnos", "Hacer un gráfico donde comparen a los chicos con las chicas, por ejemplo un diagrama de sectores donde se puede observar muy bien la proporción"*. También en estos casos, además del formato, la variable afecta al contenido, pues se añade el de los gráficos citados y el implicado en su realización (proporcionalidad, amplitud angular, etc.).

Respecto a las variables del contenido del problema encontramos alumnos que se centrarían en variar la pregunta, de modo que se pidan las probabilidades en las que aparezca sucesos complementarios, como por ejemplo: *"Calcular la probabilidad de que sea chico y no le guste el tenis", "Darle los totales y los chicos que les gusta el tenis y pedirle que calculen los que no les gusta"*. En el siguiente ejemplo el alumno sugiere varias variables, las primeras para aumentar la dificultad del problema sin variar el contenido matemático y las últimas para variar dicho contenido.

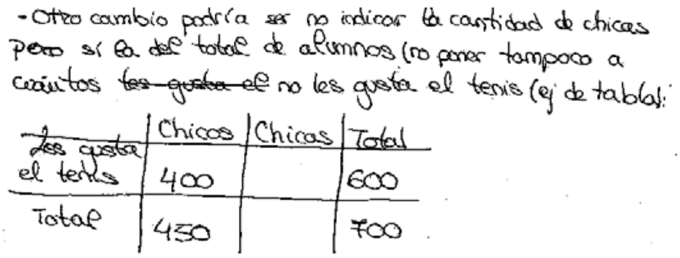
> *"Variables a cambiar, por ejemplo podría ser la obtención de datos de algún otro curso, dificultando así y extendiendo el ejercicio.*
>
> *O bien cambiando el gusto por otro deporte, albergando preguntas como cuantos chicos – chicas les gusta ambos deportes…*
>
> *Dichas variables mantienen el contenido matemático, para cambiarlo, podríamos pedir que se elabore gráficas de barras o un diagrama de sectores y ver si así las habilidades de representación gráfica de nuestros alumnos."*

Otros futuros profesores proponen variar la pregunta del enunciado, por ejemplo, para pedirles a los alumnos que mostrasen las diferentes estrategias utilizadas para su resolución: *"Pedirles diferentes estrategias para resolverlo", "Que el alumno justifique como ha encontrado las soluciones y que nos diga porque cree que la suya es la correcta"*. En otros casos, se sugiere que, una vez resuelto el problema, que añadan cuestiones para ver si han llegado a una interpretación correcta y saben aplicar los resultados: *"Una vez calculado todo pedirles que reflexionen si ellos montarían una pista de tenis en ese colegio"*.

Una de las variables más citadas por los alumnos fue variar el tipo de datos proporcionados. Por ejemplo, como vemos en la Figura 3.7.1, se sugiere eliminar algunos elementos en la tabla de datos, ya fuera excluir los totales o incluir estos y excluir algunas frecuencias dobles con la intención de que el alumno tuviese que calcularlos para poder resolver el problema: *"No poner el total de los alumnos", "Pedirles que comparen las probabilidades", "No indicar la cantidad de chicas pero si el total de alumnos", "Se puede dar la tabla y que falte algún dato, si le pedimos que rellenen los huecos y después que realicen el problema tal como está enunciado, este problema será más complejo y potenciará también la suma".*

También encontramos futuros profesores partidarios de omitir la tabla de datos y que los alumnos tengan que interpretar los datos desde el enunciado *"Expresar de otra forma los datos, en lugar de dárselos en una tabla, podrían aparecer en el enunciado",* o pasar de números enteros a porcentajes, con lo cual cambia el contenido matemático: *"Dar los datos a través de porcentajes, por ejemplo el 25% de 700, que es el total, corresponde a las chicas y que calculen el resto de datos".* Otros alumnos plantean el cálculo de probabilidades compuestas, como *"Calcular la probabilidad de que dos alumnos uno sea chica sabiendo que no le gusta el tenis y el otro sea chica y le guste el tenis".*

Figura 3.7.1. Variable del problema correctamente interpretada



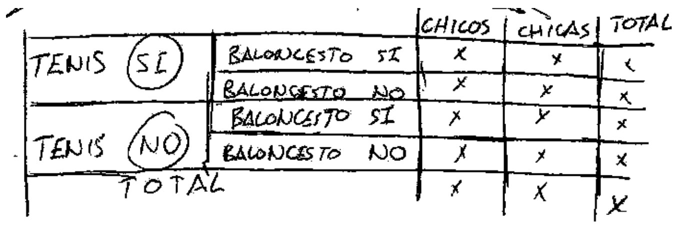
- Otro cambio podría ser no indicar la cantidad de chicas pero sí la del total de alumnos (no poner tampoco a cuántos ~~les gusta el~~ no les gusta el tenis (ej de tabla):

	Chicos	Chicas	Total
~~Les gusta~~ el tenis	400		600
Total	430		700

Otros posibles cambios se refieren a completar la resolución o presentar los resultados de una manera definida *"No dejar indicado el problema sino que lo calculen", "Para variar el contenido matemático se podría preguntar, además de la probabilidad, expresar los resultados en porcentajes", "Pedirles porcentajes, que sumen la probabilidad de dos casos distintos, que comparen los resultados de ambos sexos"* o *"Preguntar cuánto más le gusta a los chicos jugar al tenis".* La variante más demandada consistía en añadir otras variables *"Añadir más datos a la tabla, como quienes practican y cuantos no", "Añadir otra variable".*

En la Figura 3.7.2 mostramos la sugerencia de un estudiante que propone usar una tabla triple, que construye correctamente, aunque no incluye datos, pues su sugerencia es que los alumnos hagan una encuesta en clase para completarla. En cuanto a hacer que el enunciado sea más simple algunos alumnos indican variaciones como: *"Tomar números más pequeños", "Para hacer el ejercicio más fácil podríamos cambiar los datos de los números, por ejemplo poniendo un total de 100 alumnos el ejercicio resultaría más fácil", "En lugar de diferenciar entre chicos y chicas podría darse sólo los datos generales, así la tarea tendría menos dificultad.* Otros ejemplos son los siguientes: *"Enfocar la probabilidad hacia un tema que sea de mayor interés para los alumnos", "Cambiar tenis por futbol que les gusta más". "Una posible variable sería enfocar la probabilidad hacia un tema que sea de mayor interés para los alumnos, como por ejemplo niños que bailan la peonza, en vez de niños que les gusta el tenis. Otra variable podría ser tratar con números más pequeños en el caso de que el alumno presente algún tipo de dificultad en la realización del ejercicio".*

Figura 3.7.2. Variable del problema correctamente interpretada



En la Tabla 3.7.1 describimos los resultados sobre el número de respuestas correctas dadas. En ella vemos que un 12,6% de los futuros profesores no fueron capaces de dar una sola respuesta correcta a este apartado y que el 18,6% ni siquiera contestó a la cuestión. La mayoría de los estudiantes, un 60,1% fueron capaces de identificar una o dos variables correctamente que permiten una variación en su dificultad o en las estrategias con que los alumnos tratan de resolverlo. Por lo que el porcentaje de futuros profesores que expresaron 3 o más variables del problema (el máximo estuvo en 5) se reduce al 8,8%. Vemos en el siguiente ejemplo que el alumno identifica correctamente varias variables.

"Podemos cambiar las variables del problema, a continuación, propongo varias opciones:

- *No adjuntar la tabla en el problema, para que la realicen ellos mismos.*
- *Poner la tabla con algunos huecos sin rellenar, para que los puedan calcular ellos.*
- *Que comparen algunas probabilidades.*

- *Cambiar el enunciado poniendo otro tema.*
- *Realizar preguntas de otro tipo, que no sean de probabilidad."*

Si sólo tenemos en cuenta a aquellos alumnos que responden, el 15,4% falla en la identificación de variables de tarea y el 73,8% fue capaz de contestar a una o dos variables correctas. Como ocurría anteriormente, el porcentaje menor es el de alumnos que son capaces de dar más de dos variables 10,7%.

Tabla 3.7.1. Número de variables del problema correctamente definidas por los futuros profesores

n	Frecuencia Total (%)	Porcentaje valido
0	23 (12,5)	15,4
1	61 (33,3)	40,9
2	49 (26,8)	32,9
3	8 (4,4)	5,5
4	6 (3,3)	4,0
5	2 (1,1)	1,3
No responde	34 (18,6)	100,0
Total	183	Total responden 149

Variables de tarea incorrectamente identificadas

Respecto a los errores en la identificación de las variables de tarea, encontramos alumnos que no identifican correctamente la probabilidad conjunta y la condicionada, por lo que confunden los apartados b y c: *"Cambiar las preguntas b y c ya que pueden llevar a confusión, son la misma, el alumno se puede confundir y pensar que se le pide otra cosa"*. También se da el caso de estudiantes que confunden la probabilidad condicional con la probabilidad simple: *"Poner las cuestiones a y b como la c, es decir planteándolas de manera que el alumno tenga que fijarse mejor en la tabla, analizando dicha cuestión"*, es decir no discrimina el contenido del problema.

Algunos alumnos pretenden calcular estadísticos sin sentido para los datos presentados en el problema, como calcular la moda o mediana de algunos sucesos: *"Calcular la moda, media y mediana de chicos que no les gusta el tenis"*, *"Trabajar conceptos como moda, mediana"*, *"Extraer la moda y otros conceptos"*, *"Para variar el contenido podrían preguntar la moda o la mediana de alumnos que prefieren cada deporte"*. Otros alumnos a la hora de introducir nuevas variables tienen problemas de interpretación de los resultados ya que el problema no tendría sentido con los cambios sugeridos: *"Para variar el contenido matemático añadir edades a la tabla"*, *"Destacar un balance de edad de los chicos y chicas para sumar más datos"*, introduciendo incluso variables que ya aparecen: *"Añadir por ejemplo la variable de que no les gusta el tenis, es decir una tercera variable aumentaría el grado de dificultad"*.

Encontramos participantes que a partir de los datos pretenden pedir a los alumnos resultados imposibles de obtener, ya que solicitan datos que no están en la tarea: *"Preguntarles el grado en que aprecian el tenis: les gusta mucho el tenis, les gusta, les gusta poco, no les gusta"*. Otros proponen cambios que no afectan la dificultad: *"Cambiar las cifras de los chicos por las de las chicas"*.

Otras respuestas son muy imprecisas y las hemos considerado incorrectas, pues tan sólo pretenden con el enunciado cambiando el lenguaje, haciéndolo más técnico o llevar al error de los alumnos preguntándole más cuestiones: *"Podría aumentar la dificultad del problema si el lenguaje no es el adecuado a su nivel"; "Para variar la dificultad podríamos hacer más preguntas, creando con ellas un mayor nivel de dificultad haciendo que el alumno se cuestione nuevas preguntas"*.

En la Tabla 3.7.2 se clasifica a los futuros profesores según el número de respuestas incorrectas que hayan dado para la variable problema. El 47% de los futuros profesores no dio ninguna respuesta incorrecta, mayoritariamente porque solían centrar la respuesta del variable problema en cambiar las cantidades de la tabla de enunciados para que fuese más fácil de calcular las probabilidades o en incluir variables nuevas. Un 24,6% de los futuros profesores dio al menos un resultado incorrecto y el 9,8% restante dio más de uno, destacando un alumno que identificó incorrectamente seis variables.

Tabla 3.7.2. Variables problemas incorrectamente definidas por los futuros profesores

n	Frecuencia Total (%)	Porcentaje valido
0	86 (47,0)	57,7
1	45 (24,6)	30,2
2	11 (6,0)	7,4
3	6 (3,3)	4,0
6	1 (0,5)	0,7
No responde	34 (18,6)	100,0
Total	183	Total responden 149

3.7.3. IDENTIFICACIÓN DE VARIABLES DEL SUJETO

Variables del sujeto correctas

Algunos ejemplos correctos de respuestas, que ya hemos comentado en el apartado anterior, también pueden considerarse como variables del sujeto que hacen referencia a tratar de aumentar la motivación del alumnado, ya sea simplemente cambiando la variable del problema como haciendo el conteo con los propios alumno. Otro ejemplo de variable del sujeto es el siguiente: *"Cambiar lo que le gusta al alumnos por otra cosa o aspecto que despierte interés en los alumnos de clase. Por ejemplo si en la clase gusta mucho el futbol puede proponerse el problema para que así los alumnos se sientan motivados"*. Igualmente un alumno indica *"Una variable sería que los alumnos no estén motivados"*. *"Se puede cambiar la variable del sujeto haciendo el ejercicio los niños de la clase, es decir si les gusta el tenis o no y resolver las probabilidades entre todos los alumnos, así se sentirán más motivados"*.

En otros casos se sugiere cambiar el nivel educativo donde se plantea el problema "*Proponer el problema a niños de secundaria", "Que lo resuelvan los alumnos de universidad"*. O bien hacen referencias a los conocimientos previos del estudiante, que también afecta a la dificultad del problema: *"Si el problema se realiza con alumnos que hayan estudiado el tema"*.

Hemos encontrado pocas respuestas respecto a la variable del sujeto, lo cual era de esperar, teniendo en cuenta que no estaba explícitamente especificada en el enunciado del problema, aunque se incluía en el estudio de variables que afectan a la resolución del problema. Sólo un 15,8% de los futuros profesores interpretó que también se podría modificar la variable sujeto, de ellos un 75,9% definió una variable, cuatro futuros profesores pudieron dar dos y sólo tres de ellos erraron en la definición de las variables del sujeto.

Tabla 3.7.3 Variables del sujeto correctamente definidas por los futuros profesores

n	Frecuencia Total (%)	Porcentaje valido
0	3 (1,6)	10,3
1	22 (12,0)	75,9
2	4 (2,2)	13,8
No responde	154 (84,2)	100,0
Total	183	Total responden 29

Variables del sujeto incorrectas

La mayoría de los ejemplos incorrectos de variantes sobre la variable sujeto se refieren a que no se puede influir en la variable sujeto: *"No se puede actuar sobre la variable sujeto porque para ello tendríamos que conocer a los alumnos a los que se les pone el problema"*. En el ejemplo presentado a continuación, además de no ser capaz de pensar en variables del sujeto, indica una variable que se refiere más a la enseñanza y no al problema a resolver.

> *"Podemos cambiar también variables en los sujetos (pero esto no podríamos hacerlo desde aquí pues no tenemos delante a los alumnos, pero un ejemplo de ello sería lo siguiente, cambiaremos la forma de explicarles la probabilidad, centrándonos en unos determinados conceptos)".*

Respecto las variables sujeto incorrectas tenemos que sólo ocho alumnos, un 27,6% de los que contestaron, respondieron incorrectamente a las variables del sujeto. Fueron muy pocos, los que cometieron errores en esta variable.

Tabla 3.7.4. Variables del sujeto incorrectamente definidas por los futuros profesores

n	Frecuencia Total (%)	Porcentaje valido
0	21 (11,5)	72,4
1	8 (4,4)	27,6
No responde	154 (84,1)	100,0
Total	183	Total responden 29

3.7.4. IDENTIFICACIÓN DE VARIABLES DE LA SITUACIÓN

Variables de situación correctas

Algunos estudiantes hicieron referencias a las variables de la situación de resolución. Como ejemplo una respuesta dada por los futuros profesores de primaria es la posibilidad de trabajar en grupo o individualmente para resolver el problema. Otros ejemplos son los siguientes: *"Hacer el trabajo en grupo"*, *"Dejar que los alumnos de los grupos hablen entre sí, dejar más o menos tiempo"*. *"Podríamos también emplear una variable en la resolución, diciéndoles a los alumno que es un ejercicio individual, y no se puede hacer en grupo, ya que en grupo, puede ser uno el que siempre dé las respuestas"*.

Otros estudiantes se centran en el material para resolver la actividad: *"Hacer el ejercicio de manera práctica en vez de sobre papel"*, *"Este tipo de actividad se podría realizar en formato digital, trabajando también las tecnologías"*, *"Que resuelvan el problema con o sin calculadora, en grupos o por parejas"*, *"Solucionar el problema mediante materiales didácticos, por ejemplo si fueran número pequeños se podría repartir canicas para representar a los alumnos o representar los datos mediante dibujos o gráficas de árbol"*. Y por último, encontramos algunos que sugieren la posibilidad de realizar el problema en casa para que los alumnos lo consultasen con sus padres: *"Se podría mandar para casa en lugar de hacerlo en clase para que así le pudieran ayudar sus padres"*.

En la Tabla 3.7.5 incluimos los resultados de las variables de situación. Solamente un 10,9% de los alumnos dio alguna variante de la variable situación correcta. Destaca un alumno que fue capaz de dar cuatro variantes.

Tabla 3.7.5. Variables de la situación correctamente definidas por los futuros profesores

n	Frecuencia Total (%)	Porcentaje valido
1	15 (8,2)	75
2	2 (1,1)	10
3	2 (1,1)	10
4	1 (0,5)	5
No responde	163 (89,1)	100,0
Total	183	Total responden 20

Variables de situación incorrectas

En este caso sólo tenemos una respuesta incorrecta, en el que el alumno confunde el la variable con el problema que trata: *"Que los alumnos tengan que trabajar en pequeños grupos y no se entiendan con los compañeros de clase, o por el contrario, que no quieran resolver la actividad sólo"*. Este dato sólo representa el 1% del total y un 5% de los que responden por lo que se puede tomar como un resultado aislado.

Tabla 3.7.6. Variables de la situación incorrectamente definidas por los futuros profesores

n	Frecuencia Total (%)	Porcentaje responden (frecuencia)
0	19 (9,9)	95,0
1	1 (1,0)	5,0
No responde	163 (89,1)	100,0
Total	183	Total responden 20

3.8. CONCLUSIONES DEL ESTUDIO 1

3.8.1. CONCLUSIONES SOBRE LOS CONOCIMIENTOS MATEMÁTICOS

El Estudio 1 se orientó a realizar un estudio exploratorio de evaluación de algunos conocimientos matemáticos y didácticos de los futuros profesores de Educación Primaria en relación con objetos matemáticos que aparecen implícitos en la interpretación de tablas de contingencia 2x2, es decir con dos filas y dos columnas. El primer objetivo de dicho estudio era el siguiente:

Objetivo 1. Analizar el conocimiento común de probabilidad que ponen de manifiesto los futuros profesores de educación primaria en el cálculo de probabilidades simples, compuestas y condicionales a partir de datos presentados en una tabla de contingencia 2x2.

Para lograr este objetivo se planteó un problema elemental en el que se pidió calcular estas probabilidades a una muestra de profesores en formación. Los resultados se han analizado con detalle a lo largo de la primera parte del capítulo, donde se realizó un análisis semiótico detallado de las soluciones dadas, categorizando las respuestas encontradas, identificando los conflictos semióticos de los estudiantes. Se compararon los resultados con los obtenidos por Estrada y Díaz (2007) en una tarea semejante y una muestra reducida de futuros profesores de educación primaria de la Universidad de Lleida que habían seguido previamente un curso optativo de didáctica de la estadística. Resultados de este estudio se han publicado en Contreras, Estrada, Díaz y Batanero, (2010) y Contreras, Batanero, Díaz y Fernándes (2011).

Por otro lado, formulamos la hipótesis de que *esperábamos que nuestros resultados mostrasen un conocimiento insuficiente de los futuros profesores de la muestra sobre la probabilidad condicional.* Dicha hipótesis se ha confirmado, pues el análisis y resultados de la primera parte del estudio nos indican que la lectura de tablas estadísticas de doble entrada y el cálculo de probabilidades a partir de ellas no es una tarea trivial, para los futuros profesores de Educación Primaria de nuestra muestra.

Las conclusiones sobre este objetivo e hipótesis es que la mayoría de los estudiantes es incapaz de responder correctamente a dos de las tareas propuestas. En el caso de la tarea primera, referente al cálculo de la probabilidad simple, un tercio de los futuros profesores cometen algún tipo de error, pero en las tareas posteriores referentes al cálculo de probabilidades compuestas y condicionales este porcentaje aumenta. Hay que tener en cuenta que el problema se agrava con la gran cantidad de estudiantes que dejaron en blanco las tareas, aproximadamente un tercio de los encuestados, lo que indica una pobreza de conocimientos básicos de probabilidad.

La mayor frecuencia de errores en el cálculo de la probabilidad simple se deben a confusión de probabilidad con frecuencia absoluta, un error que, aunque con poca frecuencia, nos parece importante, dado que implica la no apreciación de los axiomas de probabilidad. También encontramos el error de dar la probabilidad de un elemento al azar de la población, lo que indica que el estudiante es incapaz de leer la tabla de doble entrada. Otros errores importantes son dar la probabilidad mayor que la unidad.

Mayor variedad de errores aparecen en las respuestas, respecto al cálculo de la probabilidad simple, destacando la confusión de probabilidades condicionales y conjuntas ya encontrado por Einhorn y Hogarth (1986) y Ojeda (1995) y Estrada y Díaz (2007). Agrupando los resultados, el 25% de los estudiantes confunde unas probabilidades con otras. Aparece también un error no citado en la literatura consistente en calcular la probabilidad de obtener un sujeto del total de la población o bien con las condiciones dadas (chica a la que gusta el tenis) al azar entre todas ellas. Destacamos también el hecho de que 10 estudiantes confunden probabilidad simple y compuesta un error que no fue encontrado por Díaz y de la Fuente (2007).

La principal confusión en el cálculo de la probabilidad condicional se produce entre probabilidad condicional y compuesta coincidiendo con Einhorn y Hogarth (1986) y Ojeda (1995) y Estrada y Díaz (2007). Encontramos de nuevo algunos casos de alumnos que confunde, la probabilidad pedida con probabilidad del suceso elemental, error que como hemos indicado no se ha descrito anteriormente. Aparecen diversos errores producidos al intentar aplicar las fórmulas directamente sin una comprensión adecuada de las mismas.

Como resumen del análisis, indicamos que se encontró un número medio de 1,7 conflictos por estudiante. Los errores citados por Falk (1986) y Estrada y Díaz (2007) consistentes en confundir una probabilidad condicional con una conjunta o confundir una probabilidad condicional con su inversa, fueron los más frecuentes. Aun así, no explican por sí mismos los altos porcentajes de errores (cercanos al 40%) en las tareas de cálculo de probabilidad condicional y compuesta.

En nuestro análisis, además de los ya citados en estudios previos, aparecieron nuevos conflictos como confundir un suceso con su complementario, errores en la fórmula de cálculo, errores de cálculo, o no ser capaces de identificar los datos en el enunciado. Con menor frecuencia se presentan otros conflictos referidos a la confusión de frecuencia y probabilidad o invertir la fórmula de cálculo de probabilidades, obteniendo valores mayores que uno. Finalmente algunos estudiantes no son capaces de dar una respuesta.

Aunque, como indica Falk (1986), parte de las dificultades pueden ser originadas por la ambigüedad del lenguaje cotidiano utilizado en la interpretación de la probabilidad, el futuro educador debe dominar el lenguaje del azar para transmitirlo en la enseñanza.

Dada la frecuencia de aparición de estas tablas en la prensa, Internet y material profesional, consideramos de gran importancia su correcta interpretación para la toma de decisiones a partir de ellas. Por tanto, es preocupante la variedad de errores y el número de estudiantes que no han podido dar una respuesta en nuestro estudio. Todo ello indica la necesidad de mejorar la educación estadística que los futuros profesores reciben durante su formación en las facultades de educación.

3.8.2. CONCLUSIONES SOBRE EL CONOCIMIENTO DIDÁCTICO

Nuestro segundo objetivo fue a*nalizar algunos conocimientos didácticos de los futuros profesores de educación primaria sobre la tabla de contingencia 2x2 y el cálculo de probabilidades a partir de la misma.*

La segunda y tercera preguntas planteadas estaban dirigidas a evaluar algunos conocimientos didácticos del contenido de los futuros profesores, respecto a dos componentes: (a) conocimiento especializado del contenido, y (b) conocimiento del contenido y la enseñanza. Respecto a esta segunda parte del estudio la hipótesis era, *que nuestros resultados mostrarían un conocimiento insuficiente de los futuros profesores de la muestra en relación a las componentes evaluadas del conocimiento didáctico del contenido.*

Esta hipótesis se ha visto confirmada. En relación al *conocimiento especializado del contenido* se pidió a los estudiantes analizar los objetos matemáticos implícitos en la tarea. Los estudiantes identificaron algunos objetos matemáticos, aunque no todos los necesarios para resolver el problema. Respecto a los correctamente identificados, los más frecuentes han sido los conceptos, con un promedio de al menos dos correctamente identificado por alumno, seguido de los procedimientos.

Pensamos que es debido a que en otras materias se hace mucho énfasis en lo conceptual y procedimental. También el lenguaje ha resultado sencillo, en comparación con las propiedades y argumentos. No obstante los resultados son pobres. Pensamos es que una razón es que se dedica muy poco tiempo al estudio de la probabilidad en el curso de "Matemáticas y su didáctica" y en consecuencia los estudiantes tienen un escaso dominio del conocimiento común del contenido (como ya se vio en la primera parte del capítulo). Ello incide en que su conocimiento especializado del contenido sea asimismo muy escaso. La variabilidad entre estudiantes es muy alta; podemos observar como hay estudiantes que son capaces de identificar hasta 9 conceptos correctos y hasta 8 ejemplos de lenguaje o procedimientos correctos, mientras que hay alumnos que no consiguen dar una respuesta correcta.

Para comparar el número de objetos matemáticos correctamente e incorrectamente identificados en el total de la muestra se realizó pruebas *t* de Student de diferencia de medias para muestras relacionadas, encontrado diferencias estadísticamente significativas en los objetos matemáticos situaciones, lenguaje, conceptos y procedimientos. En todos estos casos el valor medio del número de objetos correctamente identificados es superior al incorrecto. Sin embargo las diferencias desaparecen si se tienen en cuenta sólo los alumnos que responden. No se comportan de igual manera las propiedades y los argumentos en el que el *p*-valor es mayor que el nivel de significación establecido (*alfa*= 0,05) cuando se tienen en cuenta todos los alumnos, por lo que no existen diferencias significativas entre ambos conjuntos.

Para valorar algunos aspectos del *conocimiento del contenido y la enseñanza* se pidió a los estudiantes identificar algunas variables que se pudiesen cambiar en el problema propuesto para aumentar o disminuir la dificultad o bien variar el contenido. Se consideraron algunos tipos de variables tenidas en cuenta en la bibliografía sobre resolución de problemas (Lester, 1983, Goldin y McClintock, 1984; Schoenfeld, 1985, Castro, 1991), diferenciado el número de variables correctas e incorrectamente identificadas respecto a la tarea, al sujeto y a la situación de resolución. Un resumen descriptivo de los resultados muestra la Tabla 3.7.7 donde vemos que, en promedio cada alumno identifica entre una y dos variables correctas en cada categoría.

Tabla 3.7.7. Variables correctamente identificadas en la tarea

Tipos	Mínimo	Máximo	Media	Desviación típica
Del problema	0	5	1,46	1,036
Del sujeto	0	2	1,03	0,499
De la situación	1	4	1,45	0,887

Al identificar variables de tarea se obtuvo 1,46 respuestas, por lo que se puede concluir que la mayoría de los alumnos puede aplicar alguna variable para modificar la dificultad del enunciado del problema. Las otras variables tuvieron una media de respuesta similar, 1,03 en el caso de la variable sujeto y 1,45 en el caso de la de situación, en estos casos la variabilidad ha sido menor que en el de la variable problema. Fueron pocas las variables incorrectamente interpretadas (Tabla 3.7.8), destacando la de la tarea, donde el número de respuesta oscila entre 0 y 6, pero la media de errores está cercana a 0,61 por lo que algunos resultados están influyendo en la dispersión. Para el resto de variables el máximo de errores fue de uno, y ocurrió en muy pocos caso por lo que los resultados no son muy significativos.

Tabla 3.7.8. Variables incorrectamente identificadas en la tarea

Tipos	Mínimo	Máximo	Media	Desviación típica
Del problema	0	6	0,61	0,913
Del sujeto	0	1	0,28	0,455
De la situación	0	1	0,05	0,05

En resumen, la mayoría de la muestra fue capaz de identificar una o dos variables en cada categoría y son más las respuestas correctas que las incorrectas. Los resultados indican que en los futuros profesores son capaces de encontrar variaciones en las cuestiones para con ello modificar los contenidos y la dificultad de las tareas propuestas con el fin de llegar a una comprensión mayor por parte del alumnado, pero con gran dificultad, pues el número de respuestas es pequeño. Todo ello confirma de nuevo la necesidad de reforzar la preparación didáctica de los profesores en el campo de la probabilidad.

CAPÍTULO 4

EVALUACIÓN DE CONOCIMIENTOS DE FUTUROS PROFESORES DE SECUNDARIA

4.1. INTRODUCCIÓN

Como se analizó en el capítulo 2, es muy común encontrar una variedad de errores en el razonamiento condicional, que pueden llevar a decisiones incorrectas con consecuencias indeseables. El profesor que tiene que explicar este tema, en la Educación Secundaria Obligatoria o Bachillerato, debiera ser capaz de reconocer estos sesgos en sus estudiantes y ayudarles a superarlos. De este modo, podrá prepararlos para la correcta toma de decisiones en su vida personal y profesional. Pero es posible que los mismos profesores tengan estos razonamientos erróneos, con lo que podrían incluso transmitirlos a sus alumnos (Stohl, 2005).

El Estudio 2, que se presenta en este capítulo, está orientado a evaluar la presencia de los sesgos más comunes sobre el razonamiento en probabilidad condicional, que se describieron en el capítulo 2, en una muestra de futuros profesores de Educación Secundaria y Bachillerato. Al mismo tiempo se desea evaluar el conocimiento común de dicho contenido, pues dicho conocimiento es esencial en la toma de decisiones instruccionales (Ball, Lubienski y Mewborn, 2001). Para realizar esta evaluación se utilizará un instrumento con evidencia de validez, fiabilidad y de sencilla aplicación, ya probado en unos trabajos anteriores con alumnos de Psicología, y desarrollado por Díaz (2007). Este cuestionario es útil para la finalidad del Estudio 2, ya que proporciona un diagnóstico de las dificultades de los estudiantes, sus conocimientos y capacidad de resolución de problemas. Los resultados proporcionarán la base para el diseño de actividades formativas dirigidas al profesorado, para mejorar su formación específica en la enseñanza de este tema y se utilizarán en los Estudios 3, 4 y 5.

En lo que sigue se describen los objetivos, material y método, contexto educativo y resultados del Estudio 2.

4.2. OBJETIVOS E HIPÓTESIS DEL ESTUDIO 2

Aunque son muy escasas las investigaciones de evaluación del razonamiento probabilístico de futuros profesores de secundaria, los estudios en futuros profesores de primaria (por ejemplo, Azcárate, 1995; Begg y Edward, 1999; Watson, 2001; Pereira-Mendoza, 2002; Batanero, Godino y Cañizares, 2005; Tarr y Lannin, 2005; Ortiz, Nordin, Batanero, Serrano y Rodríguez, 2006), indican que éstos presentan sesgos relacionados con la probabilidad y particularmente con la probabilidad condicional y la independencia. Por otro lado, como se estudió en el capítulo 2, estos sesgos son también frecuentes en los alumnos de secundaria, por lo que los futuros profesores debieran haberlos superado y saber evaluarlos en sus alumnos. En consecuencia, es necesario realizar un estudio de evaluación de las dificultades que los futuros profesores de Educación Secundaria y Bachillerato tienen con la probabilidad condicional.

En concreto, un primer objetivo de este estudio es *evaluar entre los futuros profesores la existencia de las dificultades que se describen a continuación,* indicando el ítem del cuestionario donde se evalúan:

- *Falacia de las tasas base* (ítem 1). Tversky y Kahneman (1982b) investigaron este sesgo como parte de su trabajo sobre la heurística de representatividad. Los autores denominan *falacia de las tasas base* al hecho de ignorar la probabilidad a priori de un suceso en problemas que involucran la probabilidad inversa. Es decir, consiste en calcular una probabilidad condicional como si fuese una probabilidad simple, ignorando algunos datos del problema. Generalmente, este sesgo ocurre debido a una percepción incorrecta de la dependencia entre los sucesos implicados y aparece en problemas bayesianos. Estudios psicológicos muestran estimaciones incorrectas de riesgos asociados a la toma de decisiones (Eddy, 1982; Totohasina, 1992; Teigen, Brun y Frydenlund, 1999). Este tipo de razonamiento también ha sido descrito por Serrano, Batanero, Ortiz y Cañizares (2001) y Díaz y de la Fuente (2007).

- *Confusión entre sucesos independientes y mutuamente incluyentes* (ítem 2). Creer que dos sucesos son independientes si y sólo si son excluyentes es un error muy extendido, incluso entre futuros profesores de secundaria según Sánchez (1996). Investigaciones como la de Kelly y Zwiers (1986) indican que este error puede ser debido a la imprecisión del lenguaje ordinario, en que "independiente" puede significar, a veces, separado. Las dificultades aparecen porque no se identifica con claridad la intersección de los sucesos en el espacio muestral producto (Sánchez, 1996; Truran y Truran, 1996). En ocasiones no es sencillo identificar si dos variables son independientes o están asociadas en situaciones experimentales (Kelly y Zwiers, 1986; Estepa, 1994).
- *Confundir probabilidades conjuntas y condicionales* (ítem 3). Pollatsek, Well, Konold y Hardiman (1987) sugieren que los enunciados que usan la conjunción "y" pueden llevar a confundir una probabilidad conjunta y una probabilidad condicional. Ojeda (1995) indica que los enunciados en los que aparece la intersección de sucesos son más difíciles de interpretar que aquellos donde aparecen la probabilidad condicional, ya que mayoritariamente se interpreta la intersección como condicionamiento.
- *Falacia de la conjunción* (ítem 4). Se denomina *falacia de la conjunción* a la creencia de que es más probable la intersección de dos sucesos que cada uno de ellos por separado (Tversky y Kahneman, 1982a). Díaz (2005), en su estudio sobre la falacia de la conjunción, hace hincapié en la influencia de los sucesos que intervienen en los problemas. Cuando uno de los sucesos tiene una probabilidad muy alta en comparación con el otro, la intersección de los dos se ve más probable que el suceso de mayor probabilidad, por lo que aparece la falacia. Según Tversky y

Kahneman (1982b) el error se produce como resultado de considerar la conjunción como más representativa que cada evento por separado.

- *Falacia de la condicional transpuesta* (ítem 5). Se conoce como falacia de la condicional transpuesta a la confusión de las probabilidades *P(A/B) y P(B/A)* (Falk, 1986), que aparece con frecuencia en contextos médicos (Eddy, 1993; Gotzsche y Olsen, 2000), así como al interpretar erróneamente el significado del nivel de significación en un contraste de hipótesis (Vallecillos, 1994). Otras investigaciones como las de Pollatsek y cols. (1987) y Ojeda (1995) resaltan las dificultades de los estudiantes en el razonamiento de la probabilidad condicional a la hora de interpretar los sucesos condicionantes.
- *Falacia del eje de tiempo* (ítems 6 y 7b). Es la creencia de que un suceso no puede condicionar otro que ocurra anteriormente. Gras y Totohasina (1995) indican que los estudiantes asocian el condicionamiento con el orden temporal de los sucesos y no encuentran natural que se condicione un suceso por otro que ocurre con posterioridad. Denominan a esta dificultad la concepción cronologista de la probabilidad condicional. Ello es debido a una confusión entre condicionamiento y causación. Aunque una relación condicional indica que una relación causal es posible, ésta no es segura. Este sesgo se ha detectado también en las investigaciones de Batanero, Estepa, Godino y Green (1996), Falk (1986) y Díaz (2007).

Un *segundo objetivo* del Estudio 2 sería comparar los resultados de la evaluación en futuros profesores de secundaria, con los obtenidos por Díaz (2007) en estudiantes de Psicología. En el citado trabajo se detectó un alto porcentaje de sesgos relacionados con la probabilidad condicional en tareas que se encuentran con frecuencia en la vida cotidiana, incluso algunas que formalmente se asemejan a otras que aparecen en el proceso de evaluación del aprendizaje. Estos sesgos inciden en la resolución de problemas y en el aprendizaje, por lo que los profesores han de estar formados para poder detectarlos en sus estudiantes. El primer paso para detectarlos será superarlos (en caso que los presenten) y por ello cobra especial importancia su evaluación.

Como *tercer objetivo* del Estudio 2, se desea evaluar la competencia en resolución de problemas de probabilidad condicional por los futuros profesores y su interrelación con los sesgos anteriormente descritos. Como hemos indicado, la probabilidad condicional no sólo tiene importancia en sí misma, sino como base de muchos otros conceptos estadísticos, como la correlación, asociación y regresión, la distribución muestral, intervalos de confianza y contraste de hipótesis. Es importante asegurar una correcta comprensión formal matemática y una capacidad de resolver problemas en estos futuros profesores, si queremos asegurar su competencia en la enseñanza de los temas mencionados, incluidos en el currículo de educación secundaria y Bachillerato.

Por ello, en el cuestionario utilizado en nuestro trabajo, además de los sesgos mencionados, se evalúa el conocimiento formal de la probabilidad condicional y la competencia para la resolución de los tipos de problemas incluidos en la enseñanza en secundaria y Bachillerato:

- *Comprensión conceptual de la probabilidad condicional* (ítem 8). Se pretende evaluar el conocimiento de la definición del concepto, y si el estudiante es capaz de discriminarlo respecto a una probabilidad simple.
- *Cálculo de la probabilidad condicional.* El ítem 9 evalúa la capacidad de resolver problemas de probabilidad condicional en un experimento simple utilizando únicamente la definición. El ítem 7a evalúa el cálculo de probabilidades condicionales en situaciones sin reemplazamiento. El ítem 11 evalúa si los alumnos son capaces de resolver correctamente problemas de probabilidad condicional dentro de un contexto de muestreo con reposición, distinguiendo entre sucesos dependientes e independientes.
- *Cálculo de la probabilidad conjunta*. El ítem 12 evalúa la capacidad de resolver problemas de probabilidad conjunta en una situación sincrónica (sucesos que ocurren de forma simultánea) a partir de la probabilidad simple y condicional. Autores como Falk (1989) y Ojeda (1995) indican que las situaciones sincrónicas son situaciones estáticas, en las que no subyace una secuencia de experimentos, sino que los experimentos aleatorios se realizan simultáneamente. Por otra parte el ítem 14 evalúa el cálculo de la probabilidad compuesta haciendo uso de la regla del producto a partir de sucesos independientes.
- *Teorema de la probabilidad total*. (ítem 10) Evalúa la capacidad de resolver problemas que involucren el teorema de la probabilidad total.
- *Teorema de Bayes*. El ítem 13 pretende conocer si se interpreta correctamente el cálculo de probabilidad condicional en situaciones de probabilidad inversa.

Hipótesis

Las hipótesis que planteamos respecto a los anteriores objetivos en el Estudio 2 son las siguientes:

Esperamos, en primer lugar, encontrar una proporción alta de futuros profesores en la muestra que presenten los sesgos evaluados en el estudio. Nos basamos para apoyar esta hipótesis en la persistencia con que dichos sesgos han sido denunciados en las investigaciones previas con diferentes especialidades de alumnos universitarios. Por otro lado, en la investigación de Díaz (2007), los ítems relacionados con la evaluación de los sesgos no tuvieron capacidad de discriminación entre alumnos con alto o bajo conocimiento matemático respecto a la probabilidad condicional. Es de esperar, por tanto, que los resultados respecto a los sesgos en futuros profesores sean similares a los obtenidos por Díaz (2007) con alumnos de Psicología.

Por el contrario, *esperamos que los resultados en los ítems relacionados con la resolución de problemas en nuestro estudio sean mejores que los encontrados por la citada autora*, debido a la mejor preparación en estadística y probabilidad de los participantes en nuestro estudio. Nos basamos para formular esta hipótesis en las propiedades discriminantes que mostraron los ítems relacionados con la resolución de problemas en el estudio de Díaz (2007).

Finalmente, *la tercera hipótesis de este estudio es la falta de relación entre la competencia matemática de los estudiantes para resolver problemas de probabilidad condicional y los sesgos observados en su razonamiento probabilístico*. Esta hipótesis también se ve apoyada por los resultados del estudio de Díaz (2007) quien observó esta falta de relación en su muestra. Esperamos que esta falta de relación se confirme, asimismo, en futuros profesores.

4.3. MATERIAL Y MÉTODO

El material utilizado en el Estudio 2, que se recoge en el anexo 1 y se analiza con detalle en el apartado 4.6, fue una parte del cuestionario RPC (Razonamiento sobre Probabilidad Condicional), que fue elaborado por Díaz (2004, 2007) y validado por su autora con un riguroso proceso metodológico (Díaz y de la Fuente, 2007).

La recogida de datos se realizó durante el segundo cuatrimestre del curso 2009-2010 en las distintas universidades que forman parte de la muestra. Las instrucciones para cumplimentarlo, tiempo disponible e instrucciones fueron iguales en todos los grupos de estudiantes y los datos se recogieron dentro de una de las asignaturas cursadas por los alumnos. Para la recogida de datos se contactó con algunos profesores del último curso de la Licenciatura de Matemáticas y en la especialidad de matemáticas del Máster Universitario de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanza de Idiomas. En los casos en que las fechas permitieron la recogida de datos, todos los profesores contactados se interesaron en el trabajo, y se ofrecieron para colaborar en la obtención de datos.

A dichos profesores se les mandó el cuestionario, junto con una descripción detallada de su contenido, indicando para cada ítem, las soluciones correctas y causas de las incorrectas. Además se les envió una presentación PowerPoint para que posteriormente pudiesen discutir con los alumnos las soluciones y los posibles fallos y sesgos que se encontrasen. De este modo la actividad de completar el cuestionario, junto con la sesión posterior de corrección y análisis de los sesgos se convirtió en una actividad formativa sobre didáctica de la probabilidad en las universidades y cursos participantes.

El cuestionario se rellenó de forma individual, por escrito, de forma que cada futuro profesor expresase sus conocimientos en probabilidad condicional, tanto en la comprensión conceptual y resolución de problemas como en la superación de posibles sesgos en distintas situaciones. Una vez recogidos los cuestionarios se codificaron los datos siguiendo, en primer lugar, el procedimiento de Díaz (2007) y posteriormente ampliando la categorización en los ítems abiertos.

4.4. CONTEXTO EDUCATIVO

Como se ha indicado, los estudiantes participantes provienen del último curso de la Licenciatura de Matemáticas o del Máster de Educación Secundaria. A continuación se describe de forma resumida el contexto educativo de estos estudiantes, con especial atención a la formación que han recibido sobre probabilidad.

4.4.1. LICENCIATURA EN MATEMÁTICAS

Un primer grupo de estudiantes cursaban quinto curso de la Licenciatura de Matemáticas (actualmente en proceso de transformación a un título de Grado). El plan de estudios de los alumnos participantes, de una duración de cinco cursos académicos varía de una a otra universidad, aunque incluía en todas ellas las materias básicas de matemáticas: Análisis, álgebra, geometría, estadística, y un buen número de asignaturas que se ocupan de los aspectos más aplicados de las matemáticas. La Licenciatura en Matemáticas proporciona una formación adecuada para desarrollar las capacidades analíticas y el pensamiento lógico y riguroso del estudiante. Permite al licenciado utilizar los conocimientos teóricos y prácticos adquiridos, en la definición de problemas y en la búsqueda de soluciones en contextos académicos y empresariales.

Los licenciados en matemáticas están capacitados para trabajar tanto en empresas, instituciones como en la administración pública y en la enseñanza en centros de Educación Secundaria. Los planes de estudios recientes, al incidir en los aspectos prácticos de la disciplina, hacen que estos titulados sean muy valorados para puestos de responsabilidad en empresas e instituciones. Sin embargo, una de las salidas profesionales más frecuentes es la docencia en diversos niveles educativos, entre ellos, Educación Secundaria o Bachillerato. Un problema actual en esta titulación es la disminución de los estudiantes que se interesan por esta licenciatura, lo que también incide en el menor nivel de los estudiantes que acceden a ella.

Respecto al estudio de la probabilidad condicional, los alumnos de la Licenciatura de Matemáticas han cursado varias asignaturas específicas de probabilidad y estadística a lo largo de la carrera en las que se incluye el tema en las primeras lecciones. Seguidamente describimos las principales características de la titulación en las diferentes universidades donde se han recogido datos.

Universidad de Granada

La Licenciatura en Matemáticas se imparte desde 1964 en la Facultad de Ciencias de la Universidad de Granada y en ella se imparten varias asignaturas en las que se trata explícitamente la probabilidad condicional. Algunas de estas asignaturas son de carácter troncal, como la asignatura de segundo curso “Probabilidades y Estadística” de 10,5 créditos o la de tercer curso “Ampliación de Estadística” de 9 créditos, las dos anuales. Otras asignaturas, de carácter optativo, que incluyen la enseñanza de la probabilidad condicional, son las asignaturas de “Seminario de Historia de la Matemática” y “Teoría de la Probabilidad” de segundo ciclo y con seis créditos cada una. Además de estas materias, el alumno tiene la opción de elegir algunas otras asignaturas en las, que aunque no se imparte la probabilidad condicional en el temario, si aparece implícita en los contenidos. Estas asignaturas son “Análisis de Datos” de primer ciclo y “Análisis

Multivariante", "Estadística Computacional", "Inferencia Estadística", "Modelos de la Investigación Operativa" y "Procesos Estocásticos", incluidas en el segundo ciclo.

Universidad de La Laguna

En la Licenciatura en Matemáticas de la Universidad de La Laguna existen dos asignaturas cuatrimestrales que se centran en aspectos probabilísticos. La asignatura de segundo curso "Probabilidades I" y la de tercer curso "Probabilidades II" de 6 créditos cada una. Los contenidos principales de estas asignaturas son los fundamentos matemáticos del cálculo de probabilidades y del tratamiento de los fenómenos de azar y su significado en nuestro lenguaje cotidiano. Además de estas materias, el alumno tiene otras asignaturas de carácter troncal como: "Estadística descriptiva" de primer curso, "Estadística I" de segundo e "Historia de las matemáticas" de tercer curso, donde implícitamente aparece definida la probabilidad condicional dentro del temario.

Universidad de Salamanca

En la Universidad de Salamanca, la Licenciatura en Matemáticas consta de 305 créditos de los cuales sólo 12 son dedicados a la enseñanza de la probabilidad, de forma obligatoria para todos los estudiantes, dentro de la asignatura de "Probabilidades y Estadística" de segundo curso. Existen otras dos asignaturas que tratan la probabilidad de forma explícita, las dos de carácter optativo asociadas al segundo ciclo: La asignatura de "Probabilidad y Medida" y la de "Teoría de la Probabilidad" ambas de 7,5 créditos. Otras asignaturas, también de carácter optativo, tratan la probabilidad de forma implícita. Estas asignaturas son: "Procesos Estocásticos" y "Análisis de Datos Multivariantes", las dos de 7,5 créditos.

4.4.2. MÁSTER UNIVERSITARIO DE EDUCACIÓN SECUNDARIA OBLIGATORIA Y BACHILLERATO, FORMACIÓN PROFESIONAL Y ENSEÑANZA DE IDIOMAS

Recientemente se ha implantado el Máster Universitario de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanza de Idiomas, orientado a los estudiantes que se preparan profesionalmente para la docencia en dichos niveles educativos, y respondiendo a la obligatoriedad de cursar estudios de máster para ejercer la docencia en estos ámbitos, dispuesta en la Ley Orgánica de Educación 2/2006 de 24 de mayo de 2006 y la regulación establecida para estos máster en la Orden 3858/2007 de 27 de diciembre.

Más de 40 universidades españolas imparten este máster, cuya duración es de un año académico (60 créditos) e incide en las competencias que menos dominan los licenciados, como la formación psicológica, pedagógica y sociológica. La estructura de las enseñanzas del citado Máster es común y se estructuran en dos módulos: Uno genérico con formación básica y común para todos los alumnos (entre 12 y 15 créditos según la universidad) y otro específico que desarrolla las competencias relacionadas con las diferentes disciplinas o especialidades docentes (entre 20 y 30 créditos). Además, es obligatorio realizar un periodo de prácticas en centros de Educación Secundaria y un trabajo fin de máster (entre 16 y 20 créditos).

Los contenidos de los módulos varían de una a otra universidad. Aunque cada universidad interpreta de un modo estas asignaturas, en casi todas se imparten algunas sesiones sobre probabilidad y su didáctica, variando la extensión de esta formación; pues algunas universidades dedican hasta 3 créditos completos a la didáctica de la estadística y probabilidad, mientras que en otras el tema sólo se incluye a través de la posibilidad de elegirlo para el trabajo fin de Máster. Describimos brevemente a modo de ejemplo algunos planes de estudio del Máster en Educación de algunas universidades de la muestra.

Universidad de Alicante

En esta universidad el módulo genérico consta de tres asignaturas con un total de 14 créditos: “Aprendizaje y desarrollo de la personalidad” (6 créditos), “Procesos y contextos educativos” (5 créditos), “Sociedad, familia y educación” (3 créditos). El módulo específico para los alumnos de la especialidad de matemáticas (30 créditos), se compone de 3 materias. La primera, “Aprendizaje y enseñanza de las matemáticas” (15 créditos), está dividida en tres partes: “Enseñanza de las matemáticas en Educación Secundaria”, “Aprendizaje de las matemáticas en Educación Secundaria” y “Aproximación didáctica a la resolución de problemas de matemáticas”, cada una de 5 créditos. Las otras dos materias son “Introducción a las metodologías y técnicas básicas de investigación en Didáctica de la Matemática” (9 créditos) e “Innovación en la enseñanza de la matemática e investigación en didáctica de la matemática” (6 créditos).

El tercer módulo está compuesto de dos Prácticas en centros (de 4 y 6 créditos respectivamente), y de un trabajo fin de máster (6 créditos).

Durante el Máster se trata el tema de la probabilidad y su didáctica en las asignaturas: “Enseñanza de las matemáticas en educación secundaria” (donde se pasó el cuestionario), “Aprendizaje de las matemáticas en educación secundaria” (que incluye un tema titulado “Características del aprendizaje de las nociones de azar y probabilidad y manejo de datos”), y “Aproximación didáctica a la resolución de problemas de matemáticas”. En “Innovación en la enseñanza de la matemática e investigación en didáctica de la matemática” se estudia el tema “Introducción a las líneas actuales de investigación en Didáctica de la Matemática” y en “Introducción a las metodologías y técnicas básicas de investigación en Didáctica de la Matemática” se hizo un resumen de una tesis doctoral sobre resolución de problemas de probabilidad condicional, haciendo hincapié en los instrumentos de recogida y de análisis de datos.

Universidad de Cádiz

En la Universidad de Cádiz, el módulo genérico consta de 3 asignaturas con un total de 12 créditos: “Sociedad, Familia y Educación, Aprendizaje y Desarrollo de la Personalidad” y “Procesos y Contextos Educativos” de 4 créditos cada una. El módulo específico para los alumnos de matemáticas costa de 24 créditos, y se compone de 3 materias: “Complementos disciplinares” (6 créditos), “Enseñanza y Aprendizaje” (12 créditos) e “Investigación e Innovación” (6 créditos), junto con otras materias optativas por un total de 8 créditos. El tercer módulo está compuesto de unas prácticas docentes de 10 créditos en centros y de un trabajo fin de máster (6 créditos).

Específicamente no se trata la probabilidad o la didáctica de la probabilidad en ninguna asignatura. En la asignatura “Complementos Disciplinares” se utilizó el análisis de la evolución histórica de la probabilidad como referente para el aula, junto con algunas actividades y ejemplos sobre probabilidad pero dentro del contexto y no como objeto de estudio. Los alumnos, sin embargo, podrían elegir el tema en su trabajo fin de Máster.

Universidad de Granada

En la Universidad de Granada, el módulo genérico está compuesto de tres asignaturas de 4 créditos cada una: “Proceso y contextos educativos”, “Aprendizaje y desarrollo de la personalidad” y “Sociedad, familia y escuela”. El módulo específico compuesto de tres módulos: “Aprendizaje y enseñanza de las materias” (12 créditos), “Complementos de formación disciplinar” (6 créditos) e “Innovación docente e investigación educativa” (6 créditos), en el que el alumno deberá cursar obligatoriamente, por especialidad, las tres materias que componen el módulo.

Dentro del área de Didáctica de la Matemática, en la especialidad de matemáticas, se imparte dos asignaturas: “Aprendizaje y enseñanza de las Matemáticas” (12 créditos) e “Innovación docente e investigación educativa” (4 créditos). Los contenidos de la primera se centran en bloques de contenido, tales como Competencia matemática y análisis curricular, Conocimiento profesional del profesor de matemáticas, Aprendizaje matemático o Características del razonamiento matemático de los alumnos/ as de secundaria. En la segunda se tratan temas de investigación en innovación en educación matemática. La probabilidad condicional podría contemplarse en cualquiera de las dos materias en las actividades prácticas, tanto individuales como colectivas

El módulo de Prácticum está compuesto de dos submódulos: “Prácticas docentes” (10 créditos) en que los estudiantes consiguen una experiencia práctica en enseñanza de la matemática en un centro educativo y el “Trabajo fin de máster” (6 créditos), donde también podría elegirse un tema relacionado con la didáctica de la probabilidad. Por último esta universidad incluye un módulo de libre disposición en el que los alumnos elegirán dos asignaturas de 4 créditos cada una de entre las siguientes: “Atención a la diversidad y multiculturalidad”, “Atención a los estudiantes con necesidades especiales”, “Hacia una cultura de paz”, “Educación para la igualdad” y “Organización y gestión de centros educativos”.

Universidad Pública de Navarra

En esta universidad el módulo genérico costa de 3 asignaturas con un total de 14 créditos: “Aprendizaje y desarrollo de la personalidad” (5 créditos), “Procesos y contextos educativos” (6 créditos) y “Sociedad, familia y educación” (3 créditos). El módulo específico de matemáticas consta de las asignaturas: “Complementos para la formación disciplinar en matemáticas” (Estadística y Probabilidad) de 6 créditos, “Intensificación disciplinar en matemáticas” (Análisis, Álgebra y Geometría) de 6 créditos, “Aprendizaje y enseñanza de las matemáticas” de 9 créditos e “Iniciación a la investigación en didáctica de las matemáticas e innovación” de 6 créditos. También hay que elegir entre dos asignaturas optativas dentro de la especialidad: “Aprendizaje integrado de matemáticas e inglés” y “Aprendizaje de matemáticas basado en problemas y proyectos” de 3 créditos cada una. El tercer módulo está compuesto de dos Practicums, uno de 3 y otro de 7 créditos, para cada cuatrimestre del máster, donde el alumno consigue una experiencia práctica en enseñanza de la matemática. Además realiza un trabajo fin de máster (6 créditos).

Respecto a la docencia sobre de la probabilidad y su didáctica, esta se imparte como parte de la asignatura “Complementos para la formación disciplinar en matemáticas (Estadística y Probabilidad)” donde se dedican 1,5 créditos de los 6 previstos. En “Aprendizaje y enseñanza de las matemáticas”, se dedica 1 crédito de los 9 previstos y en “Iniciación a la investigación en didáctica de las matemáticas e innovación” 0,5 créditos de los 6 previstos.

Universidad de Santiago de Compostela

En este caso el módulo genérico costa de 5 asignaturas con un total de 16 créditos: “Didáctica, Currículum y Organización Escolar” (4,5 créditos), “La función tutorial y la orientación académica” (2.5 créditos), “Educación y lenguas en Galicia” (2 créditos), “Desarrollo psicológico y aprendizaje escolar” (3,5 créditos) y “Educación, sociedad y política educativa” (3,5 créditos).

El módulo específico de la especialidad en Ciencias Experimentales, Matemáticas, y Tecnología e Informática consta de las siguientes materias: “La ciencia y su metodología para el profesorado de Educación Secundaria” (3 créditos), “Diseño de investigaciones educativas en ciencias en la Educación Secundaria” (3 créditos), “Matemáticas para el profesorado de Educación Secundaria” (5 créditos), “Análisis didáctico de la Aritmética, el Álgebra y el Análisis en la Educación Secundaria” (6 créditos), “Análisis didáctico de la Geometría, la medida de Magnitudes, la Probabilidad y la Estadística en la Educación Secundaria” (6 créditos) y “Diseño, Planificación y Evaluación de Propuestas Didácticas en Matemáticas en la Educación Secundaria” (3 créditos). El tercer módulo está compuesto de dos Practicums, uno de 4 y otro de 8 créditos, para cada cuatrimestre del máster, donde el alumno consigue una experiencia práctica en enseñanza de la matemática. Además realiza un trabajo fin de máster (6 créditos).

En la asignatura “Análisis didáctico de la geometría, la medida de magnitudes, la probabilidad y la estadística en la educación secundaria" de 6 créditos se dedican 2,5 a la didáctica de la probabilidad y estadística. Como contenido tratan la problemática didáctica del proceso de enseñanza-aprendizaje de la combinatoria, probabilidad y estadística en educación secundaria, las concepciones de los estudiantes, sus obstáculos y dificultades.

Universidad de Salamanca

En Salamanca el módulo genérico consta de 5 asignaturas obligatorias comunes a todas las especialidades: "Psicología de la Educación", "Atención a la diversidad en Educación", "Organización e Historia del Sistema Educativo", "Orientación Educativa (tutorial y familiar)" y "Sociología de la Educación", todas ellas de 3 créditos.

El módulo específico de Matemáticas consta de las Materias: "Didáctica de la especialidad de Matemáticas", "Recursos en la especialidad de Matemáticas", "Innovación docente en la especialidad de Matemáticas", "Iniciación a la investigación educativa en la especialidad de Matemáticas", "Historia de la especialidad de Matemáticas", "Evaluación en la especialidad de Matemáticas", "Metodología en la especialidad de Matemática" y "Contenidos en el contexto de la especialidad de matemáticas", todas ellas de 3 créditos. El tercer módulo está compuesto del "Practicum de Observación en la especialidad de Matemáticas" y del "Practicum de Intervención en la especialidad de Matemáticas", de 6 créditos cada uno, donde el alumno consigue una experiencia práctica en enseñanza de la matemática. Además realiza un trabajo fin de máster (6 créditos).

La probabilidad como tal no se trata en ninguna de las materias del máster. Se dedica un crédito a la Estadística en la asignatura "Contenidos en el contexto de la especialidad de matemáticas".

4.5. DESCRIPCIÓN DE LA MUESTRA

En la muestra del estudio participaron 196 futuros profesores que, como muestra la Tabla 4.5.1, están divididos en 95 alumnos de la Licenciatura de Matemáticas y 101 alumnos del Máster Universitario de Educación Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanza de Idiomas.

Tabla 4.5.1. Sujetos encuestados por estudios

	Frecuencia (%)
Matemáticas	95 (48,5)
Máster de secundaria	101 (51,5)

Las universidades donde se tomó la muestra fueron las de Alicante, Cádiz, Barcelona, Extremadura, Granada, La Laguna, Navarra, Salamanca, Santiago de Compostela y Valladolid, lo que supone una variedad de comunidades autónomas y situación geográfica. El mayor número de participantes corresponde a estudiantes de la Licenciatura de Matemáticas de dos cursos sucesivos de la Universidad de Granada, como se describen en la Tabla 4.5.2, con un 35,2%. El porcentaje del resto de universidades está comprendido entre el 3,1% de la Universidad Pública de Navarra y el 10,2% de la Universidad de Salamanca.

La submuestra de alumnos de matemáticas ha sido tomada en tres universidades. Mayoritariamente los cuestionarios han sido contestados por alumnos de la universidad de Granada (60%), el resto pertenece la Universidad de la La Laguna (23,2%) y de Salamanca (16,8%). La submuestra de alumnos del Máster de Secundaria, ha sido tomada en las Universidades de Alicante, Cádiz, Barcelona, Extremadura, Granada, Navarra, Salamanca, Santiago de Compostela y Valladolid, destacando la Universidad de Santiago de Compostela y la de Barcelona con un 16,8% y un 15,8% de participantes.

Tabla 4.5.2. Sujetos encuestados por Universidad y tipo (Frecuencia y porcentaje)

	Matemáticas	Máster Secundaria	Total
Universidad de Alicante	-	14 (13,9)	14 (7,1)
Universidad de Barcelona	-	16 (15,8)	16 (8,2)
Universidad de Cádiz	-	12 (11,9)	12 (6,1)
Universidad de Santiago de Compostela	-	17 (16,8)	17 (8,7)
Universidad de Extremadura	-	13 (12,9)	13 (6,6)
Universidad de Granada	57 (60,0)	12 (11,9)	69 (35,2)
Universidad de La Laguna	22 (23,2)	-	22 (11,2)
Universidad Pública de Navarra	-	6 (5,9)	6 (3,1)
Universidad de Salamanca	16 (16,8)	4 (4,0)	20 (10,2)
Universidad de Valladolid	-	7 (6,9)	7 (3,6)

Al analizar los estudios previos de los alumnos del Máster, vemos que la mayoría de ellos son licenciados en Matemáticas (44,5%), pero también encontramos alumnos de diferentes ramas relacionadas con las ingenierías, arquitectura, etc. como podemos ver en la Tabla 4.5.3. Esto puede ocasionar un problema añadido, al no tener todos los estudiantes la misma formación inicial, y tener una menor preparación matemática que sus compañeros. También pueden tener una menor vocación hacia la enseñanza, pues su elección inicial de formación se orientaba a otras profesiones y por diversos motivos han considerado la posibilidad de dedicarse a la enseñanza.

Tabla 4.5.3. Alumnos de Máster en Educación según titulación de acceso

	Frecuencia	Porcentaje
Matemáticas	45	44,5
Arquitectura	6	5,9
Ingeniería Industrial	11	10,9
Ingeniería de Montes	4	4,0
Ingeniería Informática	6	5,9
Licenciatura de Economía	6	5,9
Ingeniería Química	8	7,9
Ingeniería Técnica en Electricidad	1	1,0
Ingeniería de Telecomunicaciones	5	5,0
Ingeniería Aeronáutica	1	1,0
Licenciatura en Ciencias y Técnicas Estadísticas	2	2,0
Licenciatura de Física	2	2,0
Administración y Dirección de Empresas	4	4,0

4.6. ANÁLISIS DEL CUESTIONARIO

El cuestionario utilizado es el RPC (Razonamiento sobre Probabilidad Condicional) elaborado por Díaz (2007) construido con un proceso riguroso, siguiendo las normas psicométricas APA, AERA y NCME (1985, 1999) para este tipo de trabajos. A continuación se describe la construcción de dicho cuestionario y se analizan los ítems elegidos para esta investigación.

4.6.1. PROCESO DE CONSTRUCCIÓN

En el proceso de construcción del cuestionario de Díaz (2007) se siguieron las siguientes fases:

Fase 1. Especificación del contenido de la variable objeto de medición. Se basó en la revisión bibliográfica de las investigaciones previas, que permitió conocer los principales errores relacionados con este concepto y a su vez recopilar algunos ítems usados en dichas investigaciones. Una definición más precisa de la variable se fundamentó en el análisis del contenido sobre probabilidad condicional reflejado en una muestra de 21 libros de texto de estadística aplicada, elegida entre los que se usan en la

Licenciatura de Psicología para las asignaturas de análisis de datos.

Fase 2. Construcción de la versión piloto del cuestionario RPC. Este estudio consta de dos etapas: En la primera (pruebas piloto de ítems), un total de 49 ítems recogidos de investigaciones previas y de libros de texto se probaron en muestras de entre 49 y 117 estudiantes de primer curso de Psicología, para determinar su índice de dificultad, discriminación y la distribución de respuestas entre los distractores. En una segunda etapa se hizo la selección final de unidades de contenido y de los ítems del cuestionario a partir de un juicio de expertos (participaron 10 expertos en estadística o didáctica de la estadística).

Fase 3. Prueba piloto del cuestionario RPC. Una vez que se dispuso de una primera versión del instrumento, se llevó a cabo una prueba piloto del cuestionario RPC con nuevas muestras de 37 y 57 estudiantes, para obtener una primera determinación de sus características psicométricas. También se estimó el tiempo necesario para completarlo (entre una hora y hora y cuarto).

Fase 4. Revisión de la prueba piloto. Se revisó el instrumento piloto a partir de juicio de expertos (8 expertos en metodología) para mejorar su legibilidad y disminuir la dificultad de algunos ítems, finalizando, de este modo, el proceso de construcción del cuestionario.

4.6.2. CARACTERÍSTICAS PSICOMÉTRICAS

Fiabilidad

Para comprobar las características psicométricas del cuestionario RPC se estudió la fiabilidad de consistencia interna del test, mediante el cálculo del coeficiente Alfa de Crobach, que dio un valor de 0,797. Díaz (2007) informa que fue estadísticamente significativa la correlación corregida ítem-test en todos los ítems, oscilando sus valores entre 0,154 y 0,624, excepto en dos ítems (falacia de la condicional transpuesta y falacia del eje de tiempos) en que fue próxima a cero. La autora explica esta falta de discriminación por el hecho de que estas falacias afectan por igual a estudiantes con bajos o altos conocimientos formales sobre probabilidad condicional.

Los coeficientes de fiabilidad basados en el análisis factorial dieron los siguientes valores: Theta de Carmines θ=0,82; Omega de Heise y Bohrnsted, Ω=0,896, cumpliéndose la relación: $\alpha < \theta < \Omega$. Estos valores, junto con el hecho de que el primer autovalor en el análisis factorial explicó mucho mayor porcentaje de varianza que los siguientes, y que la mayoría de los ítems tenían pesos apreciables en el mismo antes de la rotación, apoyan una cierta unidimensionalidad del cuestionario y por tanto, la validez de constructo (Carmines y Zéller, 1979). Los índices de dificultad varían entre 0,25 (ítem 6) y 0,84 (ítem 8 e ítem 2).

En un segundo estudio se obtuvo un alto valor de los coeficientes de correlación en la fiabilidad test-retest, tanto si fue obtenida con la correlación de Pearson (0,876) como con la de Spearman (0,867), oscilando las correlaciones del mismo ítem en las dos aplicaciones entre 0,299 y 0,799 y siendo todas ellas estadísticamente significativas. Todo ello asegura una alta estabilidad del cuestionario, medido por la fiabilidad test-retest (Carmines y Zéller, 1979).

Validez

La extracción inicial del análisis factorial produce 7 factores con autovalor mayor que 1, que explicaron el 54,5% de la varianza total. El primer factor explica un 22,3% de la varianza, mientras que los siguientes entre el 4,6% y el 9,98% cada uno. La importancia relativa del primer factor proporciona una evidencia de *validez de constructo*, puesto que el porcentaje es claramente superior al explicado por los restantes (Morales, 1988).

El primer factor (de mayor peso relativo) incluye la mayoría de los ítems de respuesta abierta. En particular el problema relacionado con el Teorema de Bayes, presenta la contribución mayor, seguido del relacionado con la probabilidad total y los de probabilidad compuesta. Todos estos problemas requieren un proceso de resolución de al menos dos pasos, el primero de los cuales es el cálculo de una probabilidad condicional, que se utiliza seguidamente en un segundo paso (ej. en la regla del producto). La autora interpreta este factor como la habilidad de resolver problemas complejos de probabilidad condicional.

El segundo factor explica un 6,2% de la varianza. El cálculo de una probabilidad simple, probabilidad conjunta y probabilidad condicional aparece en este componente separado, probablemente porque el formato de la tarea afectará a la ejecución, como indican Ojeda (1995) y Gigerenzer (1994) entre otros investigadores.

Un tercer factor (7,94% de la varianza) muestra la relación entre la definición, el espacio muestral y el cálculo de probabilidades condicionales en situaciones con y sin reemplazamiento; habilidades que se alcanza en el nivel máximo de comprensión de la clasificación de Tarr y Jones (1997). La autora denomina a este factor “comprensión conceptual de la probabilidad condicional”. El conjunto de los tres primeros factores (resolución de problemas complejos, lectura de tablas y comprensión conceptual) englobaría la comprensión formal de la probabilidad condicional.

El resto de factores describen los diferentes sesgos que afectan al razonamiento condicional, que no se relacionan con la comprensión formal, apoyando nuestra hipótesis en este sentido. Cada uno de los sesgos (condicional transpuesta, falacia del eje temporal, falacia de la conjunción, confusión entre independencia y mutua exclusividad, así como falacia de tasas base) aparecen separados de los primeros factores y entre sí. Algunos de ellos aparecen opuestos o relacionados con algunas unidades semánticas del razonamiento probabilístico condicional. Por ejemplo, la independencia está ligada a la falacia de las tasas base (donde los sujetos tienen que juzgar si los sucesos son independientes o no) y opuesta a la idea de dependencia.

En el estudio 3, la prueba Lambda de Wilks (0,631) y correlación canónica (0,697) indican el poder discriminante del conjunto de ítems y es un indicio de *validez de criterio* del cuestionario RPC respecto al criterio dicotómico de haber recibido o no instrucción sobre el tema. Todos los ítems obtuvieron una diferencia estadísticamente significativa en las pruebas univariantes a favor del grupo con instrucción, excepto los ítems 2 y 4 (falacia del eje de tiempos) y 3 (falacia de la condicional transpuesta) relacionados con sesgos en el razonamiento condicional, que parecen no mejorar con la instrucción. La *validez de criterio* también se apoya por los resultados del análisis discriminante, que muestra un 82% de clasificaciones correctas de alumnos con y sin instrucción. La autora justifica finalmente la *validez de contenido* del cuestionario mediante el juicio de expertos realizado y mediante el análisis de contenido de los ítems, que cubren el contenido pretendido. En resumen, el cuestionario muestra unas buenas propiedades de validez y fiabilidad, y se dispone de datos de comparación de una alta muestra de estudiantes de la Licenciatura de Psicología, por lo que hemos decidido usarlo.

4.6.3. ANÁLISIS A PRIORI DE LOS ÍTEMS DEL CUESTIONARIO

En este apartado llevamos a cabo un análisis teórico de las respuestas esperadas a cada uno de los ítems que constituyen el instrumento. Para cada ítem analizamos la respuesta correcta y los sesgos contenidos en los distractores. Asimismo describimos el contenido evaluado por el ítem, que consta de dos partes: (a) Contenido primario, aquél para cuya evaluación se ha elaborado el ítem; (b) contenido secundario, o contenido relacionado que también se requiere para resolver el ítem.

4.6.3.1. ÍTEMS DE OPCIÓN MÚLTIPLE Y SESGOS EVALUADOS

Ítem 1. Un taxi se vio implicado en un accidente nocturno con choque y huida posterior. Hay dos compañías de taxis en la ciudad, la Verde y la Azul. El 85% de los taxis de la ciudad son Verde y el 15% Azul. Un testigo identificó el taxi como Azul. El tribunal comprobó la fiabilidad del testigo bajo las mismas circunstancias que había la noche del accidente y llegó a la conclusión de que el testigo identificaba correctamente cada uno de los colores en el 80% de las ocasiones y fallaba en el 20%. ¿Cuál es la probabilidad de que el taxi implicado en el accidente fuera en efecto Azul?

a. 80 %
b. 15%
c. (15/100) x (80/100)
d. 41 %

El ítem se ha incluido para evaluar el sesgo conocido como "*falacia de las tasas base*", consistente en ignorar la probabilidad a priori del suceso en la toma de decisiones en problemas de probabilidad inversa y que en el estudio de Díaz (2007) apareció en el 15% de sujetos. Este tipo de razonamiento también se ha encontrado en investigaciones en educación matemática (Serrano, Batanero, Ortiz y Cañizares, 2001). Este sesgo aparece en las repuestas (a) y (b), que corresponden a estudiantes que no tienen en cuenta todos los datos del problema. Por otro lado el distractor (c) corresponde a aquellos que confunden probabilidad condicional con la probabilidad conjunta.

La respuesta correcta de este ítem es la (d) 41%. Para determinarla se debe aplicar el teorema de Bayes. Dicho teorema calcula la probabilidad condicionada de un suceso *A* dado otro suceso *B*, con probabilidad distinta de cero, mediante la siguiente expresión:

$$P(A/B)=\frac{P(A\cap B)}{P(B)}=\frac{P(B/A)P(A)}{P(B)}\,.$$

Previamente ha de calcular $P(B)$ mediante la fórmula de la probabilidad total:

$$P(B)=P(B/A)P(A)+P(B/\overline{A})P(\overline{A})\,.$$

Aplicado a este caso, la solución sería la siguiente: El alumno ha de interpretar que, para calcular la probabilidad de que el taxi implicado en el accidente fuera en efecto azul, tenemos que calcular la probabilidad de que sea azul condicionada a la fiabilidad del testigo. Denominemos *A* al suceso "*Color de taxi azul*" y *C* al suceso "*Fiabilidad al identificar correctamente el color*". Por tanto la probabilidad condicionada viene dada por $P(A/C)=\frac{P(A\cap C)}{P(C)}$, siendo la probabilidad de la fiabilidad del testigo $P(C)=P(C/A)P(A)+P(\bar{C}/\bar{A})P(\bar{A})$, es decir la probabilidad de acertar multiplicada por la probabilidad de ser azul más la probabilidad de fallar multiplicada por la probabilidad de no ser azul.

$$P(A/C)=\frac{P(A\cap C)}{P(C)}=\frac{P(C/A)P(A)}{P(C/A)P(A)+P(\bar{C}/\bar{A})P(\bar{A})}=\frac{0,8\cdot 0,15}{0,8\cdot 0,15+0,2\cdot 0,85}=0,41.$$

Por tanto, el resultado de que el taxi implicado fuese azul es del 41%.

En resumen, este ítem tiene como contenido primario la *"falacia de las tasas base"*. Como contenidos secundarios aparecen: "*Aplicar correctamente el cálculo de probabilidad condicional en situaciones de probabilidad inversa (Teorema de Bayes)"*, "*aplicar correctamente la regla de la probabilidad total"*, *"resolver problemas de probabilidad condicional a partir de probabilidades conjuntas y simples"*, *"resolver correctamente problemas condicionales cuando se invierte el eje de tiempos"*, *"distinguir probabilidad condicional con inversa"* y "*resolver correctamente problemas en situaciones sincrónicas"*.

Ítem 2. Se extrae una carta al azar de una baraja española (40 cartas con números del 1 al 7, sota caballo y rey): sea A el suceso "se extrae una carta de oros" y B el suceso "se extrae un rey" ¿Los sucesos A y B son independientes?

a. No son independientes porque en la baraja hay un rey de oros.
b. Sólo si sacamos primero una carta para ver si es rey y se vuelve a colocar en la baraja y luego sacamos una segunda carta para mirar si es oros.
c. No, porque P(rey de oros)= P(rey) x P(oros)
d. Sí, en todos los casos.

Este ítem tiene como objetivo evaluar si los futuros profesores son capaces de *"distinguir entre sucesos independientes y mutuamente excluyentes"*. Este error, descrito por Kelly y Zwiers (1986), consiste en confundir estos términos, lo que es debido, según los autores a la imprecisión del lenguaje ordinario, en el que "independiente" puede significar, a veces, "separado". La respuesta correcta sería la descrita en el apartado (d) ya que una de las definiciones de sucesos independientes afirma que: "*Dos sucesos son independientes cuando la probabilidad de la intersección de dos sucesos es igual al producto de las probabilidades simples, es decir:* $P(A \cap B) = P(A) \times P(B)$". Esta propiedad se cumple en el ítem, puesto que:

$$P(A \cap B) = \frac{1}{40};\; P(A) \times P(B) = \frac{10}{40}\frac{4}{40} = \frac{1}{40}.$$

Por tanto los sucesos *A* y *B* son independientes. Los distractores tratan de detectar diferentes errores descritos en la literatura:

- En el distractor (a) se trata de detectar la posible confusión entre sucesos excluyentes y sucesos independientes descrito por Kelly y Zwiers (1986), puesto que los sucesos no son excluyentes, aunque si independientes.
- En el distractor (b) se trata de detectar la creencia errónea que sólo pueden ser independientes los sucesos de experimentos diacrónicos. Totohasina (1992) indica que algunos estudiantes tienen dificultad en reconocer la independencia en sucesos sincrónicos como en este problema.
- En el distractor (c) se presenta la definición correcta de la regla del producto, junto con la afirmación incorrecta que esta regla no se cumple en el caso de sucesos independientes. En este caso se cumple porque los sucesos son independientes.

En resumen, el contenido primario de este ítem es "*distinguir entre sucesos independientes y mutuamente excluyentes*". El alumno debe también saber "*resolver correctamente problemas de probabilidad compuesta haciendo uso de la regla del producto en experimentos independientes*", que sería el contenido secundario.

Ítem 3. La probabilidad de que una mujer de 40 años tenga una mamografía positiva es el 10.3%. La probabilidad de que una mujer de 40 años tenga cáncer de pecho y una mamografía positiva es de 0.8%. Una mujer de 40 años se realiza una mamografía y resulta positiva. ¿Cuál es la probabilidad de que realmente tenga cáncer?

a. $\frac{0,8}{10,3} = 0,0776$ **, probabilidad de 7,76%**

b. $10,3 \times 0,8 = 8,24$, probabilidad del 8,24%

c. 0,8 %

Este ítem se incluye para comprobar si los estudiantes confunden la probabilidad conjunta y condicional, sesgo descrito por Pollatsek, Well, Konold y Hardiman (1987), y Einhorn y Hogarth (1986) en la interpretación de los enunciados de los problemas de probabilidad. Tversky y Kahneman (1982b) y Díaz (2005) lo describen en relación con la falacia de la conjunción o creencia de que es más probable la intersección de dos sucesos que cada uno de ellos por separado.

Para resolverlo, el alumno debe identificar los datos del problema e interpretar las diferentes probabilidades. Si denotamos por + al suceso “mujer de 40 años con una mamografía positiva” y *C* al suceso “mujer de 40 años con cáncer de pecho”. Entonces tenemos que $P(+) =$ 0,103 y $P(C \cap +) = 0{,}008$. Además debe recordar la fórmula para calcularla: $P(C/+) = \frac{0'8}{10.3} = 0'0776$.

- Si el alumno elige el distractor (b), está aplicando la regla del producto, en lugar de dividir las probabilidades, por tanto no recuerda la fórmula de la probabilidad condicional.
- Si el alumno elige el distractor (c) confunde probabilidad conjunta y condicional, por tanto otro contenido secundario sería *“confundir probabilidad conjunta, condicional y simple”*.

El contenido primario de este ítem es “*resolver correctamente problemas de probabilidad condicional a partir de probabilidades conjuntas y simples”*. Como contenidos secundarios encontramos “*resolver correctamente problemas de probabilidad condicional en situaciones sincrónicas” y* “*confundir probabilidad conjunta, condicional y simple”*.

Ítem 4. Supón que Rafa Nadal alcanza la final de Roland Garros en 2004. ¿Cuál de los siguientes sucesos consideras más probable?

a. **Rafa Nadal pierde el primer set**
b. Rafa Nadal pierde el primer set pero gana el partido
c. Los dos sucesos son iguales de probables

Este ítem evalúa el sesgo conocido como “*falacia de la conjunción*”, descrito por Tversky y Kahneman (1982b), que consiste en la creencia de que es más probable la intersección de dos sucesos que cada uno de ellos por separado. La respuesta correcta es la (a) ya que la probabilidad conjunta se puede descomponer como producto de dos probabilidades. Estas dos probabilidades son simples si los sucesos que intervienen en la probabilidad conjunta son mutuamente independientes. Por tanto, tenemos el producto de dos números entre cero y uno, por lo que el resultado siempre va a ser menor que cada una de las probabilidades por separado.

- Si el alumno elige el distractor (b) está usando la "falacia de la conjunción", asignando mayor valor a la probabilidad conjunta que a la simple de cada uno de los sucesos que intervienen.
- Si el alumno elige el distractor (c) está considerando como equiprobables la probabilidad simple y conjunta. Este razonamiento erróneo, definido como "*sesgo de equiprobabilidad*", ha sido descrito por Lecoutre y Durand (1988), Lecoutre y Cordier (1990) y Lecoutre (1992).

En resumen, este ítem corresponde al contenido primario "*falacia de la conjunción*" y como secundario "*distinguir entre probabilidad conjunta, condicional y simple*" ya que el alumno ha de interpretar las opciones (a) y (b) respectivamente como probabilidades simples y conjuntas.

Ítem 5. Un test diagnóstico de cáncer fue administrado a todos los residentes en una gran ciudad. Un resultado positivo en el test es indicativo de cáncer y un resultado negativo es indicativo de ausencia de cáncer. ¿En cuál de las siguientes predicciones tienes más confianza?

a. Predecir que una persona tiene cáncer si ha dado positivo en el test de diagnóstico.
b. Predecir un resultado positivo en el test de diagnóstico si la persona tiene cáncer.
c. Tengo la misma confianza en ambas predicciones.

Este ítem evalúa la "*falacia de la condicional transpuesta*", descrita por Falk (1986) como la incorrecta discriminación entre las dos direcciones de la probabilidad condicional $P(A/B)$ y $P(B/A)$. Eddy (1982) expone que este error se observa en enunciados de contextos médicos, donde se confunde la probabilidad de tener una enfermedad cuando el test de diagnóstico resulta positivo, con la probabilidad de un resultado positivo en el test de diagnóstico, dado que se tiene la enfermedad.

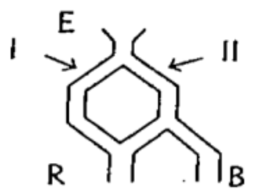
El resultado (b) es el más probable ya que este tipo de test se diseña para detectar una cierta enfermedad. La probabilidad de que una persona sana tenga un resultado positivo (falso positivo) en la prueba es muy pequeña, pero no imposible y el alto número de personas sanas en la población hace que la probabilidad de estar enfermos si el test ha sido positivo no sea, en general, demasiado alta (Eddy, 1982).

- Si el alumno elige el distractor (a) está incurriendo en la denominada como "falacia de la condicional traspuesta", ya que está confundiendo la condición con el condicionado (probabilidad condicional con su inversa).
- Si el alumno elige el distractor (c) está incurriendo en un error, ya que afirma que la probabilidad de que tenga cáncer es igual a la probabilidad de que el test de positivo, es decir $P(A/B)=P(B/A)$ si y sólo sí $P(A)=P(B)$. Si esto ocurriera el test sería fiable al 100% cosa que no es cierta ya que en tal caso no habrían falsos positivos. Es decir se manifiesta confusión entre una situación condicional y una diagnóstica.

En resumen este ítem evalúa como contenido primario la "*falacia de la condicional transpuesta*" y como contenido secundario si el alumno "*distingue entre probabilidad condicional y su inversa*" (distractor a).

Ítem 6. Una bola se suelta en E. Si sale por R, ¿Cuál es la probabilidad de que haya pasado por el canal I?

a. 0'50
b. 0'33
c. 0'66
d. No se puede calcular



Este ítem evalúa la "*falacia del eje de tiempo*", descrita por Falk (1989), que consiste en creer que un suceso no puede condicionar otro que ocurra anteriormente. Este problema también ha sido descrito como concepción cronologista de la probabilidad condicional por Gras y Totohasina (1995). Autores como Estrada, Díaz y de la Fuente (2006) indican que la comprensión de la relación de condicionalidad se dificulta si la secuencia temporal de los sucesos no coincide con el orden dado en el condicionamiento. El resultado (c) es el correcto ya que para resolver este ítem, el alumno ha de tener en cuenta que si la bola pasa por el canal *II* puede caer por *R* o por *B*, y si pasa por *I*, sólo puede caer por *R*. El espacio muestral de este experimento sería: {(*I*,*R*), (*II*,*R*), (*II*,*B*)} siendo:

$$P(I,R)=\frac{1}{2},\ P(II,R)=\frac{1}{4}\ \text{y}\ P(II,B)=\frac{1}{4}.$$

Por tanto la probabilidad condicionada viene dada por:

$$P(I/R)\ =\ P(I,R)/\ P(R)\ =\ \frac{1/2}{3/4}=\ \frac{2}{3}.$$

Otro razonamiento es que a *R* llega el doble de bolas desde *I* que desde *II*. Luego, $P(I/R) = 2/3$.

- Si se elige el distractor (a) no se está teniendo en cuenta las bolas que caen por el orificio *B*, es decir no se tiene en cuenta el condicionamiento por un suceso posterior, haciendo una incorrecta restricción del espacio muestral.
- Si se elige el distractor (b) se está confundiendo el suceso condicionado ya que se está calculando la probabilidad de que habiendo salido la bola por *R*, la bola haya pasado por *II*.
- Si se elige el distractor (d) se comete la "falacia del eje de tiempos", ya que se piensa que no se puede calcular la probabilidad de un suceso condicionado a otro que ocurra anteriormente.

En resumen este ítem evalúa como contenido primario si los alumnos son capaces de *"resolver correctamente problemas condicionales cuando se invierte el eje del tiempo (falacia del eje de tiempos)" y "resolver correctamente problemas de probabilidad condicional en situaciones diacrónicas"*. Como objetivo secundario, el contenido *"resolver correctamente problemas de probabilidad condicional a partir de probabilidades conjuntas y simples"*.

Ítem 7.a. Una urna contiene dos bolas blancas y dos bolas negras. Extraemos a ciegas dos bolas de la urna, una detrás de otra, sin reemplazamiento. ¿Cuál es la probabilidad de extraer una bola negra en segundo lugar, habiendo extraído una bola negra en primer lugar? P (N2/ N1)

a. 1/ 2
b. 1/ 6
c. 1/ 3
d. 1/ 4

Este ítem evalúa la capacidad de resolver problemas de probabilidad condicional en un contexto de muestreo sin reposición. El alumno ha de tener en cuenta la composición de la urna una vez sacada la bola negra, es decir, el hecho de que *la probabilidad condicional supone una restricción del espacio muestral.* Una vez sacada una bola negra de la urna tenemos una bola negra y dos blancas, por lo que la probabilidad de obtener una bola negra es igual a 1/3.

- Si el alumno eligió el distractor (a) está confundiendo muestreo con y sin reposición, ya que está considerando que la extracción de la bola negra no modifica el espacio muestral y, por tanto, la probabilidad es igual a 1/2.

- Si el alumno eligió el distractor (b) está confundiendo la probabilidad condicional con la conjunta ya que está aplicando la regla del producto: (1/2)(1/3)=1/6.
- Si el alumno eligió el distractor (d) está confundiendo probabilidad condicional y probabilidad simple, ya que sólo está considerando la probabilidad de obtener una bola negra en la primera extracción (1/4).

En resumen, este ítem incluye como contenido primario "*resolver correctamente problemas en contexto de muestreo sin reposición*" y como secundarios "*resolver correctamente problemas de probabilidad condicional en situaciones diacrónicas*", así como "*distinguir sucesos dependientes e independientes*".

Ítem 7.b. ¿Cuál es la probabilidad de extraer una bola negra en primer lugar, sabiendo que hemos extraído una bola negra en segundo lugar? P (N1/ N2)

a. 1/ 3
b. 1/ 4
c. 1/ 6
d. 1/ 2

Este ítem evalúa la falacia del eje de tiempo (Falk, 1986) descrita anteriormente. Este problema parte de una inferencia inversa, lo que requiere un razonamiento probabilístico que es indiferente al orden temporal (Díaz, 2007). Para resolverlo, el alumno tiene que entender la diferencia entre muestreo con reposición y sin reposición. Además tiene que darse cuenta de que si en segundo lugar, hay una bola negra, ésta queda eliminada en la primera extracción, por lo que el espacio muestral de la primera extracción queda restringido. Por tanto la solución parte de que al conocer que la segunda bola es negra el espacio muestral se restringe a $\{(n,n);(b_1,n);(b_2,n)\}$ y por tanto $P(n,n)=\frac{1}{3}$.

- Si el alumno elige el distractor (d) no se está teniendo en cuenta el condicionamiento a un suceso que ocurre posteriormente, ya que se calcula la probabilidad simple, de que la primera bola sea negra. Con este distractor se evalúa la "falacia del eje temporal" al no tener en cuenta un suceso posterior.
- Si el alumno elige el distractor (b) está confundiendo probabilidad condicional con probabilidad conjunta, ya que está calculando la probabilidad conjunta de sacar una

bola negra en primer lugar y una bola negra en segundo lugar.

Este ítem tiene como contenido primario *"resolver correctamente problemas de probabilidad condicional cuando se invierte el eje de tiempo"* y como secundarios "*resolver correctamente problemas en contexto de muestreo sin reposición", "reconocer que la probabilidad condicional supone una restricción del espacio muestral"* y *"distinguir entre probabilidad condicional, conjunta y simple"*.

4.6.3.2. ITEMS ABIERTOS: DEFINICIÓN Y PROBLEMAS DE PROBABILIDAD CONDICIONAL

Ítem 8. Explica con tus propias palabras la diferencia entre una probabilidad simple y una probabilidad condicional.

Este ítem permite valorar la comprensión del contenido *"definición de la probabilidad condicional"*, contenido secundario en todos los ítems del instrumento, puesto que es necesario conocerla para poder calcularla. Además el contenido *"distinguir probabilidad conjunta, condicional y simple"* sería un contenido secundario.

Consideramos correctas las respuestas donde el alumno, o bien da una definición correcta o pone un ejemplo adecuado de cada una de las probabilidades. También admitimos los estudiantes que indican las diferencias entre los conceptos o ejemplos en que se vean claramente las diferencias.

Ítem 9. Hemos lanzado dos dados y sabemos que el producto de los dos números obtenidos ha sido 12 ¿Cuál es la probabilidad de que ninguno de los dos números sea un 6?

Para calcular la probabilidad pedida en un sólo experimento, el alumno debe reconocer que *"la probabilidad condicional supone una restricción del espacio muestral"*, ya que sólo debe tomar como casos posibles el conjunto {(2, 6) (3,4), (4,3) y (6,2)}. La probabilidad pedida es igual a 1/2 ya que sólo hay dos casos favorables. Maury (1986) indica que parte de la dificultad de los problemas se debe a que los alumnos no son capaces de identificar y restringir correctamente el espacio muestral en la probabilidad condicional.

Para resolver correctamente el problema, el alumno debe ser capaz de enumerar el espacio muestral obtenido al lanzar dos dados (36 casos posibles), como indica la Figura 4.6.1, identificando los casos en que el producto es 12. Debe ser también capaz de identificar la situación como de "muestreo sin reposición" puesto que cada valor obtenido del primer dado puede repetirse en el segundo. Tiene también que tener en cuenta que el orden debe considerarse al resolver el problema.

Figura 4.6.1. Espacio muestral de lanzar dos dados

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

En consecuencia, este ítem corresponde al contenido primario "*calcular la probabilidad condicional en un sólo experimento,* y como secundarios "*la probabilidad condicional supone una restricción del espacio muestral*", y "*calcular la probabilidad condicional en muestreo sin reposición*".

Ítem 10. En una ciudad hay 60 hombres y 40 mujeres por cada 100 habitantes. 50 de cada 100 hombres y 25 de cada 100 mujeres fuman. Si cogemos 200 personas al azar, ¿cuántas de ellas fumarán?

El objetivo de este problema es evaluar la capacidad del estudiante para resolver problemas que involucren la fórmula de la probabilidad total. Denotando como *F* al suceso "fumar", *M* al suceso "ser mujer" y *H* al suceso "ser hombre" y aplicando la fórmula de la probabilidad total tenemos que la solución de este ítem es:

$$P(F) = P(H \cap F) + P(M \cap F) = P(H)P(F/H) + P(M)P(F/M)$$

$$P(F) = 0,6 \cdot 0,5 + 0,4 \cdot 0,25 = 0,4 .$$

Por tanto el número esperado de personas que fuman, de un total de 200 sería: $200 \cdot 0,4 = 80$.

Otra forma para resolver este problema es utilizando una tabla de doble entrada como la Tabla 4.6.1, donde se trabaja con datos enteros y hay que aplicar una proporción para calcular los datos de las celdas. Si de cada 100 personas fuman 40, de 200 personas, fumarán 80.

Tabla 4.6.1. Tabla de doble entrada del problema

	Fuma	No fuma	
Hombres	30	30	60
Mujeres	10	30	40
	40	60	100

En consecuencia, el contenido primario sería *"aplicar correctamente la regla de la probabilidad total"*. Además, como contenidos secundarios, se pretende conocer si son capaces de *"calcular la probabilidad compuesta haciendo uso de la regla del producto en caso de sucesos dependientes", "resolver correctamente problemas en situaciones sincrónicas"* y *"distinguir probabilidad condicional con inversa"*.

Ítem 11. Una persona lanza un dado y anota si saca un número par o impar. Estos son los resultados:

par, impar, impar, par, par, impar, par, par, par, par, impar, impar, par, par, par.

Lanza el dado de nuevo. ¿Cuál es la probabilidad de sacar un número par en la siguiente tirada?

Este ítem pretende analizar la *"resolución de problemas de cálculo de probabilidad condicional en contexto de muestreo con reposición"*, pues cada lanzamiento del dado se puede considerar como extracción al azar de un número del 1 a 6. También se puede comprobar si se da en los alumnos la *"falacia del jugador"*, esto es, considerar que para el cálculo de la probabilidad pedida se debe tener en cuenta los resultados anteriores, es decir pensar que porque un hecho no ha ocurrido en el pasado es más fácil que ocurra en el futuro.

La solución, al tratarse de experimentos independientes, es de 1/2 ya que tendríamos un caso favorable de dos posibles y los diferentes lanzamientos son independientes unos de otros. El ítem requiere que el alumno comprenda que las ocurrencias anteriores no afectan la probabilidad de un suceso cuando los experimentos son independientes, como es el caso del muestreo sin reposición.

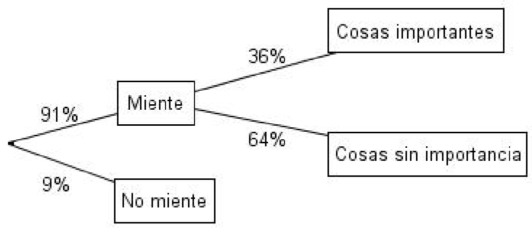
Como no se especifica si el dado es sesgado, se podría admitir también la solución 10/15, si el alumno estima la probabilidad mediante la frecuencia relativa. Por tanto este ítem tiene como contenido primario *"resolver correctamente problemas en contexto de muestreo con reposición"* y como secundario *"resolver correctamente problemas de* *probabilidad condicional en situaciones diacrónicas”*, así como *“distinguir entre sucesos dependientes e independientes”*.

Ítem 12. En una encuesta publicada en un periódico se informa que el 91% de los habitantes de una ciudad mienten usualmente y de ellos el 36% mienten sobre cosas importantes. Si nos encontramos con una persona al azar de esta ciudad, ¿cuál es la probabilidad de que mienta sobre cosas importantes?

Este ítem pretende comprobar si los alumnos son capaces de *“calcular una probabilidad conjunta en una situación sincrónica”* en caso de sucesos dependientes. Para resolver el problema, el alumno podría utilizar un diagrama en árbol como el presentado en la Figura 4.6.2. Los sucesos son dependientes, pues si una persona no miente, tampoco miente sobre cosas importantes. Por tanto aplicando la regla del producto para el caso de sucesos dependientes, tenemos:

$$P(I) = P(M \cap I) = P\left(M\right)P\left(I/M\right) = 0{,}91 \cdot 0{,}36 = 0{,}327.$$

Figura 4.6.2. Diagrama de árbol del problema



Observamos que el estudiante ha de utilizar la fórmula de probabilidad condicional, aplicándola directamente, una vez identificados los sucesos de interés y sus probabilidades. En resumen, este ítem evalúa como contenido primario *“resolver correctamente problemas de probabilidad condicional en situaciones sincrónicas” y “resolver correctamente problemas de probabilidad compuesta en caso de sucesos dependientes”*, y como secundario, el contenido *“calcular una probabilidad condicional dentro de un único experimento”*.

Ítem 13. Una fábrica dispone de dos máquinas M1 y M2 que fabrican bolas. La máquina M1 fabrica el 40 % de las bolas y la M2 el 60%. El 5% de las bolas fabricadas por M1 y el 1% de las fabricadas por M2 son defectuosas. Tomamos una bola al azar que resulta ser defectuosa. ¿Cuál es la probabilidad de que haya sido fabricada por M1?

Este ítem pretende conocer si los alumnos reconocen la necesidad del uso del teorema de Bayes y de la fórmula de la probabilidad total para resolver el problema, a la vez que comprobamos si utilizan correctamente la fórmula de ambas. Además se quiere conocer si los alumnos presentan la falacia de la tasa base. Por tanto el alumno ha de utilizar correctamente la fórmula del teorema de Bayes, definida como:

$$P(M_1 / D) = \frac{P(D \cap M_1)}{P(D)} = \frac{P(D / M_1)P(M_1)}{P(D)}.$$

Previamente para calcular $P(D)$ ha de hacer uso de la fórmula de la probabilidad total: $P(D) = P(D / M_1)P(M_1) + P(D / M_2)P(M_2)$. El resultado correcto por tanto es:

$$P(D) = 0,4 \cdot 0,05 + 0,6 \cdot 0,01.$$

$$P(M_1 / D) = \frac{0,4 \cdot 0,05}{0,4 \cdot 0,05 + 0,6 \cdot 0,01} = 0,7692.$$

Como vemos, el ítem tiene como contenido primario *"aplicar correctamente el cálculo de probabilidad condicional en situaciones de probabilidad inversa (Teorema de Bayes)"*. Y como contenidos secundarios los siguientes: *"aplicar correctamente la regla de la probabilidad total"*, *"resolver problemas de probabilidad condicional a partir de probabilidades conjuntas y simples"*, *"resolver correctamente problemas condicionales cuando se invierte el eje de tiempos"* y *"resolver correctamente problemas en situaciones sincrónicas"*.

Ítem 14. Un grupo de alumnos de un colegio se examina de Matemáticas e Inglés. La proporción de alumnos que aprueban Matemáticas es del 80% y la proporción de alumnos que aprueba Inglés es del 70%. Suponiendo que las notas en cada una de las asignaturas son independientes, ¿Cuál es la probabilidad de que un alumno escogido al azar haya aprobado ambas asignaturas?

Este ítem pretende conocer si los alumnos son capaces de calcular la probabilidad conjunta para sucesos independientes. Por tanto para resolver este ítem hay que hacer uso de la regla del producto, teniendo en cuenta que los sucesos son independientes entre sí. Utilizando la propiedad de independencia de dos sucesos, dos sucesos son independientes si y sólo si cumplen que la probabilidad conjunta es igual al producto de las probabilidades simples de cada uno de los sucesos:

$$P\left(M \cap I\right) = P\left(M\right) \cdot P\left(I\right) = 0,8 \cdot 0,7 = 0,56.$$

En resumen, este ítem evalúa como contenido principal *"calcular la probabilidad compuesta haciendo uso de la regla del producto en caso de sucesos independientes"*, y como secundarios *"distinguir entre sucesos dependientes e independientes"*, y *"resolver correctamente problemas de probabilidad compuesta haciendo uso de la regla del producto en caso de sucesos dependientes"*.

Resumen de contenidos evaluados

Tabla 4.6.2. Contenidos primarios y secundarios evaluados por los ítems

Contenido	Primario	Secundario
1. Definición de la probabilidad condicional; diferenciar de simple y compuesta	8	Todos
2. Reconocer que la probabilidad condicional supone una restricción del espacio muestral	9	7a, 7b
3. Distinguir una probabilidad conjunta y su inversa; falacia de la condicional transpuesta	5	1, 10
4. Distinguir entre probabilidad condicional y conjunta	3	
5. Probabilidad conjunta menor que probabilidad simple; falacia de la conjunción	4	
6. Distinguir sucesos independientes, dependientes y mutuamente excluyentes	2	7a, 11, 14
7. Calcular una probabilidad condicional dentro de un único experimento	9	12
8. Resolver correctamente problemas de probabilidad condicional en un contexto de muestreo con reposición	11	9
9. Resolver correctamente problemas de probabilidad condicional en un contexto de muestreo sin reposición	7a	7b
10. Resolver correctamente problemas de probabilidad condicional a partir de probabilidades conjuntas y simples	3	1, 6, 7a, 11, 13, 14
11. Resolver correctamente problemas condicionales cuando se invierte el eje de tiempo; falacia del eje de tiempos	6, 7b	13
12. Resolver correctamente problemas en situaciones diacrónicas	6	7a, 11
13. Resolver correctamente problemas en situaciones sincrónicas	12	1, 3, 10, 13
14. Resolver correctamente problemas de probabilidad compuesta haciendo uso de la regla del producto en caso de sucesos independientes	14	2
15. Resolver correctamente problemas de probabilidad compuesta haciendo uso de la regla del producto en caso de sucesos dependientes	12	10, 13
16. Aplicar correctamente el cálculo de la probabilidad condicional en situaciones de sucesos múltiples (regla de la probabilidad total)	10	1, 13
17. Aplicar correctamente el cálculo de la probabilidad condicional en situaciones de probabilidad inversa (regla de Bayes)	13	1
18. Falacia de las tasas base	1	

En la Tabla 4.6.2 resumimos los distintos contenidos que se evalúan en los ítems que incluiremos en el instrumento, diferenciando si el contenido de interés del ítem se consideraba como primario o secundario. En total hemos considerado 18 contenidos, de modo que cada contenido primario es representado por uno o dos ítems. Además, algunos de estos contenidos aparecen como secundarios en otros. Como vemos, los ítems elegidos para llevar a cabo la evaluación permiten estudiar, tanto el conocimiento formal, representado por el conocimiento de la definición y la capacidad de resolver problemas, como los sesgos de razonamiento.

4.7. RESULTADOS

En este apartado se muestran los resultados obtenidos en las respuestas de los futuros profesores a los diferentes ítems. Presentamos por un lado, la comparación de los datos en nuestra muestra de alumnos con la proporcionada por Díaz (2007). También se comparan los resultados para los dos grupos de nuestra muestra por separado (alumnos de la Licenciatura de Matemáticas con alumnos de Máster), junto con un análisis del número de sesgos dados por los alumnos.

4.7.1. SESGOS EN EL RAZONAMIENTO SOBRE PROBABILIDAD CONDICIONAL

Comenzamos el análisis de resultados con el estudio de los ítems de opción múltiple, en la mayoría de los cuáles se evalúa la existencia de algunos sesgos sobre el razonamiento condicional. Comentaremos los resultados en cada uno de los distractores que componen el ítem.

Resultados en el ítem 1. Falacia de las tasas base

Ítem 1. Un taxi se vio implicado en un accidente nocturno con choque y huida posterior. Hay dos compañías de taxis en la ciudad, la Verde y la Azul. El 85% de los taxis de la ciudad son Verde y el 15% Azul. Un testigo identificó el taxi como Azul. El tribunal comprobó la fiabilidad del testigo bajo las mismas circunstancias que había la noche del accidente y llegó a la conclusión de que el testigo identificaba correctamente cada uno de los colores en el 80% de las ocasiones y fallaba en el 20%. ¿Cuál es la probabilidad de que el taxi implicado en el accidente fuera en efecto Azul?

a. 80 %
b. 15%
c. (15/100) x (80/100)
d. 41 %

Compararemos en primer lugar los resultados en este ítem en el global de nuestra muestra (Tabla 4.7.1) con los resultados de Díaz (2007) en estudiantes de Psicología. El ítem resultó difícil, pues solamente el 7,7% de los futuros profesores fue capaz de dar una respuesta correcta, a la cuestión planteada. El resultado ha sido mucho peor que los obtenidos por Díaz, donde un 49,8% de estudiantes de Psicología respondieron correctamente al problema; por tanto Díaz obtuvo un 42,1% más de aciertos que los obtenidos en nuestro estudio.

Tabla 4.7.1. Resultados en el ítem 1

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. 80 %	49	25,0	33	8,0
b. 15%	29	14,8	29	7,0
c. (15/100) x (80/100)	100	51,0	121	29,2
d. 41 %	15	7,7	206	49,8
En blanco.	3	1,5	25	6,0
Total	196	100,0	414	100,0

El 39,8% de los estudiantes de nuestra muestra eligen las repuestas (a) y (b), que indican *la falacia de las tasas base*. El 25% de ellos respondió dando la solución (a) en la que sólo se tiene en cuenta el porcentaje de identificaciones correctas por parte del testigo y el 14,8% optó por la solución (b), en la que sólo se tiene en cuenta el porcentaje de taxis azules de la ciudad, por lo que ambas respuestas indican que este grupo de alumnos interpretan el problema como de probabilidad simple. Los resultados fueron mejores en los estudiantes de Psicología, quienes debido a sus estudios de psicología del razonamiento, pueden haber leído alguna literatura relacionada con los sesgos en el razonamiento estocástico, lo que es poco probable en los estudiantes de nuestra muestra. En todo caso, nuestros resultados son mejores que los obtenidos por Pollatsek y cols. (1987), quienes encontraron que el 69% de los sujetos de su estudio presentaban esta falacia.

Fue muy alta la proporción de participantes en nuestra muestra (51%) que respondió a la alternativa (c), confundiendo en este problema probabilidad condicional y conjunta. Esta fue la segunda respuesta mayoritaria de alumnos de Psicología 29,2%, aunque también con menor frecuencia que en los futuros profesores. Los alumnos de Psicología dejan más respuestas en blanco en este ítem que los de nuestra muestra, aunque el porcentaje es, de todos modos, pequeño.

En segundo lugar compararemos los resultados en los dos grupos de nuestra muestra (Tabla 4.7.2), los alumnos que estudian Matemáticas y los alumnos del Máster. No se observan diferencias en el porcentaje de alumnos que respondieron correctamente (7,4% en los alumnos de Matemáticas y 7,9% en los alumnos del Máster). Si existe un poco más de diferencia en los alumnos que respondieron incorrectamente al distractor (b), ya que el porcentaje de alumnos del Máster que eligió este distractor fue un 12,4%

mayor. El total de alumnos que manifiestan la falacia de las tasas base es algo menor en los estudiantes de la Licenciatura de Matemáticas (31,6% de los alumnos de matemáticas y 47,5% de los alumnos del Máster). Por el contrario, los alumnos de Matemáticas tienden mayoritariamente (un 58,9%) a confundir la solución con un problema de probabilidad conjunta, lo que también ocurre con los alumnos del máster pero en menor medida (un 43,6%).

Tabla 4.7.2. Resultados en el ítem 1 (Matemáticas y Máster de Secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. 80 %	22	23,2	27	26,7
b. 15%	8	8,4	21	20,8
c. (15/100) x (80/100)	56	58,9	44	43,6
d. 41 %	7	7,4	8	7,9
En blanco.	2	2,1	1	1,0
Total	95	100,0	101	100,0

Resultados en el ítem 2. Confusión entre sucesos excluyentes e independientes

Ítem 2. Se extrae una carta al azar de una baraja española (40 cartas con números del 1 al 7, sota caballo y rey): sea A el suceso "se extrae una carta de oros" y B el suceso "se extrae un rey" ¿Los sucesos A y B son independientes?

a. No son independientes porque en la baraja hay un rey de oros.
b. Sólo si sacamos primero una carta para ver si es rey y se vuelve a colocar en la baraja y luego sacamos una segunda carta para mirar si es oros.
c. No, porque P(rey de oros)= P(rey) x P(oros)
d. Sí, en todos los casos.

Como vimos anteriormente, este ítem evalúa si los alumnos distinguen sucesos dependientes, independientes y excluyentes. En la muestra global un 26% de participantes, (Tabla 4.7.3) interpretan correctamente el problema dando la solución (d). La proporción es muy similar a la obtenida por Díaz, cuyos estudiantes respondieron correctamente en un 29%. En ambos casos es una proporción muy pequeña, teniendo en cuenta que el contexto de juegos de cartas es bien conocido y que basta aplicar directamente la definición de independencia para dar la respuesta correcta.

El 32,1% de los futuros profesores y el 28,5% de los psicólogos confundieron la solución correcta con el distractor (a), por lo que se trata de alumnos que tienen problemas a la hora de diferenciar entre sucesos excluyentes y sucesos independientes. Este resultado está de acuerdo con las investigaciones de Kelly y Zwiers (1986), quienes sugieren que una posible causa del alto porcentaje de alumnos que cometen este sesgo puede ser debido a la imprecisión del lenguaje ordinario. La proporción fue algo mayor en los futuros profesores aunque no demasiado.

Tabla 4.7.3. Resultados en el ítem 2

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. No son independientes porque en la baraja hay un rey de oros.	63	32,1	118	28,5
b. Sólo si sacamos primero una carta para ver si es rey y se vuelve a colocar en la baraja y luego sacamos una segunda carta para mirar si es oros.	64	32,7	61	14,7
c. No, porque P(rey de oros) = P(rey) x P(oros)	17	8,7	82	19,8
d. Sí, en todos los casos.	51	26,0	120	29,0
En blanco.	1	0,5	33	8,0
Total	196	100,0	414	100,0

El 32,7% de futuros profesores y el 14,7% de los psicólogos respondieron mediante el distractor (b), es decir se trata de alumnos que creen que solamente pueden ser independientes los sucesos de experimentos diacrónicos. Las situaciones sincrónicas dificultan el cambio de espacio muestral, como comprueba Ojeda (1995) al plantear el problema a alumnos de entre 14 y 16 años, el 60% de los cuáles dio una respuesta errónea. El porcentaje menor, en el caso de los futuros profesores (8,7%) e intermedio en el caso de los psicólogos (19,8%), fue para el distractor (c). Estos alumnos son aquellos que interpretan incorrectamente la regla del producto, ya que afirman que no se ha de cumplir esta regla en el caso de sucesos independientes.

Como resumen vemos que, en general, hay problemas a la hora de diferenciar entre independencia, dependencia y sucesos excluyentes ya que un 74% de los futuros profesores y un 71% de los psicólogos erraron en la solución correcta.

En la Tabla 4.7.4 tenemos los resultados respecto a los dos grupos de nuestra muestra. Observamos que el porcentaje de respuestas correctas es un 15,8% mayor en los alumnos del máster, ya que los alumnos de matemáticas se decantaron mayoritariamente por los distractores (a) y (b). El porcentaje de errores en ambos grupos es muy alto, aunque los alumnos de matemáticas hay mayor confusión entre sucesos excluyentes e independientes, pues un 16,3% se decantaron por el distractor (a). Ocurre lo mismo con el distractor (b), un 4% mayor en el caso de los alumnos de la licenciatura, es decir creen que solamente pueden ser independientes los sucesos de experimentos diacrónicos.

Tabla 4.7.4. Resultados en el ítem 2 (Matemáticas y Máster de Secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. No son independientes porque en la baraja hay un rey de oros.	39	41,1	24	23,8
b. Sólo si sacamos primero una carta para ver si es rey y se vuelve a colocar en la baraja y luego sacamos una segunda carta para mirar si es oros.	33	34,7	31	30,7
c. No, porque P(rey de oros) = P(rey) x P(oros)	6	6,3	11	10,8
d. Sí, en todos los casos.	17	17,9	34	33,7
En blanco.	-	-	1	1,0
Total	95	100,0	101	100,0

Resultados en el ítem 3. Confusión entre probabilidad condicional y conjunta

Ítem 3. La probabilidad de que una mujer de 40 años tenga una mamografía positiva es el 10.3 %. La probabilidad de que una mujer de 40 años tenga cáncer de pecho y una mamografía positiva es de 0.8%. Una mujer de 40 años se realiza una mamografía y resulta positiva. ¿Cuál es la probabilidad de que realmente tenga cáncer?

a. $\frac{0,8}{10,3}=0,0776$**, probabilidad de 7,76%**

b. $10,3\times 0,8=8,24$, probabilidad del 8,24%

c. 0,8 %

Con este ítem se pretende evaluar la confusión entre probabilidad conjunta y condicional, error descrito, entre otros, por Einhorn y Hogarth (1986), Falk (1986) y Totohasina (1992). La Tabla 4.7.5 muestra los resultados, que son mejores que los de Díaz pues el 43,9% de nuestra muestra y el 33,8% de la de psicología dieron una respuesta correcta a este ítem. Por tanto se tratan de alumnos que identifican correctamente los datos del problema e interpretan las diferentes probabilidades para construir la probabilidad condicional $P(C/+)$ y a su vez recuerdan la fórmula para calcularla: $P(C/+)=\frac{0,8}{10,3}=0,0776$.

El 23,5% de la muestra de futuros profesores y el 30,7% de psicólogos optó por el distractor (b), por lo que están aplicando la fórmula del producto, en lugar de dividir las probabilidades. Un 31,6% de la muestra de futuros profesores y el 32,6% de la de psicólogos eligió el distractor (c), es decir tienen problemas a la hora de distinguir entre probabilidad conjunta y condicional y vemos que los resultados son similares en ambas muestras. Este porcentaje está de acuerdo con el estudio de Einhorn y Hogarth (1986) los cuales sugieren que esto puede ser debido a la mala interpretación de la conjunción "*y*" en los enunciados. Falk (1986) achaca la prevalencia de este error al uso de un lenguaje ambiguo en el enunciado de los problemas de probabilidad condicional.

Tabla 4.7.5. Resultados en el ítem 3

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. $\frac{0,8}{10,3}=0,0776$, probabilidad de 7,76%	86	43,9	140	33,8
b. $10,3\times 0,8=8,24$, probabilidad del 8,24%	46	23,5	127	30,7
c. 0,8 %	62	31,6	135	32,6
En blanco.	2	1,0	12	2,9
Total	196	100,0	414	100,0

En la Tabla 4.7.6 se comparan los dos grupos de nuestra muestra. En ella observamos que el porcentaje de respuestas correctas es similar, con un 45,5% en alumnos del Máster y un 42,1% en los alumnos de matemáticas. En el distractor (b) (olvido de la fórmula) el porcentaje de alumnos de matemáticas que eligió dicha respuesta casi dobla al de alumnos del máster con un 30,5%. En cambio el porcentaje de alumnos del Máster que confunde probabilidad condicional con probabilidad conjunta, distractor (c), es mayor que los alumnos de la licenciatura en un 10,3%.

Tabla 4.7.6. Resultados en el ítem 3 (Matemáticas y Máster de Secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. $\frac{0,8}{10,3}=0,0776$, probabilidad de 7,76%	40	42,1	46	45,5
b. $10,3\times 0,8=8,24$, probabilidad del 8,24%	29	30,5	17	16,9
c. 0,8 %	25	26,3	37	36,6
En blanco.	1	1,1	1	1,0
Total	95	100,0	101	100,0

Resultados en el ítem 4. Falacia de la conjunción

Ítem 4. Supón que Rafa Nadal alcanza la final de Roland Garros en 2004. ¿Cuál de los siguientes sucesos consideras más probable?

a. Rafa Nadal pierde el primer set
b. Rafa Nadal pierde el primer set pero gana el partido
c. Los dos sucesos son iguales de probables

Este ítem evalúa la falacia de la conjunción, (Tversky y Kahneman, 1982b). En este caso la solución correcta es la (a), ya que al ser sucesos independientes esta va a ser menor que la probabilidad dada en el distractor (b), en la que intervienen los sucesos ganar el segundo y tercer set. Como vemos en la Tabla 4.7.7, el 57,1% de los alumnos de la muestra global dio una solución correcta; por tanto estos alumnos distinguen correctamente entre probabilidad conjunta, condicional y simple. Los resultados son bastante mejores en este ítem que los de Díaz, cuyo porcentaje de respuestas correctas fue sólo del 24,9%, un 32,2% menor que nuestra muestra.

Tabla 4.7.7. Resultados en el ítem 4

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. Rafa Nadal pierde el primer set.	112	57,1	103	24,9
b. Rafa Nadal pierde el primer set pero gana el partido.	23	11,8	39	9,4
c. Los dos sucesos son iguales de probables.	58	29,6	256	61,8
En blanco.	3	1,5	16	3,9
Total	196	100,0	414	100,0

Un 11,7% de nuestra muestra y un 9,4% de la de Psicología optaron por la opción (b); se tratan de estudiantes que están incurriendo en la falacia de la conjunción. En consecuencia, la falacia de la conjunción tiene poca incidencia en ambos estudios. Un 29,6% de nuestra muestra y un 61,8% de la de Psicología eligieron la opción (c), manifestando el sesgo de equiprobabilidad descrito en los experimentos de Lecoutre (1992). Esta autora describe la creencia de los sujetos en la equiprobabilidad de todos los sucesos asociados a cualquier experimento aleatorio, incluso en aquellos en que no es aplicable el principio de indiferencia o donde no hay una simetría física. Este sesgo también se observó en el estudio de Serrano (1996) con futuros profesores de primaria.

Tabla 4.7.8. Resultados en el ítem 4 (Matemáticas y Máster de Secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. Rafa Nadal pierde el primer set.	40	42,1	72	71,3
b. Rafa Nadal pierde el primer set pero gana el partido.	13	13,7	10	9,9
c. Los dos sucesos son iguales de probables.	40	42,1	18	17,8
En blanco.	2	2,1	1	1,0
Total	95	100,0	101	100,0

En la Tabla 4.7.8 presentamos los resultados para los dos grupos de nuestra muestra. En este ítem, el porcentaje de alumnos del Máster que responden correctamente es del 71,3%, un 29,2% mayor que las respuestas correctas del grupo de alumnos de la licenciatura, por lo que vemos que, como en otros ítems anteriores, el alto conocimiento matemático de estos alumnos no los prepara para enfrentarse a sus intuiciones incorrectas en el campo de la probabilidad. El porcentaje de alumnos que incurren en la falacia de la conjunción es bajo en ambos grupos con un 13,7% y un 9,9% respectivamente. Por el contrario el porcentaje de alumnos de la licenciatura que manifiesta el sesgo de equiprobabilidad es 24,3% mayor que la del máster, confirmando nuestra anterior opinión.

Resultados en el ítem 5. Falacia de la condicional transpuesta

Ítem 5. Un test diagnóstico de cáncer fue administrado a todos los residentes en una gran ciudad. Un resultado positivo en el test es indicativo de cáncer y un resultado negativo es indicativo de ausencia de cáncer. ¿En cuál de las siguientes predicciones tienes más confianza?

a. Predecir que una persona tiene cáncer si ha dado positivo en el test de diagnóstico.
b. Predecir un resultado positivo en el test de diagnóstico si la persona tiene cáncer.
c. Tengo la misma confianza en ambas predicciones.

Este ítem intenta evaluar si se incurre en la falacia de la condicional transpuesta. Como muestra la Tabla 4.7.9, nuestros resultados son en este caso mejores que los de Díaz, pues de la muestra global un 41,3% dio la respuesta correcta (b), por lo que, además de distinguir entre las distintas situaciones, también son capaces de reconocer la diferencias entre probabilidad condicional y su inversa. En la muestra de Díaz el porcentaje de aciertos fue menor (32,1%).

Tabla 4.7.9. Resultados en el ítem 5

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. Predecir que una persona tiene cáncer si ha dado positivo en el test de diagnóstico.	27	13,8	24	5,8
b. Predecir un resultado positivo en el test de diagnóstico si la persona tiene cáncer.	81	41,3	133	32,1
c. Tengo la misma confianza en ambas predicciones.	83	42,3	245	59,2
En blanco.	5	2,6	12	2,9
Total	196	100,0	414	100,0

El 42,3% y 59,2% de las respectivas muestras optó por el distractor (c), por tanto manifiestan confusión entre situación condicional y diagnóstica, mostrando la falacia de la condicional transpuesta, con alta frecuencia en ambas muestras, aunque algo menor en la nuestra. Estos resultados también se obtienen en Batanero, Estepa, Godino y Green (1996), aunque con menor frecuencia que en nuestro estudio.

Finalmente algunos sujetos que eligieron el distractor (a), el 13,8% de nuestra muestra y el 5,8% de la de Díaz, incurrieron en la falacia de la condicional transpuesta. Estos resultados son semejantes, aunque menores, a los dados por Batanero, Estepa, Godino y Green (1996).

En la Tabla 4.7.10 observamos que el porcentaje que responde correctamente al ítem 5 es un 5,6% mayor en la muestra de alumnos de la licenciatura. En cambio, el porcentaje de alumnos que incurren en la falacia de la conjunción es algo menor en los alumnos de la licenciatura con un 14,7% por un 12,9% de los alumnos del Máster. El porcentaje de alumnos que manifiestan confusión entre situaciones condicionales y diagnósticas es mayor en los estudiantes del máster (45,5%), siendo un 38,9% de los alumnos de la licenciatura. Las proporciones en el distractor (a) son semejantes.

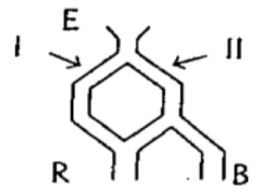
Tabla 4.7.10. Resultados en el ítem 5 (Matemáticas y Máster de Secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. Predecir que una persona tiene cáncer si ha dado positivo en el test de diagnóstico.	14	14,7	13	12,9
b. Predecir un resultado positivo en el test de diagnóstico si la persona tiene cáncer.	42	44,3	39	38,6
c. Tengo la misma confianza en ambas predicciones.	37	38,9	46	45,5
En blanco.	2	2,1	3	3,0
Total	95	100,0	101	100,0

Resultados en el ítem 6. Falacia del eje de tiempos

Ítem 6. Una bola se suelta en E. Si sale por R, ¿Cuál es la probabilidad de que haya pasado por el canal I?

a. 0’50
b. 0’33
c. 0’66
d. No se puede calcular



Con este ítem se pretendía evaluar si los futuros profesores son capaces de resolver correctamente problemas de probabilidad condicional en situaciones donde invierte el eje del tiempo (Falk, 1986). Como muestra la Tabla 4.7.11, solamente un 23,5% de los alumnos de la muestra global fue capaz de dar una respuesta correcta al problema, por lo que esta falacia parece muy arraigada en nuestro estudio.

Comparando con la muestra de Díaz, vemos que el porcentaje de respuestas correctas dada por los alumnos de Psicología es mucho menor con un 9,9% frente a un 23,5%. Destaca el 67,3% de nuestra muestra y el 76,8% de la de Díaz que dio como respuesta el distractor (a), indicando que no se ha tenido en cuenta las bolas que caen por *B*, es decir no se ha tenido en cuenta el condicionamiento por un suceso posterior. El 0,5% de nuestra muestra y el 2,7% de la de Díaz incurrieron en la falacia del eje de tiempo reflejada por el distractor (d). Por tanto, agrupando los resultados de los distractores (a) y (d), vemos que el 68,3% de los sujetos de nuestra muestra y el 78,5% de la muestra de Díaz incurrió en la falacia del eje de tiempos. El distractor (b) dado por un 7,7% y el 8,9% indica que se han considerado equiprobables los tres caminos, cayendo otra vez en el sesgo de equiprobabilidad (Lecoutre, 1992).

Si tenemos en cuenta nuestros dos grupos, Tabla 4.7.12, vemos que los resultados son semejantes en ambos. El porcentaje de respuestas correctas es similar en los alumnos del Máster y de la licenciatura con un 23,1 % y un 23,8% respectivamente. Destaca el 67,4% del grupo de la licenciatura y el 69,3% del máster que incurrió en la falacia del eje de tiempo.

Tabla 4.7.11. Resultados en el ítem 6

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. 0,50	132	67,3	318	76,8
b. 0,33	15	7,7	37	8,9
c. 0,66	46	23,5	41	9,9
d. No se puede calcular.	2	1,0	7	1,7
En blanco.	1	0,5	11	2,7
Total	196	100,0	414	100,0

Tabla 4.7.12. Resultados en el ítem 6 (Matemáticas y Máster de Secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. 0,50	62	65,3	70	69,3
b. 0,33	9	9,5	6	5,9
c. 0,66	22	23,1	24	23,8
d. No se puede calcular.	2	2,1	0	0
En blanco.	0	0	1	1,0
Total	95	100,0	101	100,0

Resultados en el ítem 7a

Ítem 7.a. Una urna contiene dos bolas blancas y dos bolas negras. Extraemos a ciegas dos bolas de la urna, una detrás de otra, sin reemplazamiento. ¿Cuál es la probabilidad de extraer una bola negra en segundo lugar, habiendo extraído una bola negra en primer lugar? P (N2/ N1)

a. 1/ 2
b. 1/ 6
c. 1/ 3
d. 1/ 4

Este ítem pretendía ver si los futuros profesores eran capaces de resolver correctamente problemas en un contexto de muestreo sin reposición y como comparación con la segunda parte (7b) que evalúa la falacia del eje de tiempo. En la Tabla 4.7.13 observamos que el porcentaje mayoritario de alumnos de nuestra muestra (64,3%), fue capaz de contestar correctamente al problema. Comparando estos datos con los de Díaz, vemos que se asemeja el porcentaje de respuestas correctas, aunque este es un poco mayor con un 68,8%. El problema fue sencillo en ambos grupos.

Un 3,6% de sujetos de nuestra muestra y un 5,6% de la de Díaz confundieron muestreo con y sin reposición ya que eligieron el distractor (a), considerando que la extracción de la bola negra no modifica el espacio muestral y, por tanto, identificaban la probabilidad como 1/2. Otro 26,5% y 16,9%, respectivamente, confundieron la probabilidad condicional y la conjunta (distractor b) aplicando la regla del producto (1/2)(1/3)=1/6, por lo que no distinguen entre probabilidad conjunta, condicional y simple. Y por último un 4,1% y un 7,3%, respectivamente, confunde probabilidad condicional y simple (distractor d) ya que consideran únicamente la probabilidad simple de obtener una bola negra en la primera extracción (1/4).

Tabla 4.7.13. Resultados en el ítem 7.a

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. 1/ 2	7	3,6	23	5,6
b. 1/ 6	52	26,5	70	16,9
c. 1/ 3	126	64,3	285	68,8
d. 1/ 4	8	4,1	30	7,3
En blanco.	3	1,5	6	1,4
Total	196	100,0	414	100,0

Tabla 4.7.14. Resultados en el ítem 7.a (Matemáticas y Máster de Secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. 1/ 2	4	4,2	3	3,0
b. 1/ 6	23	24,2	29	28,7
c. 1/ 3	62	65,3	64	63,3
d. 1/ 4	4	4,2	4	4,0
En blanco.	2	2,1	1	1,0
Total	95	100,0	101	100,0

Los resultados de los dos grupos de nuestra muestra, Tabla 4.7.14, indican una homogeneidad entre ambos grupos, ya que el porcentaje de respuestas se asemeja. Sólo destaca un 4,5% más de alumnos del Máster que se decantaron por el distractor (b), por lo que confunden la probabilidad condicional y la conjunta.

Resultados en el ítem 7b. Falacia del eje de tiempos

Ítem 7.b. ¿Cuál es la probabilidad de extraer una bola negra en primer lugar, sabiendo que hemos extraído una bola negra en segundo lugar? P (N1/ N2)
a. 1/ 3
b. 1/ 4
c. 1/ 6
d. 1/ 2

Este ítem pretendía conocer como los futuros profesores resuelven problemas de probabilidad condicional cuando se invierte el eje de tiempo en contexto de muestreo sin reposición, lo que solamente realizó correctamente un 29,6% de nuestra muestra y un 23,9% de los alumnos de la muestra de Díaz, Tabla 4.7.15. Los resultados confirman la hipótesis de Falk (1986), de que la dificultad de los problemas de probabilidad condicional crece cuando se invierte el eje temporal, pues formalmente, el ítem es similar a la primera parte (7a) donde la respuesta mayoritaria fue correcta en ambos grupos. La mejor preparación matemática en nuestro estudio no sirvió para superar este sesgo.

Tabla 4.7.15. Resultados en el ítem 7b

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. 1/3	58	29,6	99	23,9
b. 1/4	41	20,8	103	24,9
c. 1/6	25	12,8	38	9,2
d. 1/2	64	32,7	151	36,5
En blanco.	8	4,1	23	5,5
Total	196	100,0	414	100,0

Un 32,7% de los participantes en la muestra global y un 36,5% de la de Díaz no entendieron que se puede condicionar un suceso con otro que ocurra posteriormente, ya que eligieron el distractor (d) que evaluaba la falacia del eje temporal. Un 12,8% y un 9,2% de sujetos, respectivamente, se decantaron por el distractor (c) tratándose de alumnos que confunden entre probabilidad condicional y simple. Por último, un 20,8% y un 24,9%, respectivamente, eligieron la opción (b) confundiendo la probabilidad condicional y conjunta, error señalado por Falk (1986).

En la Tabla 4.7.16 se presentan los resultados de los dos grupos de nuestra muestra. En este caso el porcentaje de alumnos que respondieron correctamente fue mayor por parte de los alumnos del Máster en un 12,4% que el de los alumnos de la licenciatura. El porcentaje de alumnos que incurrió en la falacia del eje de tiempos, distractor (d), fue similar en ambos grupos con un 33,7% y un 31,7% respectivamente. En el resto de distractores, llama la atención que el distractor (b) fue considerado en un 8,5% más en el grupo de alumnos de la licenciatura, por el contrario el porcentaje de alumnos que se decantó por el distractor (c) fue similar en ambas muestras.

Tabla 4.7.16. Resultados en el ítem 7b (Matemáticas y Máster de Secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
a. 1/3	22	23,2	36	35,6
b. 1/4	24	25,3	17	16,8
c. 1/6	12	12,6	13	12,9
d. 1/2	32	33,7	32	31,7
En blanco.	5	5,2	3	3,0
Total	95	100,0	101	100,0

Síntesis de resultados en sesgos

Para poder realizar una síntesis sobre la presencia de sesgos en el razonamiento condicional de los futuros profesores de educación secundaria, en la Tabla 4.7.17 presentamos un resumen de la frecuencia y porcentaje con que se dieron los diferentes sesgos para cada ítem.

Como vemos, el porcentaje que incurre en las diferentes falacias o sesgos es alto, tanto en nuestra muestra como en la de Díaz (2007). Destaca por su incidencia, la falacia del eje de tiempos que se dio en un alto porcentaje, 68,3% y 78,5% respectivamente en el ítem 6 y en el 32,7% y 36,5% de respuestas en el ítem 7b. Estos sujetos tienen dificultad para resolver un problema de probabilidad condicional si el suceso condicionante ocurre después del condicionado. Sin embargo, esta es una situación muy común en la vida real (por ejemplo en la evaluación de los alumnos, el suceso condicionante seria el resultado del examen y el condicionado los conocimientos del estudiante). También aparece con frecuencia en inferencia estadística donde conceptos como intervalo de confianza o nivel de significación están definidos mediantes condicionales en que el eje de tiempos se ha invertido.

En menor medida, aunque el porcentaje es alto, tenemos la falacia de la tasa base, que en nuestra muestra se dio en un porcentaje del 39,8%, pero que en la muestra de Díaz sólo alcanzó el 15%. En estos casos se ignora una parte de los datos al resolver un problema de probabilidad condicional.

Tabla 4.7.17. Síntesis de sesgos

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Ítem 1 - Falacia tasas base	78	39,8	62	14,9
Ítem 2 - Confusión entre sucesos independientes y mutuamente excluyentes	63	32,1	118	28,5
Ítem 3 - Confusión entre probabilidad conjunta y condicional	62	31,6	135	32,6
Ítem 4 – Falacia de la conjunción	23	11,7	39	9,4
Ítem 5 – Falacia de la condicional transpuesta	27	13,8	24	5,8
Ítem 6 – Falacia del eje de tiempos	134	68,3	325	78,5
Ítem 7b- Falacia eje de tiempos	64	32,7	151	36,5

El siguiente sesgo por orden de frecuencia de aparición es la confusión entre probabilidad conjunta y condicional, con uno de cada tres encuestados aproximadamente, posiblemente por la dificultad de expresar condicionales en el lenguaje ordinario y con algo mayor frecuencia en los futuros profesores de secundaria. El resto de sesgos, falacia de la conjunción y de la condicional transpuesta, se dieron en un porcentaje bajo, alrededor del 10% en nuestra muestra y en menor porcentaje en la de Díaz.

Para comprobar si existe relación entre la frecuencia de aparición de los diferentes sesgos estudiados y el tipo de estudiante hemos realizado un test Chi-cuadrado de contraste de hipótesis para determinar si se rechaza o no la idea de independencia entre las diferentes muestras y sesgos, obteniendo Chi-cuadrado= 44,36; g.l.= 6; p= 0,00001. Por tanto, los resultados son estadísticamente muy significativos y se rechaza la hipótesis de independencia. Como conclusión, los diferentes sesgos encontrados en cada ítem tienen relación con la muestra de estudiantes, siendo mayor en la falacia de las tasas base, confusión entre sucesos independientes y mutuamente excluyentes, más para los futuros profesores de matemáticas que entre los estudiantes de psicología, mientras que la falacia del eje de tiempo es ligeramente mayor en el segundo grupo. El resto de sesgos tiene poca incidencia y es similar en ambos grupos.

En la Tabla 4.7.18 presentamos los resúmenes de los diferentes sesgos encontrados, junto con la frecuencia o porcentaje para cada ítem en los dos grupos de nuestra muestra. Como ocurría en la muestra global, el sesgo que más aparece es la falacia del eje de tiempos, con un 67,4% en los alumnos de la licenciatura y un 69,3% en la muestra de alumnos del Máster. La mayoría de los sesgos tienen resultados semejantes en ambos grupos, excepto en la falacia de la tasa base, ya que el porcentaje de alumnos del Máster que cometió errores es superior en un 15,9% a los alumnos de Matemáticas, y en la confusión entre sucesos independientes y mutuamente excluyentes donde el porcentaje de alumnos de Matemáticas es mayor en un 17,3% que en los alumnos del Máster.

Tabla 4.7.18. Síntesis de sesgos (Matemáticas y Máster de Secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Ítem 1 - Falacia tasas base	30	31,6	48	47,5
Ítem 2 - Confusión entre sucesos independientes y mutuamente excluyentes	39	41,1	24	23,8
Ítem 3 - Confusión entre probabilidad conjunta y condicional	25	26,3	37	36,6
Ítem 4 – Falacia de la conjunción	13	13,7	10	9,9
Ítem 5 – Falacia de la condicional transpuesta	14	14,7	13	12,9
Ítem 6 – Falacia del eje de tiempos	64	67,4	70	69,3
Ítem 7b- Falacia eje de tiempos	32	33,7	32	31,7

Al igual que en la comparación entre la muestra global y la de Díaz, hemos realizado un test Chi-cuadrado para determinar si se rechaza o no la idea de independencia entre los diferentes grupos de nuestra muestra y los diferentes sesgos. Los resultados del análisis son: Chi-cuadrado= 10,12; g.l.= 6; *p*-valor= 0,1198, por lo que los resultados no son estadísticamente significativos y no podemos rechazar la hipótesis de que los resultados de los sesgos de ambos grupos de nuestra muestra sean independientes. En consecuencia, no observamos diferencia en la presencia de los diferentes sesgos en los dos grupos que componen la muestra.

Tabla 4.7.19. Número de sesgos por alumno

	Lic. Matemáticas		Máster Secundaria	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Alumnos con 0 sesgos	3	3,2	10	9,9
Alumnos con 1 sesgo	19	20,0	16	15,8
Alumnos con 2 sesgos	36	37,9	28	27,7
Alumnos con 3 sesgos	24	25,3	31	30,7
Alumnos con 4 sesgos	11	11,5	11	10,9
Alumnos con 5 sesgos	2	2,1	5	5,0
Total	95	100,0	101	100,0

En la Tabla 4.7.19 presentamos la distribución del número de sesgos encontrados por alumno, comparando los resultados en la Licenciatura de Matemáticas y del Máster. Observamos que son minoría los estudiantes sin sesgos, especialmente en la Licenciatura de Matemáticas. El 96,8% de los alumnos de Matemáticas y el 90,1% del

Máster incurrió por lo menos en un sesgo y el 76,8% de Matemáticas y 74,3% del Máster en dos o más, lo que es causa de preocupación, pues estos docentes podrían transmitir dichos sesgos de razonamiento a sus futuros estudiantes. Los mayores porcentajes se dieron para alumnos con dos y tres sesgos, encontrando alumnos hasta con 5, aunque el porcentaje de ocurrencia fue pequeño.

En la Tabla 4.7.20 realizamos un análisis descriptivo del número de sesgos detectado por sujeto para ambos grupos de la muestra. Observamos similitud del número medio de sesgos por alumno del grupo, aunque la media y dispersión ha sido algo mayor en los estudiantes del Máster pero la diferencia no fue estadísticamente significativa.

Tabla 4.7.20. Análisis descriptivo de sesgos por alumno

	Lic. Matemáticas	Máster Secundaria
Media	2,28	2,32
Desv. Típica	1,078	1,288
Mínimo	0	0
Máximo	5	5

Tabla 4.7.21. Número de ítems correctos por alumno, ítems 1 al 7b

	Lic. Matemáticas		Máster Secundaria		Muestra Global	
	Frecuencia	Porcentaje	Frecuencia	Frecuencia	Frecuencia	Frecuencia
0	2	2,1	2	2,0	4	2,0
1	16	16,8	15	14,9	31	15,8
2	29	30,5	28	27,7	57	29,1
3	25	26,3	13	12,9	38	19,4
4	15	15,8	17	16,8	32	16,3
5	5	5,3	17	16,8	22	11,2
6	3	3,2	3	3,0	6	3,1
7	0	0	6	5,9	6	3,1
Total	95	100,0	101	100,0	196	100,0

En la Tabla 4.7.21 hemos analizado el número de respuestas correctas para los ítems de opción múltiple. El número de respuestas correctas ha fluctuado entre cero y seis, para la muestra de la licenciatura, y entre cero y siete, para la muestra de alumnos del Máster. Respecto al grupo de Matemáticas tenemos que la mayoría de alumnos, 91,8%, contestó entre uno y cinco ítems correctamente. En el 89,1% los alumnos del máster respondieron correctamente entre uno y cinco ítems. El porcentaje de alumnos que no contestó correctamente a ningún ítem fue sólo de 2%, similar en ambas muestras. Los valores descriptivos de ambas muestras, Tabla 4.7.22, indican que el valor medio de respuestas correctas fue inferior para el grupo de la licenciatura (con 2,65 de respuestas correctas de media) respecto al 3,19 de media de respuestas correctas de la muestra del Máster. Los resultados de la muestra de la licenciatura están más agrupados que la del Máster.

Tabla 4.7.22. Análisis descriptivo de ítems correctos por alumno, ítems 1 al 7b

	Lic. Matemáticas	Máster Secundaria	Muestra Global
Media	2,65	3,19	2,93
Desv. Típica	1,31	1,76	1,58
Mínimo	0	0	0
Máximo	6	7	7

Como resumen del estudio de sesgos, observamos que son pocos los futuros profesores que se encuentran libres de los mismos en las situaciones planteadas, muchas de las cuales aparecen en la vida cotidiana o profesional. Los resultados en algunos de los ítems han sido peores que los obtenidos por Díaz (2007) en estudiantes de Psicología, a pesar de la mayor preparación matemática de los alumnos de nuestra muestra. La presencia de sesgos es similar en licenciados de Matemáticas y en estudiantes del Máster de Secundaria, muchos de los cuales provienen de otras titulaciones. Será en consecuencia necesario que la preparación de los futuros profesores tenga en cuenta estos sesgos y se organicen actividades formativas que los ayuden a superarlos. En el Capítulo 6 describiremos una de estas situaciones mostrando cómo al confrontar a los futuros profesores con los sesgos, se puede mejorar su razonamiento.

4.7.2. DEFINICIÓN DE LA PROBABILIDAD SIMPLE Y CONDICIONAL

Además de los ítems de opción múltiple, el cuestionario incluye otros de respuesta abierta, que se analizan a continuación y que evalúan el conocimiento formal del tema. Para analizar los resultados, en primer lugar, se clasifican las respuestas según grado de completitud con una escala que varía de 0 a 2 o de 0 a 4, según el ítem, y que fue propuesta por Díaz (2007) en el análisis de sus resultados. El uso de estas escalas permitirá comparar los resultados con los de la citada autora. Por otro lado, cada una de las categorías citadas se desglosa en otras nuevas para analizar con mayor profundidad las dificultades y conflictos semióticos de los participantes en nuestra investigación.

En lo que sigue, se describen las diferentes categorías de respuesta a estos ítems, analizando resumidamente un ejemplo de cada una y presentando las frecuencias con que aparecen en el global de la muestra. Seguidamente, se comparan nuestros datos con los obtenidos en Díaz (2007) y finalmente se comparan los dos grupos de estudiantes que conforman nuestra muestra. Es importante resaltar que los estudiantes de Psicología respondieron al cuestionario después de haber estudiado el mismo curso el tema (aproximadamente un mes después), mientras que algunos futuros profesores lo habían estudiado varios años antes. Por ello, algunas respuestas son mejores en el primer grupo, a pesar de su menor formación matemática.

Ítem 8. Explica con tus propias palabras la diferencia entre una probabilidad simple y una probabilidad condicional.

El primero de los ítems abiertos estudia la capacidad para dar una definición adecuada del concepto. Díaz (2007) clasificó las respuestas de este ítem siguiendo el siguiente criterio: dio el valor 0 si la respuesta es totalmente incorrecta, 1 si se define correcta, pero imprecisamente la probabilidad simple o la condicional, 2 puntos si se define correcta, y precisamente una de las probabilidades pedidas, 3 si se define correcta, pero imprecisamente las dos probabilidades pedidas y 4 si se define correctamente las dos probabilidades pedidas. En nuestro caso hemos ampliado esta clasificación, estudiando las distintas respuestas, junto con los distintos errores que se cometen para cada categoría. Las categorías que hemos encontrado en este ítem son las siguientes:

C0. *Respuesta totalmente incorrecta*. Las causas de los errores son diversas, por lo que hemos clasificado las respuestas incorrectas en tres grupos. Un primer grupo incluye los errores formales en la definición (C01), un segundo grupo engloba las definiciones con impresiones debidas a confusión de conceptos que intervienen en la misma (C02) y un tercer grupo si los errores se producen por confusiones en la fórmula de la probabilidad condicional (C03). A continuación analizamos estas categorías:

C0.1. *Errores en la definición*. Cuando el sujeto incluye el término que desea definir en la definición o bien da una definición que corresponde a otro concepto. Por ejemplo, el sujeto 123 da una definición circular y mezcla los términos suceso con "algo" y "sentencia":

> *"Como su nombre indica la probabilidad condicionada viene condicionada por otra sentencia. Al variar algo varia otro"* (Sujeto 123).

C0.2. *Errores producidos por confundir otro concepto que interviene en la definición.* En la siguiente respuesta, el sujeto 3 confunde frecuencia relativa con sucesos, pues la probabilidad seria el límite de la frecuencia relativa, no del suceso, que no tiene límite. La definición es incorrecta pues se refiere a sucesos y no a sus probabilidades:

> *"Probabilidad simple: valor límite al que tiende un suceso aleatorio, no condicionado por nada. Probabilidad condicionada: valor límite al que tiende un suceso dependiente de otro suceso, cuando se ha dado el otro suceso"* (Sujeto 3).

C0.3. *Errores por confusión en la fórmula de la probabilidad condicional*. Algunos sujetos incluyen la fórmula en la definición de la probabilidad condicional, pero confunden la fórmula de la probabilidad condicional, como el siguiente caso que invierte la fórmula, además de cambiar el suceso condicionante y de sustituir la intersección por la unión de sucesos. Por otro lado la definición es circular:

> *"Probabilidad simple: es cuando intervienen los elementos sin estar condicionados a algo. Probabilidad condicional: cuando los elementos esta condicionados a otros, cuando depende de otros.* $P(A/B)=\frac{P(A)}{P(A\cup B)}$*"* (Sujeto 162).

C1. *Define correcta, pero imprecisamente la probabilidad condicional.* Díaz (2007) asignó un punto en este ítem cuando la definición no tiene errores importantes pero es confusa o imprecisa. Como en el caso anterior, nosotros hemos desglosado esta categoría. A continuación se analizan los casos encontrados:

C1.1. *Olvida definir la probabilidad simple y define imprecisamente la probabilidad condicional.* El siguiente ejemplo define la probabilidad condicional a partir de la ocurrencia de un suceso, pero la definición es imprecisa. Introduce la palabra "probabilidad" (que se trata de definir) en la definición; además se refiere a un suceso futuro (mientras que la probabilidad también se puede aplicar a un suceso en el pasado). Subyace la falacia del eje de tiempos (Tversky y Kahneman, 1982a; Gras y Totohasina, 1995):

> *"Probabilidad condicional: ocurrido un suceso la probabilidad de ocurrir los venideros sabiendo el suceso ocurrido, varía"* (Sujeto 134).

C1.2. *Suponer un suceso dependiente del otro, aunque esto no se exija para definir la probabilidad condicional.* Podría subyacer en estos casos una confusión entre condicionamiento y causalidad que ha sido descrita, entre otros, por Falk (1986). Por ejemplo, el sujeto 54 define la probabilidad condicional como una probabilidad que puede depender de otro suceso:

> *"Una probabilidad condicional es una probabilidad que puede depender de otro suceso. Por ejemplo la probabilidad de la segunda extracción de una bola (ejercicio anterior) depende de la primera extracción. Así la probabilidad de que la segunda sea negra si la primera es blanca es 2/3, sin embargo si la primera es negra sería de 1/3"* (Sujeto 54).

C2. *Define correcta, y precisamente una de las probabilidades pedidas,* como el siguiente ejemplo, que sólo define la probabilidad condicional: *"Cuando tratamos de calcular la probabilidad de ocurrencia de un suceso, conociendo una información adicional, hablamos de probabilidad condicionada"* (Sujeto 1). Otro ejemplo define la probabilidad condicional a partir de información añadida: *"La probabilidad condicional, es una probabilidad en la que se tiene información añadida que puede influir en la probabilidad"* (Sujeto 137). Hemos clasificado estas respuestas en tres grupos: un primer grupo si se trata de errores por confusión de conceptos (C2.1), un segundo grupo si se utilizan fórmulas para definir una de las probabilidades (C2.2) y un tercer grupo si se trata de errores por utilizar la propiedad de independencia para definir la probabilidad simple (C2.3).

C2.1. *Errores en una de las dos definiciones por confundir conceptos,* por ejemplo, definir la probabilidad condicionada en la forma siguiente, donde en realidad está dando la definición de probabilidad simple:

> *"Relación entre casos favorables y posibles pero teniendo en cuenta que ha ocurrido algo que altera la probabilidad. De ahí el término condicionada"* (Sujeto 122).

C2.2. *Utilización de fórmulas para definir una de las probabilidades pedidas*. Una respuesta de este tipo sería utilizar la descomposición de la probabilidad condicional como cociente de la probabilidad conjunta y de la condición, dando una definición correcta, pero olvida dar la definición de probabilidad simple, que dice no conocer.

> *"No conozco el termino de probabilidad simple, si la de suceso simple que se refiere a un tipo particular de suceso formado por un sólo elemento. La probabilidad condicionada de un suceso A*

condicionado a otro B siendo A y B dos sucesos de un mismo experimento aleatorio con P(B) distinto de cero es la probabilidad de que ocurra A sabiendo que ha ocurrido B. Se formula: $P(A/B)=\frac{P(A\cap B)}{P(B)}$ *”* (Sujeto 6).

C2.3. *Utilización de la propiedad de independencia para definir la probabilidad simple*. Cuando utilizan como elemento intrínseco la independencia para definir la probabilidad simple, mostrando una confusión entre independencia y probabilidad simple:

“Sean A y B dos sucesos posibles donde A y B son independientes si $P(A\cap B)=P(A)P(B)$*, entonces la probabilidad respecto A y B es simple”* (Sujeto 116).

C3. *Define correcta, pero imprecisamente las dos probabilidades pedidas*. Hemos clasificado las respuestas en seis grupos: Un primer grupo si suponen que el suceso condicionante tiene que ocurrir antes que el condicionado (C3.1), un segundo grupo englobaría los futuros profesores que suponen que el suceso condicionante es dependiente (C3.2), un tercer grupo si usan fórmulas para definir una de las probabilidades (C3.3), un cuarto grupo si se indica la diferencia entre ambas probabilidades, pero de forma imprecisa (C3.4), un quinto grupo si en lugar de dar una definición, da un ejemplo (C3.5) y un sexto grupo si confunde la probabilidad simple o condicional entre sí o con la conjunta (C3.6).

C3.1. *Suponer que el suceso condicionante tiene que ocurrir antes que el condicionado.* Se observa en la respuesta la falacia del eje de tiempos (Tversky y Kahneman, 1982a) o la concepción cronológica de la probabilidad condicional descrita por Gras y Totohasina (1995). En el siguiente ejemplo, se definen las dos probabilidades pero la respuesta es imprecisa, ya que no se tiene que dar un orden para poder condicionar dos sucesos:

“La probabilidad simple es la probabilidad de un suceso aislado que estemos estudiando sin depender de otros suceso, mientras que la probabilidad condicional es la probabilidad de un suceso condicionado a que ocurra previamente otro suceso diferente al anterior” (Sujeto 7).

C3.2. *Suponer que el suceso condicionante es dependiente del suceso condicionado*, aunque como hemos indicado esto no es necesario y podría subyacer en estos casos una confusión entre condicionamiento y causalidad (Falk, 1986). El sujeto 8 identifica la probabilidad condicional solamente para sucesos independientes:

> *"La probabilidad simple está basada en sucesos únicos o independientes, mientras que la condicionada utiliza sucesos que se condiciona unos a otros, por lo que para el cálculo de sus probabilidades será necesario tener en cuenta las de los sucesos dependientes"* (Sujeto 8).

C3.3. *Usar fórmulas no demasiado precisas para definir una de las probabilidades*. Por ejemplo, el sujeto 11 utiliza la definición de la regla de Laplace para la probabilidad simple como cociente entre casos favorables y posibles, pero la notación para la probabilidad es imprecisa:

> "$Simple = \frac{Sucesos\ favorables}{Sucesos\ posibles}$" (Sujeto 11).

C3.4. Incluimos en esta categoría las respuestas que *indican la diferencia* entre probabilidad simple y compuesta*, pero son imprecisas*. En la respuesta que reproducimos a continuación, se define la diferencia haciendo alusión a la idea de dependencia, que como hemos visto no es necesaria. Este razonamiento es impreciso porque el suceso condicionante puede ser dependiente del condicionado:

> *"La probabilidad condicional se diferencia de la simple en que está afectada por un suceso que no es independiente"* (Sujeto 12).

C3.5. *En lugar de dar una definición, da un ejemplo*. Por ejemplo, el sujeto 29 utiliza una baraja de cartas para describir la diferencia entre las dos probabilidades, en lugar de utilizar sus definiciones o una descripción de tipo general:

> *"La probabilidad simple, es por ejemplo, la probabilidad de sacar un oros de una baraja española. Una probabilidad condicionada, es por ejemplo la probabilidad de sacar un oro de una baraja española condicionada a otro suceso, que la primera carta por ejemplo fue copas"* (Sujeto 29).

C3.6. *Confunde una probabilidad con otra*, por ejemplo, confunde la probabilidad simple y conjunta. Pollatsek, Well, Konold y Hardiman (1987) afirman que muchas de las dificultades que las personas tienen con la comprensión de la probabilidad pueden deberse a la redacción de los enunciados. Su hipótesis se basa en los resultados de Einhorn y Hogarth (1986). Estos autores sugieren que los enunciados que usan la conjunción "*y*" pueden llevar a confundir la probabilidad conjunta y la probabilidad condicional. El siguiente profesor en formación, en lugar de dar un ejemplo de probabilidad simple, da un ejemplo de probabilidad conjunta:

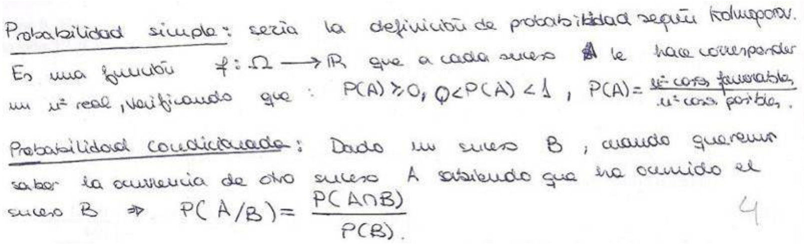
"Probabilidad simple trata dos sucesos que no guardan relación directa, por ejemplo la probabilidad que llueva en Santiago y luzca el sol en Buenos Aires. Probabilidad condicional se refiere a que los sucesos están relacionados entre sí. Por ejemplo como ocurre en el ítem 7 donde la probabilidad del segundo suceso incrementa o disminuye dependiendo del primero" (Sujeto 33).

C4. *Define correcta y de manera precisa las dos probabilidades pedidas*. Por último, algunos futuros profesores dan las dos definiciones de forma correcta o explican correctamente la diferencia entre las dos probabilidades. Hemos desglosado las respuestas en dos grupos: Un primer grupo si se trata de participantes que definen ambas probabilidades de forma correcta sin aplicar fórmulas (C4.1) y un segundo grupo si usan fórmulas para definir ambas probabilidades (C4.2).

C4.1. *Definir ambas probabilidades de forma correcta sin aplicar fórmulas*. Por ejemplo el sujeto 64 define las dos probababilidades de la siguiente manera:

"La probabilidad condicionada mide el valor de la probabilidad de un suceso A condicionada a la ocurrencia de otro suceso B. Mientras que la probabilidad simple sólo da el valor de la probabilidad de ocurrencia de un suceso independiente de la ocurrencia o no de otro suceso" (Sujeto 64).

C4.2.*Usar fórmulas para definir ambas probabilidades*. En el siguiente ejemplo (sujeto 22) se define la probabilidad simple y condicionada utilizando nomenclatura matemática. Además se hace referencia a que la probabilidad es una función de conjunto y a los axiomas de Kolmogorov.


Probabilidad simple: sería la definición de probabilidad según Kolmogorov.
Es una función $f: \Omega \longrightarrow \mathbb{R}$ que a cada suceso A le hace corresponder un nº real, verificando que: $P(A) \geq 0$, $0 < P(A) < 1$, $P(A) = \frac{\text{nº casos favorables}}{\text{nº casos posibles}}$.
Probabilidad condicionada: Dado un suceso B, cuando queremos saber la ocurrencia de otro suceso A sabiendo que ha ocurrido el suceso B $\Rightarrow P(A/B) = \frac{P(A \cap B)}{P(B)}$.
4

(Sujeto 22).

En la Tabla 4.7.23 presentamos las frecuencias y porcentajes de las diferentes categorías estudiadas anteriormente. La categoría más frecuente es la C3.2 que consiste en definir las dos probabilidades, siendo impreciso en la definición de la probabilidad condicional al suponer que los sucesos han de ser dependientes. En estas respuestas pudiera presentarse una confusión entre condicionamiento y causalidad comentada anteriormente. Si sumamos las categorías C0.1, C1.2, C2.3 y C3.2, todas las cuáles implican pensar que los sucesos han de estar condicionados, un total de 64 futuros profesores presentarían esta creencia (32,7%).

Tabla 4.7.23. Frecuencia y porcentaje de respuestas en la definición

	Frecuencia	Porcentaje
C0.1	11	5,6
C0.2	3	1,5
C0.3	2	1
C1.1	5	2,6
C1.2	1	0,5
C2.1	4	2
C2.2	4	2
C2.3	3	1,5
C2.4	1	0,5
C3.1	34	17,3
C3.2	49	25,1
C3.3	9	4,6
C3.4	10	5,1
C3.5	9	4,6
C3.5	3	1,5
C4.1	23	11,8
C4.2	8	4,1
En blanco	17	8,7

La segunda categoría más frecuente es la C3.1 (17,3%) consiste en los participantes que suponen que el suceso condicionante tiene que ocurrir antes que el condicionado, o falacia del eje de tiempo. En estos dos tipos de respuestas se translucen algunos de los sesgos sobre el razonamiento en probabilidad condicional que ya se detectaron en los ítems de opciones múltiples. La tercera categoría más frecuente ha sido la C4.1, definir ambas probabilidades de forma correcta sin aplicar fórmulas, con un 11,8%. Esta categoría, junto con la C4.2, implica que sólo un 15,9% de los participantes dio una respuesta correcta a las dos probabilidades. Es una proporción muy pequeña tratándose de futuros profesores que tendrán que impartir este tema si se incorporan a la función docente.

En la Tabla 4.7.24 se comparan las respuestas dadas por los sujetos de nuestra muestra, junto con los del estudio de Díaz (2007) con alumnos de Psicología. En ella podemos observar grandes diferencias entre los dos grupos. Por ejemplo, la categoría que más aparece en el grupo de futuros profesores ha sido la tercera, definir correctamente pero de forma imprecisa las dos probabilidades, con un 58,1% de respuestas en esa categoría. En contra, la categoría que más aparece en la muestra de Psicología ha sido la cuarta con un 30,7%, definir correctamente las dos probabilidades, algo que en nuestra muestra (15,8%). En cambio, el segundo porcentaje mayor de la muestra de Psicología se da en las respuestas en blanco o totalmente incorrectas, con un 28,7% de alumnos, mientras que ésta sólo se dio en un 16,9% de la muestra de nuestro estudio.

Tabla 4.7.24. Resultados en el ítem 8

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	33	16,9	119	28,7
C1	6	3,2	28	6,8
C2	12	6,0	90	21,7
C3	114	58,1	50	12,1
C4	31	15,8	127	30,7
Total	196	100,0	414	100,0

Tabla 4.7.25. Resultados en el ítem 8 (Matemáticas y Máster de secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	10	10,5	23	22,7
C1	4	4,2	2	2,0
C2	4	4,2	8	8,0
C3	58	61,1	56	55,5
C4	19	20,0	12	11,8
Total	95	100,0	101	100,0

En la Tabla 4.7.25 hemos dividido los resultados de los dos grupos de nuestra muestra, alumnos de la licenciatura y del máster, para ver si se presentan o no diferencias entre ellos. Observamos que el porcentaje de alumnos que contestan correctamente a las dos probabilidades es un 8,2% mayor en los alumnos de la Licenciatura. En cambio un 61,1% de estos alumnos definió las dos probabilidades de forma imprecisa comparado con el 55,5% de los alumnos del Máster. La respuesta más llamativa es que sólo un 10,5% de los alumnos de la licenciatura falló en la respuesta; este porcentaje crece al 22,7% en el caso de los alumnos del Máster en educación.

En general, se ve mejor conocimiento de las definiciones en los alumnos de la Licenciatura, aunque en los dos casos es deficiente pues son pocos los que dan las dos definiciones completas y correctas o explican de forma correcta la diferencia entre los dos conceptos.

4.7.3. CÁLCULO DE PROBABILIDAD CONDICIONADA

Ítem 9. Hemos lanzado dos dados y sabemos que el producto de los dos números obtenidos ha sido 12 ¿Cuál es la probabilidad de que ninguno de los dos números sea un 6?

Junto con la definición se incluyen en el cuestionario una serie de problemas. Este ítem pretende evaluar el cálculo de una probabilidad condicional en una situación sincrónica y sin reemplazamiento por parte de los alumnos. Siguiendo el mismo criterio de Díaz (2007) y para poder comparar con los resultados de esta autora se ha puntuado siguiendo el siguiente criterio: 0 puntos si no se llega a identificar los datos, 1 punto si el alumno identifica los casos favorables o posibles, pero no resuelve el problema y 2 si resuelve el problema correctamente por la regla de Laplace o por la fórmula del producto. Para ampliar la clasificación de respuestas de Díaz, dentro de cada categoría hemos estudiado las distintas respuestas, junto con los distintos errores que cometen los alumnos. Las categorías que hemos encontrado en este ítem son las siguientes:

C0. *No llega a identificar los datos.* Hemos clasificado las respuestas incorrectas en varios grupos dependiendo del error cometido. Un primer grupo, si se trata de errores en la identificación de casos favorables y posibles (C0.1), un segundo grupo incluye errores al identificar incorrectamente el espacio muestral de lanzar dos dados (C0.2), un tercero, si se trata de errores por no tener en cuenta el orden al identificar los casos favorables y posibles (C0.3) y un cuarto consiste en calcular la probabilidad conjunta y confundir la fórmula de la probabilidad conjunta aplicando la regla de la suma, en vez del producto (C0.4). A continuación los describimos con detalle incluyendo un ejemplo de cada uno:

C0.1 *Error en la identificación de casos favorables y posibles, lo que llevara a un error en el cálculo de la probabilidad.* En la respuesta que reproducimos, el sujeto 46 toma como posibles el conjunto de todas las variantes de las tiradas en vez de las que multiplican 12. Es decir, falla en hacer la restricción del espacio muestral, que es necesaria, según Totohasina (1992) para calcular la probabilidad condicional. Muestra también una confusión entre probabilidad simple y condicional pues calcula una probabilidad simple en lugar de la condicional:

$\Omega = \{(1,1),(1,2),(1,3), - - - -, (6,6)\}$.

$|\Omega| = 36$.

Las dos posibilidades son $3 \times 4 \wedge 4 \times 3$.

Por tanto la probabilidad es $\frac{2}{36} = \frac{1}{18}$ (Sujeto 46).

C0.2. *Calcular incorrectamente el espacio muestral*, debido a un error al considerar el orden de los dados, y además, fallo en la restricción del espacio muestral. En el siguiente caso, no se tiene en cuenta que el lanzamiento de dos dados son experimentos independientes, por lo que hay que distinguir el orden:

> *"Hallamos el espacio muestra: {(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,2), (2,3), (2,4), (2,5), (2,6), (3,3), (3,4), (3,5), (3,6), (4,4), (4,5), (4,6), (5,5), (5,6), (6,6)}"* (Sujeto 110).

C0.3. *No tener en cuenta el orden al identificar los casos favorables y posibles, aunque restringe el espacio muestral*. En estos casos se resuelve el problema, pero los hemos considerado incorrectos por no considerar el orden. Por ejemplo, el sujeto 11 toma como espacio muestral *E:*{(2,6), (3,4)} para calcular los casos posibles, sin tener en cuenta que puede variar el orden:

> *"Los casos posibles son: 2*6=12, 3*4=12. P(ninguno de los dos es 6)=1/2"* (Sujeto 11).

C0.4. *Error al identificar la probabilidad pedida.* Trata de calcular la probabilidad conjunta y confunde la fórmula aplicando la regla de la suma, en lugar del producto:

> *"Entendiendo que los dados son cubos, o sea que están numerados del 1 al 6 la probabilidad será cero ya que sólo estará la solución 6+6=12"* (Sujeto 126).

C1. *Se identifican los casos favorables o posibles, pero no se resuelve el problema*. Hemos considerado los siguientes casos: Un primer grupo (C1.1) si se plantea el problema pero no se llega a dar una solución, un segundo grupo por cometer errores en la fórmula de la probabilidad condicional (C1.2), un tercero por calcular la probabilidad conjunta de obtener un número diferente de 6 en los dos dados (C1.3), un cuarto por calcular la probabilidad conjunta en vez de la condicionada (C1.4) y un quinto por calcular la probabilidad conjunta de sacar un 3 o un 4, excluyendo del espacio muestral el valor 6 (C1.5); o bien toman el espacio muestral del experimento compuesto (C1.6).

C1.1. *Plantean el problema pero no llegan a dar una solución.* Una respuesta de este tipo la da el sujeto 31 identifica solamente los casos favorables sin llegar a calcular probabilidad alguna.

*"1, 2, 3, 4, 5, 6, 4*3=12, 3*4=12, 6*2=12, 2*6=12"* (Sujeto 31).

C1.2. *Error en la fórmula de la probabilidad condicional*. Se identifica correctamente los datos y la probabilidad pedida, pero se comete un error en la fórmula. En la siguiente solución, se multiplica el numerador de la fórmula (probabilidad compuesta de que no salga un seis y sume 12) por la probabilidad simple de que no salga ningún seis en las dos tiradas:

" $P(No\ 6 \mid \Pr = 12) = \dfrac{\frac{2}{36} \cdot \frac{25}{36}}{\frac{4}{36}} = 0{,}347$ " (Sujeto 43).

C1.3. *Calcular la probabilidad conjunta de obtener un número diferente de 6 en los dos dados*, como en la siguiente respuesta (sujeto 60), donde se ha interpretado incorrectamente el enunciado, aunque la probabilidad conjunta estaría bien calculada:

$P(\text{ninguno sea } 6) = 1 - P(\text{los dos sean } 6) = 1 - \left(\frac{1}{6} \cdot \frac{1}{6}\right) = 1 - \frac{1}{36}$ (Sujeto 60).

C1.4. *Calcular la probabilidad conjunta en lugar de la condicionada.* Estos alumnos pudieran estar confundiendo estas dos probabilidades, dificultad señalada por Pollatsek, Well, Konold y Hardiman (1987). En el siguiente ejemplo, se calcula la probabilidad de obtener una de las dos parejas (3,4) o (4,3), que cumplen la condición de que la suma es 12 y ninguno de los números es 6; es decir es una probabilidad conjunta (suceso compuesto):

$12 = 4 \cdot 3 = 3 \cdot 4$

$P((N_1 \neq 6) \cap (N_2 \neq 6)) = P(N_1 = 4) \cdot P(N_2 = 3) +$

$P(N_1 = 3) \cdot P(N_2 = 4) = \frac{1}{6} \cdot \frac{1}{6} + \frac{1}{6} \cdot \frac{1}{6} = \frac{2}{36} = \frac{1}{18}$

La probabilidad de que ninguno de los dos números sea un 6 es 1/18. (Sujeto 44).

C1.5. *Calcular la probabilidad conjunta de sacar un 3 o un 4, excluyendo del espacio muestral el valor 6.* Así, el sujeto 73 toma como espacio muestral los valores comprendidos entre 1 y 5 y calcula la probabilidad de obtener un 3 y un 4:

"Sólo hay dos posibilidades de que el producto sea 12 (4x3 y 6x2).Si no puede aparecer un 6

habría que calcular la probabilidad de sacar un 4 y un 3 pero no sobre 6 opciones sino sobre 5 porque no podemos contar con el número 6: $P(4\ y\ 3)=\frac{1}{5}\cdot\frac{1}{5}=\frac{1}{25}$ ” (Sujeto 73).

C1.6. *Se identifican los casos favorables y posibles, se plantea el problema pero se toma el espacio muestral del experimento compuesto*. En la siguiente solución, el alumno 147 identifica correctamente los datos y tiene en cuenta la restricción del espacio muestral necesaria en el cálculo de la probabilidad condicional, así como los elementos que la componen. Pero a la hora de calcular la probabilidad pedida, toma como casos posibles 36 en lugar de 4, es decir, toma los casos posibles en el experimento compuesto:

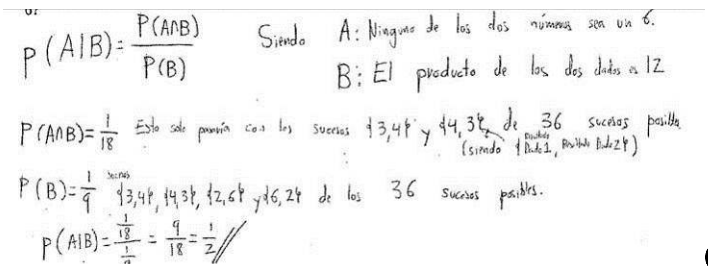
“$12=2\cdot 6=4\cdot 3=3\cdot 4$*; como* $P(12)=\frac{4}{36}$ *entonces* $\frac{2}{6}=P(12, \mathit{Ningún}\ 6)$ ” (Sujeto 147).

C2. *Resuelve el problema correctamente*, después de haber identificado tanto los datos, como la probabilidad pedida, calcula la probabilidad aplicando correctamente la fórmula, un diagrama en árbol u otro recurso. Hemos clasificado las respuestas correctas en dos grupos dependiendo de la forma de resolverlo correctamente. Un primer grupo si resuelven correctamente el problema utilizando para ello la regla de Laplace (C2.1) y un segundo grupo si resuelven correctamente el problema utilizando la fórmula de la probabilidad condicional (C2.2).

C2.1. *Resolver correctamente utilizando la regla de Laplace*. Por ejemplo, el sujeto 37 calcula el espacio muestral identificando los sucesos que intervienen. Calcula los casos favorables y posibles que intervienen y aplica la regla de Laplace:

“*E={(1,6) (6,1) (3,2) (2,3)}* $P(A)=\frac{C.\ favorables}{C.\ posibles}=\frac{2}{4}=0,5$ ” (Sujeto 37).

C2.2. *Resolver correctamente utilizando la fórmula de la probabilidad condicional.* En el siguiente caso, se utiliza la descomposición de la probabilidad condicional en función de la probabilidad conjunta y de la probabilidad del suceso condicionante. Se identifica correctamente los datos del problema, introduce una notación adecuada para los diversos sucesos que interviene, usa explícitamente la notación de probabilidad condicional, e intersección de sucesos y aplica correctamente la fórmula:


$P(A|B) = \frac{P(A\cap B)}{P(B)}$ Siendo A: Ninguno de los dos números sea un 6.
B: El producto de los dos dados es 12
$P(A\cap B) = \frac{1}{18}$ Esto solo pasaría con los sucesos {3,4} y {4,3} de 36 sucesos posibles (siendo {Dado1, Dado2})
$P(B) = \frac{1}{9}$ sucesos {3,4}, {4,3}, {2,6} y {6,2} de los 36 sucesos posibles.
$P(A|B) = \frac{\frac{1}{18}}{\frac{1}{9}} = \frac{9}{18} = \frac{1}{2}$

(Sujeto 41).

En la Tabla 4.7.26 se exponen las frecuencias y porcentajes, de las distintas categorías del ítem 9. La categoría que más aparece para la muestra de futuros profesores es la C2.1, resuelve el problema correctamente mediante la regla de Laplace, con un 53,1%. La siguiente categoría con mayor porcentaje de aparición ha sido el error definido como C0.1, identificación incorrecta de los casos favorables o posibles, con un 16,8% de respuestas. La tercera categoría que más aparece es la C1.3, calcular la probabilidad conjunta de obtener un número diferente de 6 en los dos dados, con 8,2%. El resto de categorías han tenido un porcentaje bajo de respuestas.

Tabla 4.7.26. Frecuencia y porcentaje de respuestas en calculo probabilidad condicionada

	Frecuencia	Porcentaje
C0.1	33	16,8
C0.2	7	3,6
C0.3	16	8,2
C0.4	3	1,5
C1.1	3	1,5
C1.2	3	1,5
C1.3	4	2,0
C1.4	2	1,0
C1.5	5	2,6
C1.6	3	1,5
C2.1	104	53,1
C2.2	8	4,1
En blanco	5	2,6

En la Tabla 4.7.27 se comparan los resultados dados por los futuros profesores y los alumnos de la muestra de Díaz. Observamos que la categoría con mayor frecuencia de nuestra muestra ha sido la segunda, resuelve el problema correctamente, con un 57,1%. Si comparamos los resultados de esta categoría con la muestra de Psicología, el porcentaje de alumnos que respondió correctamente fue del 34,3% de alumnos. Por lo contrario la categoría que más aparece en la muestra de Psicología ha sido la cero, con un 44,9%, es decir se trata de alumnos que resuelven incorrectamente o dejan la pregunta en blanco. El porcentaje de los que contestan incorrectamente en la muestra de futuros profesores ha sido del 32,7%, un 12,2% menos que la de Díaz. La categoría C1 se presentan en un 10,2% en nuestra muestra y en un 20,8% en la muestra de Díaz.

Tabla 4.7.27. Resultados en el ítem 9

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	64	32,7	186	44,9
C1	20	10,2	86	20,8
C2	112	57,1	142	34,3
Total	196	100,0	414	100,0

En la Tabla 4.7.28 se muestran los resultados de los dos grupos de nuestra muestra, alumnos de la licenciatura y del máster de secundaria. En ella observamos que el porcentaje de alumnos que contestan correctamente es un 21% mayor en los alumnos del Máster. También se muestra que el porcentaje de alumnos que cometieron errores en la resolución del problema fue 22,4% mayor en la muestra de los estudiantes de Matemáticas. El resto de categorías se dan de forma aproximada en ambas muestras.

Tabla 4.7.28. Resultados en el ítem 9 (Matemáticas y Máster de secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	42	44,2	22	21,8
C1	9	9,5	11	10,9
C2	44	46,3	68	67,3
Total	95	100,0	101	100,0

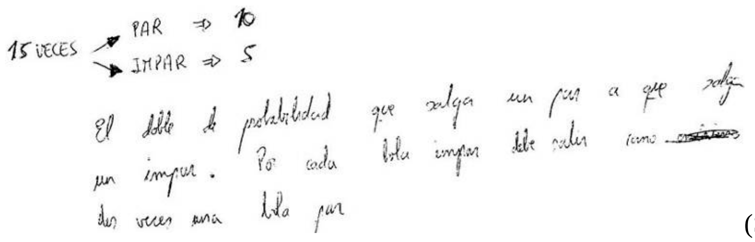
4.7.4. INDEPENDENCIA

Ítem 11. Una persona lanza un dado y anota si saca un número par o impar. Estos son los resultados:
par, impar, impar, par, par, impar, par, par, par, par, impar, impar, par, par, par.
Lanza el dado de nuevo. ¿Cuál es la probabilidad de sacar un número par en la siguiente tirada?

Este ítem se incluyó en el cuestionario para analizar si los futuros profesores tienen problemas en la comprensión de la idea de independencia. Al igual que en los otros apartados, dado que las causas de los errores son diversas, hemos clasificado las respuestas incorrectas en varios grupos dependiendo del error cometido y tratando de seguir la clasificación de Díaz (2007), para poder comparar con sus resultados, aunque desarrollando algo más la clasificación de la autora.

C0. *No se responde o responde incorrectamente.* Se han considerado varias categorías: Un primer grupo si se trata de errores por calcular la probabilidad de ocurrencia en función de los resultados anteriores, cometiendo la falacia del jugador (C0.1), donde se espera que una serie corta de ensayos se equilibre (Kelly y Zwiers, 1986; Sánchez, 1996); un segundo grupo (C0.2) trata de aplicar la fórmula de la probabilidad condicional, pero comete un error en la fórmula, un tercer grupo de respuestas (C0.3) donde se hace una interpretación incorrecta del enunciado, y el último por cometer errores de cálculo o no responder (C0.4).

C0.1. *Calcular la probabilidad de ocurrencia en función de los resultados anteriores (falacia del jugador).* Estos futuros profesores esperan que una serie corta de ensayos se equilibre y no comprenden la independencia de ensayos (Kelly y Zwiers, 1986; Sánchez, 1996). Por ejemplo, el sujeto 31 expone que habrá el doble de probabilidad de que salga par que impar ya que en las tiradas anteriores ha salido dos pares por cada impar. Manifiesta la heurística de la representatividad (Kahneman, Slovic y Tversky, 1982), esperando un cambio de resultado en una racha corta.


15 VECES → PAR ⇒ 10
→ IMPAR ⇒ 5
El doble de probabilidad que salga un par a que salga un impar. Por cada bola impar debe salir como ~~...~~ dos veces una bola par

(Sujeto 31).

Algunos futuros profesores, aunque identifican correctamente la probabilidad como 1/2 siguen pensando que depende de los resultados anteriores, como el alumno 172:

> *"La probabilidad sería la misma que todo el tiempo 3/6=1/2, sin embargo teniendo en cuenta el azar, es muy probable que sea par ya que de quince veces, 10 veces ha salido par y sólo cinco impar"* (Sujeto 172).

C0.2. *Trata de aplicar la fórmula de la probabilidad condicional, pero comete un error en la fórmula.* Así, el sujeto 80 calcula la probabilidad de que la siguiente tirada sea par como producto de la probabilidad de salir par por la probabilidad de los sucesos que ocurrieron en las tiradas anteriores. La fórmula es claramente incorrecta.

10 par —
5 impar

Se tira el dado 16 veces → probabilidad de que la última sea par sabiendo las 15 anteriores:

$P(\text{par} \mid 15 \text{ tiradas}) = \frac{1}{2} \cdot \left(\frac{10}{15} + \frac{5}{15}\right) = \frac{1}{2}\left(\frac{50}{225}\right) = \frac{25}{225} = \frac{1}{9}$ (Sujeto 80).

C0.3. *Interpretación incorrecta del enunciado*. Algunos futuros profesores no son capaces de identificar la probabilidad pedida y hacen una interpretación alternativa. En el siguiente ejemplo, se calcula la probabilidad sumando la probabilidad de sacar un número del dado más la probabilidad de sacar un número par:

$\frac{1}{6} + \frac{3}{6} = \frac{4}{6} = 6'66 \rightsquigarrow 66'6\%$

ya que tenemos tres nº pares en un dado y solo podemos sacar un resultado. (Sujeto70).

C0.4. *Cometer errores de cálculo*. Son errores de menor importancia consistentes en plantear correctamente el problema pero fallar en el cálculo. En la siguiente solución, el sujeto 97 define la probabilidad simple de salir par como 1/3 por un error elemental en la simplificación de fracciones.

$\frac{\text{Nº casos favorables}}{\text{Nº casos totales}} = \frac{3}{6} = \frac{1}{3}$

la probabilidad de sacar un nº par es 1/3 siempre que se lance un dado.

(lanzar un dado = {par, impar})
$p = \frac{1}{3}$ $p = 1/3$

Cada resultado es un suceso independiente. (Sujeto 97).

C1. *Se da una estimación frecuencial de la probabilidad*. Díaz (2007) puntuó con 1 punto este tipo de respuesta, porque el enunciado no indica si el dado está o no sesgado. Por tanto, el futuro profesor podría interpretar que se trata de un dado sesgado por la diferencia entre pares e impares en las tiradas anteriores. No tomamos como totalmente correcta esta respuesta porque los resultados obtenidos, podrían razonablemente aparecer en un dado correcto. Esta respuesta denota la heurística de la representatividad (Kahneman, Slovic y Tversky, 1982), esperando un cambio de resultado en una racha corta. Un ejemplo se muestra a continuación:

"P(P)=10/15=2/3; P(I)5/15=1/3; entonces puesto que el dado está trucado sería de 2/3" (Sujeto 20).

También tenemos casos que piensan que son muy pocos los lanzamientos como para poder aplicar la estimación frecuencial:

> *"Uno podría pensar, ya que han salido 10 pares en la muestra que volvería a salir par pues ha salido más pares que impares. Pero al ser un número de lanzamientos tan bajo, esto no nos aseguraría que el dado puede estar manipulado para que saliera más pares"* (Sujeto 57).

C2. *Calcular correctamente la probabilidad, dando el valor 1/2.* Díaz (2007) puntuó con 2 puntos esta respuesta donde el estudiante muestra una clara idea y aplicación a este caso de la independencia. Hemos clasificado está categoría en dos grupos según hayan utilizado para la resolución la regla de Laplace (C2.1) o la fórmula de la probabilidad condicional (C2.2).

C2.1. *Utilizando la regla de Laplace*. El futuro profesor reconoce la independencia en los lanzamientos sucesivos de un dado y asigna la misma probabilidad a los dos sucesos. Luego aplica la regla de Laplace explícitamente con notación adecuada. Un ejemplo (sujeto 29) se incluye a continuación:

números impares = 1, 3, 5 $\Rightarrow P(\text{impar}) = \frac{3}{6}$

números pares = 2, 4, 6 $\Rightarrow P(\text{par}) = \frac{3}{6}$

La probabilidad de sacar un número par es 3/6 porque hay 6 números en un dado y tres de ellos son pares.

(Sujeto 29).

La mayoría de los casos los futuros profesores especifican que se trata de experimentos independientes sin tener en cuenta los resultados anteriores. En la siguiente respuesta (sujeto 22) se calcula la probabilidad de cada resultado y se aplica la regla de la suma para calcular los sucesos par e impar.

$\boxed{1/2}$ porque son sucesos independientes y hay 3 números pares y 3 números impares

$P(\text{par}) = \frac{3}{6} = \frac{1}{2} \rightarrow P(2)+P(4)+P(6) = \frac{1}{6}+\frac{1}{6}+\frac{1}{6} = \frac{3}{6}$

$P(\text{impar}) = \frac{3}{6} = \frac{1}{2} \rightarrow P(1)+P(3)+P(5) = \frac{1}{6}+\frac{1}{6}+\frac{1}{6} = \frac{3}{6}$

(Sujeto 22).

C2.2. *Fórmula correcta de la probabilidad condicional.* Son los futuros profesores que identifican correctamente la independencia de ensayos y además identifican en el enunciado un problema de probabilidad condicional. Usan una notación adecuada y resuelven el problema aplicando la fórmula correspondiente. Por ejemplo, el sujeto 41 aplica la fórmula de la probabilidad condicionada para calcular la probabilidad de que salga par la siguiente tirada, teniendo en cuenta que los lanzamientos son experimentos independientes.

La probabilidad es de 0,5 por lo siguiente,

$P(I|15P)$ siendo I: Sacar número par en esta tirada.
15P: Haber sacado los resultados que aparecen en el enunciado en las 15 tiradas previas.

$$P(I|15P) = \frac{P(I \cap 15P)}{P(15P)} = \frac{P(I) \cdot P(15P)}{P(15P)} = P(I) = 0{,}5$$

(Sujeto 41).

En la Tabla 4.7.29 presentamos los resultados de las distintas categorías del ítem 11. El mayor porcentaje se da para la categoría C2.1, calcular correctamente la probabilidad haciendo uso de la regla de Laplace, con un 77% de las respuestas. La segunda categoría por orden de aparición ha sido el error C0.1, la denominada como falacia del jugador (calcular la probabilidad de ocurrencia en función de los resultados anteriores), con un 6,1%. La siguiente categoría por orden de aparición ha sido la C1 con un 5,1%. Estos futuros profesores han interpretado los resultados en función de los resultados obtenidos anteriormente o han interpretado el dado como trucado.

Tabla 4.7.29. Frecuencia y porcentaje de respuestas en el ítem 11

	Frecuencia	Porcentaje
C0.1	12	6,1
C0.2	4	2,0
C0.3	3	1,5
C0.4	5	2,6
C1	10	5,1
C2.1	151	77,0
C2.2	2	1,1
En blanco	9	4,6

Tabla 4.7.30. Resultados en el ítem 11

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	33	16,8	112	27,0
C1	10	5,1	84	20,3
C2	153	78,1	218	52,7
Total	196	100,0	414	100,0

Como se observa en la Tabla 4.7.30 el 78,1% de los sujetos de nuestra muestra y el 52,7% de los alumnos de Psicología comprenden que las ocurrencias anteriores no afectan a la probabilidad de un suceso cuando los experimentos son independientes, como es el caso del muestreo sin reposición. Un 16,8% de los nuestros participantes y un 27% de los alumnos de Psicología presentan la llamada "falacia del jugador" ya que consideran que para el cálculo de la probabilidad pedida se debe tener en cuenta los resultados anteriores. Y por último el 5,1% y el 20,3% plantean el problema como una estimación frecuencial de la probabilidad.

Tabla 4.7.31. Resultados en el ítem 11 (Matemáticas y Máster de secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	15	15,8	18	17,8
C1	5	5,3	5	5,0
C2	75	78,9	78	77,2
Total	95	100,0	101	100,0

Comparando las dos submuestras de nuestra investigación en la Tabla 4.7.31, tenemos que el porcentaje de respuestas correctas es semejante para los alumnos de la licenciatura y del máster, con un 78,9% y un 77,2% respectivamente. Ocurre lo mismo en el caso de los que interpretan los resultados en función de los resultados anteriores. En el caso de los errores, el porcentaje de los que cometen alguno de los errores descritos es mayor para la submuestra de alumnos del Máster con un 2% más que los alumnos de la Licenciatura. En resumen, este ítem fue sencillo para los futuros profesores, que aplicaron mejor el concepto de independencia en este ítem que en el de opción múltiple relacionado con este concepto.

4.7.5. CÁLCULO DE LA PROBABILIDAD COMPUESTA

Ítem 14. Un grupo de alumnos de un colegio se examina de Matemáticas e Inglés. La proporción de alumnos que aprueban Matemáticas es del 80% y la proporción de alumnos que aprueba Inglés es del 70%. Suponiendo que las notas en cada una de las asignaturas son independientes, ¿Cuál es la probabilidad de que un alumno escogido al azar haya aprobado ambas asignaturas?

En el cálculo de la probabilidad compuesta es necesario diferenciar si los sucesos son o no dependientes para poder aplicar adecuadamente la regla del producto. Este ítem pretende evaluar si los alumnos son capaces de calcular una probabilidad conjunta en el caso de sucesos independientes, mientras el ítem 12 propone un problema de cálculo de probabilidad condicional en sucesos dependientes.

Siguiendo a Díaz (2007), hemos puntuado las diferentes respuestas proporcionadas por los futuros profesores en tres categorías: (C0) no identifican correctamente los datos; (C1) identifica correctamente los datos pero no resuelve el problema debido a diferentes errores y (C2) solución correcta que se describen a continuación, así como sus diversas variantes.

C0. *Identificación incorrecta de los datos, que lleva a responder incorrectamente.* Díaz (2007) puntuó con cero puntos los casos en que no se identifica correctamente los datos y el error en la respuesta es lo suficientemente serio para impedir llegar a una estimación de la probabilidad pedida. Nosotros hemos desglosado esta categoría encontrando las siguientes:

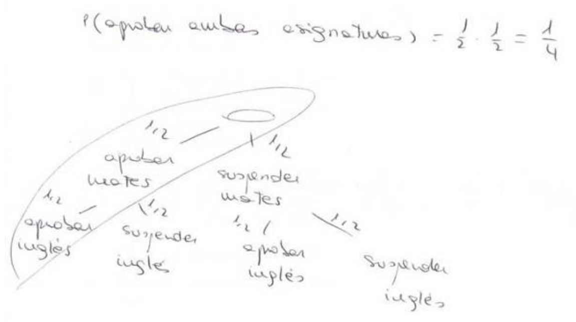
C0.1. *Confusión de probabilidades y frecuencias.* No se logra identificar correctamente los datos pues confunde las probabilidades con frecuencias, error ya encontrado en el estudio 1 realizado con profesores de primaria. Así, el sujeto 23 calcula la probabilidad de aprobar las dos asignaturas dividiendo el número de alumnos que aprueban inglés entre el número de alumnos que aprueban matemáticas, es decir, trata de aplicar la regla de Laplace, pero confunde las probabilidades del enunciado con casos favorables y posibles:

"La probabilidad de que un alumno elegido al azar haya aprobado ambas asignaturas es 70/80" (Sujeto 23).

C0.2. *Calcular el promedio de las dos probabilidades simples*. En el siguiente ejemplo, se calcula la probabilidad de aprobar las dos asignaturas calculando la media de la probabilidad de aprobar inglés y la probabilidad de aprobar matemáticas. Es claro en la respuesta que no se ha identificado correctamente ni la probabilidad pedida ni los datos del problema.

" $\frac{80+70}{2}\cdot 100 = 75\%$ *de probabilidades"* (Sujeto 32).

C0.3. *Asignar a las dos probabilidades de aprobar el valor de 1/2*. En la siguiente respuesta, el sujeto 37 calcula la probabilidad de aprobar las dos asignaturas asignando a la probabilidad de aprobar inglés y a la probabilidad de aprobar matemáticas el valor de 1/2, para después multiplicar ambas probabilidades. Tampoco identifica los datos y aunque aplica la fórmula de la probabilidad compuesta para el caso de sucesos independientes (identificando correctamente la independencia), la aplica como si no se proporcionara información previa sobre el porcentaje de aprobados en cada asignatura.

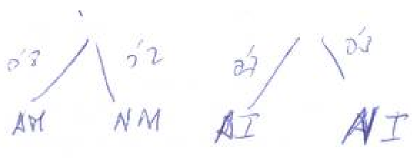


(Sujeto 37).

C0.5. *Confundir la probabilidad de un sujeto con la inversa del primer digito de los porcentajes*. Una respuesta de este tipo la da el sujeto 94 multiplica 1/8 por 1/7 para obtener la probabilidad conjunta, obteniendo un valor 1/56.

C1. *Identifica correctamente los datos y construye correctamente el árbol, pero no resuelve el problema.* Díaz (2007) puntuó con un punto en este problema a los estudiantes que, identificando los datos correctamente no llegan a acabarlo, bien porque no identifican la pregunta del enunciado o cometen un error. Entre ellos hemos encontrado las siguientes categorías:

C1.1. *Plantear el problema sin llegar resolverlo*. Por ejemplo, el sujeto 9, desarrolla un diagrama de árbol para calcular las probabilidades solicitadas. Ha identificado los sucesos dados en el enunciado, calcula las probabilidades de los complementarios (suspender cada asignatura) pero no continúa el problema. Dibuja las ramas del árbol separadas, lo que puede ser debido a identificar sucesos independientes y mutuamente excluyentes.

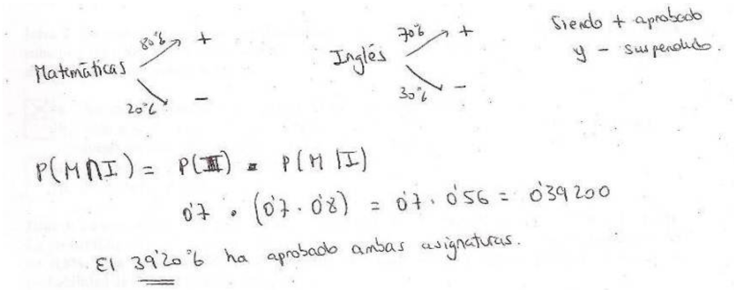


(Sujeto 9).

C1.2. *Confundir la probabilidad conjunta con la probabilidad de la unión de dos sucesos*. Se aplica la regla de la suma en lugar de la regla del producto. El siguiente futuro profesor calcula el porcentaje de alumnos que suspenden cada asignatura, identificando los datos del enunciado y obteniendo el suceso complementario en ambos casos. Una vez encontradas las probabilidades de suspender cada asignatura, se suman, para calcular la probabilidad de suspender al menos una de las dos. Es decir, aplica la regla de la suma para el caso de sucesos mutuamente excluyentes. Por tanto, confunde los sucesos independientes con sucesos excluyentes, error señalado por Sánchez (1996) y Kelly y Zwiers (1986), que ya encontramos en los ítems anteriores. Se obtiene como resultado 50%, por lo que asigna que la probabilidad de aprobar sería el complementario de ésta.

> *“20 de cada cien suspenden matemáticas; 30 de cada cien suspenden inglés. Luego la probabilidad es del 50%”* (Sujeto 45).

C1.3. *Calcular la probabilidad conjunta aplicando la regla del producto, pero, no interpretando los sucesos como independientes*. Por ejemplo, el sujeto 123 calcula la probabilidad conjunta a partir del producto de la probabilidad de aprobar inglés y de la probabilidad de aprobar matemáticas habiendo aprobado inglés, es decir, sin considerar que estos sucesos son independientes, y aplicando la regla del producto para caso de sucesos dependientes. Los datos están correctamente identificados, pero la solución es incorrecta. Además, al tratar de calcular la probabilidad condicional de aprobar matemáticas habiendo aprobado inglés, introduce un error en la fórmula.



(Sujeto 123).

C1.4. *Error en la fórmula de la probabilidad conjunta.* En el siguiente caso, se calcula la probabilidad conjunta como la suma del producto de la mitad de cada probabilidad simple. En la notación que usa en el primer término de la igualdad no está claro si interpreta el problema como la unión o intersección de sucesos.

> *“P(aprobar mates + aprobar inglés)=* $\frac{1}{2}(0,5\cdot 0,8)+\frac{1}{2}(0,5\cdot 0,7)$ *”* (Sujeto 117).

C2. *Resuelve correctamente*. Díaz puntuó con dos puntos en este problema a los estudiantes que identifican correctamente los datos y obtienen la probabilidad pedida. Nosotros hemos desarrollado en varias esta categoría.

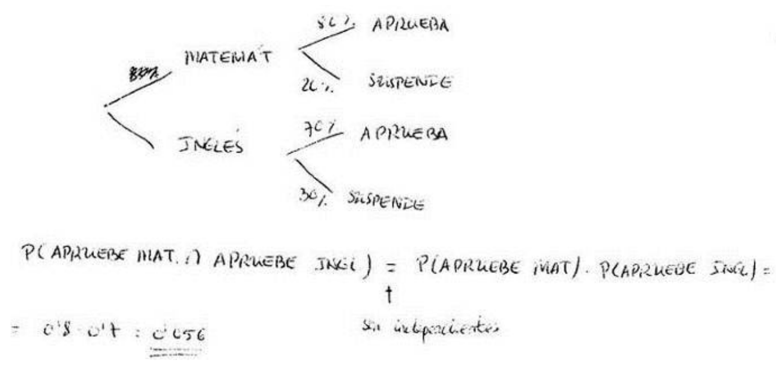
C2.1. *Resuelve correctamente sin utilizar fórmulas*. El sujeto 35 calcula la probabilidad conjunta sin utilizar notación matemática. Suponemos que ha resuelto usando cálculo mental, pues indica con palabras que ha usado la fórmula del producto y además de dar un resultado correcto, indica que los sucesos son independientes:

> "*La probabilidad de que apruebe las dos asignaturas es el producto de las dos probabilidades simples ya que son independientes, dando como resultados 0,56*" (Sujeto 35).

C2.2. *Resuelve correctamente utilizando la fórmula de la probabilidad conjunta para sucesos independientes*. En el siguiente ejemplo, se utiliza explícitamente la fórmula correcta y la propiedad de descomposición de la probabilidad conjunta para sucesos independientes.

> "*Como P(Mat)=0,8 y P(Ing)=0,7 entonces por ser independientes los dos sucesos tenemos que* $P\left(Mat \cap Ing\right) = P(Mat)\cdot P(Ing) = 0,8\cdot 0,7 = 0,56$" (Sujeto 1).

C2.3. *Resuelve correctamente interpretando el resultado como de probabilidad conjunta, utilizando un diagrama de árbol para dar la solución correcta*. Por ejemplo el sujeto 55, que aparte de interpretar el diagrama de árbol y responder correctamente a la cuestión, interpreta los dos sucesos como independientes.



(Sujeto 55).

En la Tabla 4.7.32 se observan las frecuencias de las distintas categorías del ítem 14. El mayor porcentaje se da para la categoría C2.2, que consiste en resolver correctamente utilizando la fórmula de la probabilidad conjunta para sucesos independientes, con un 41,9% de las respuestas. Esto demuestra que los futuros profesores muestran bastante competencia en la resolución de problemas de probabilidad compuesta para sucesos independientes. La segunda categoría por orden de aparición ha sido la C2.1, resuelve correctamente sin utilizar fórmulas, con un 34,3%. La siguiente ha sido la C2.3, pero ya con poca presencia (2,6% de las respuestas).

Tabla 4.7.32. Frecuencia y porcentaje de respuestas en el ítem 14

	Frecuencia	Porcentaje
C0.1	2	1,0
C0.2	3	1,5
C0.3	1	0,5
C0.4	1	0,5
C1.1	6	3,0
C1.2	3	1,5
C1.3	1	0,5
C1.4	1	0,5
C2.1	67	34,3
C2.2	82	41,9
C2.3	5	2,6
En blanco	24	12,2

Comparando nuestra muestra con la de Díaz vemos en la Tabla 4.7.33 que los resultados son mucho mejores en los futuros profesores, ya que el porcentaje de respuestas correctas de la muestra global fue un 26,1% mayor que la de Psicología, siendo también el porcentaje de errores mucho menor, un 11,3% menos. Ocurre lo mismo con la categoría 1 que se dio en un 14,8% menos en la muestra de futuros profesores. La mayor preparación matemática fue visible, por la facilidad que supuso para nuestros participantes la resolución de este problema.

Tabla 4.7.33. Resultados en el ítem 14

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	31	15,7	112	27,0
C1	11	5,5	84	20,3
C2	154	78,8	218	52,7
Total	196	100,0	414	100,0

Comparando las dos submuestras (Tabla 4.7.34), observamos pocas diferencias, aunque los resultados son ligeramente mejores en los alumnos de la Licenciatura de Matemáticas. El porcentaje de respuestas correctas del grupo de la licenciatura fue mayor en un 7,7% que la del grupo del Máster. El porcentaje de errores fue menor en la licenciatura, un 1,9% menos que en la muestra del Máster, aunque en ambos grupos muy bajos. De una forma parecida ocurre en la categoría 1 donde el porcentaje de alumnos es más alto en un 5,8%.

Tabla 4.7.34. Resultados en el ítem 14 (Matemáticas y Máster de secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	14	14,9	17	16,8
C1	2	2,1	8	7,9
C2	78	83,0	76	75,3
Total	95	100,0	101	100,0

Ítem 12. En una encuesta publicada en un periódico se informa que el 91% de los habitantes de una ciudad mienten usualmente y de ellos el 36% mienten sobre cosas importantes. Si nos encontramos con una persona al azar de esta ciudad, ¿cuál es la probabilidad de que mienta sobre cosas importantes?

En este caso se trata de evaluar la capacidad de resolución de problemas de probabilidad compuesta para sucesos dependientes, también en una situación familiar al alumno. Siguiendo a Díaz (2007) se han definido tres categorías: (C0) No llega a identificar correctamente los datos del problema; (C1) identifica los datos o realiza un diagrama en árbol correcto, pero no llega a resolver el problema y (C2) resuelve el problema correctamente. A continuación describimos estas categorías que hemos dividido en varios apartados.

C0. *Responde incorrectamente no llegando a la identificación de todos los datos*. Díaz (2007) puntuó como cero los casos en que el error en la respuesta es lo suficientemente serio para impedir llegar a una estimación de la probabilidad pedida. En esta categoría hemos incluido a aquellos que no dejan claro si identifican los datos del enunciado, pues sólo dan una respuesta donde confunden la probabilidad conjunta por la condicional. No se han realizado más categorías para este tipo de respuestas. Una respuesta de este tipo la da el sujeto 46, que interpreta incorrectamente la pregunta del enunciado como de probabilidad condicional (en lugar de probabilidad conjunta). Por otro lado interpreta incorrectamente el porcentaje del 36% como la probabilidad conjunta en lugar de la probabilidad condicional de, habiendo mentido, que la mentira sea sobre cosas importantes. Encontramos de nuevo en este alumno la confusión ya citada entre probabilidad conjunta y condicional.

"P("mientan sobre cosas importantes")=P("mientan sobre cosas importantes"/"mientan")

$P(B/A)=\frac{P(A\cap B)}{P(A)}=\frac{\frac{36}{100}}{\frac{91}{100}}$" (Sujeto 46).

C1. *Identifican los datos o realizan el diagrama en árbol correctamente, pero luego cometen algún tipo de error, bien al identificar la probabilidad que tienen que calcular, o al aplicar la fórmula.* Hemos clasificado las respuestas de esta categoría en los siguientes niveles:

C1.1. *Identifica correctamente los datos del enunciado, así como la pregunta, interpretándola como probabilidad conjunta, pero toma los sucesos como independientes*. En el siguiente ejemplo no se tiene en cuenta que el suceso "mentir sobre cosas importantes" depende del suceso "mentir", por tanto no existe independencia.

> *"P("mientan sobre cosas importantes")=* $P(A \cap B) = P(A)P(B) = 0{,}91 \cdot 0{,}36$; ya que son independientes" (Sujeto 130).

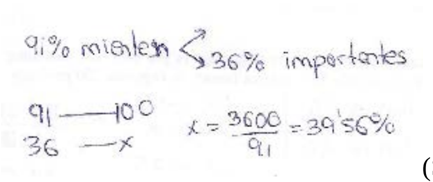
C1.2. *Error en el uso de la fórmula, habiendo identificado los sucesos y la probabilidad que se pide.* El siguiente sujeto identifica correctamente la probabilidad conjunta pero a la hora de calcularla, en lugar de aplicar correctamente la fórmula de la probabilidad conjunta, divide la probabilidad de la intersección por la probabilidad simple en lugar de multiplicarlas, error que ya habíamos encontrado en respuestas anteriores:

> "*La probabilidad de que mienta en cosas importantes es 36/91"* (Sujeto 23).

C1.3. *Errores de cálculo.* Cuando el sujeto identifica correctamente los datos y la pregunta planteada, elije una fórmula correcta para la probabilidad condicional, pero comete errores aritméticos o de otro tipo al desarrollarla. Un ejemplo se incluye a continuación:

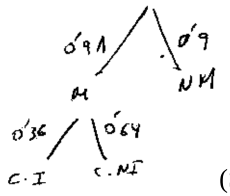
> *"Probabilidad de que un habitante mienta sobre cosas importantes = P(A)*
>
> $P(A) = \frac{91}{100}\frac{36}{100} = 0{,}91 \cdot 0{,}36 = 0{,}1476$ *Por tanto un 14,76*% " (Sujeto 25).

C1.4. *Identifica los datos y realiza el diagrama en árbol, pero no identifica la pregunta del problema, interpretándolo como un problema de proporcionalidad que resuelve una regla de tres.* Por ejemplo el sujeto 74 calcula el porcentaje de personas que mienten sobre cosas importantes dentro de los que mienten.



(Sujeto 74).

C1.5. *Interpreta correctamente los datos y realiza el diagrama de árbol, pero no calcula la probabilidad pedida.* Reproducimos la respuesta del sujeto 30, quien interpreta correctamente los diferentes sucesos y la dependencia entre ellos a partir del diagrama de árbol, pero no calcula el resultado final.



(Sujeto 30).

C2. *Resuelve correctamente el problema.* Díaz (2007) puntúa con 2 puntos los casos en que resuelven el problema correctamente empleando la notación adecuada. Hemos clasificado las respuestas de esta categoría en las siguientes:

C2.1. *Resuelve correctamente sin utilizar fórmulas.* No se explicita el tipo de probabilidad que va a calcular, pero se deduce de los cálculos realizados y la solución es correcta. En el siguiente ejemplo se calcula la probabilidad conjunta como producto de la probabilidad de mentir por la probabilidad de mentir sobre cosas importantes. Por tanto se identificó todos los datos del problema así como la pregunta, contestando adecuadamente*:*

*“P=0.91*0,36=0.33”* (Sujeto 1).

C2.2. *Resuelve correctamente, interpretando el resultado como probabilidad conjunta, utilizando la descomposición de la probabilidad conjunta como producto de una probabilidad simple por una condicionada.* Así, el sujeto 3 interpreta los sucesos como dependientes, usa una notación adecuada para representar los sucesos, sus operaciones y las diversas probabilidades y aplica la fórmula de la probabilidad condicional para calcular la probabilidad conjunta.

N_1 = mentir $\qquad P(N_1) = 0'91$

N_2 = sobre cosas importantes $\qquad P(N_2 \cap N_1)$ es la incógnita

Pero sabemos, que $P(N_2/N_1) = 0'36 = \dfrac{P(N_2 \cap N_1)}{P(N_1)} \Rightarrow$

$\Rightarrow P(N_2 \cap N_1) = P(N_2/N_1) \cdot P(N_1) =$

$= 0'36 \cdot 0'91 = \boxed{0'33}$ (Sujeto 3).

C2.3. *Resuelve correctamente, utilizando el teorema de la probabilidad total.* Es una solución un poco más complicada, pero correcta. En estos casos se deduce las probabilidades de los complementarios de algunos sucesos, necesarios para posteriormente aplicar la regla de la probabilidad total. En la siguiente respuesta, el sujeto 4 identifica correctamente los datos y la probabilidad pedida. Realiza una partición del espacio muestral, deduciendo las probabilidades de no mentir (complementario de mentir) y la probabilidad condicional de mentir sobre cosas importantes cuando no se miente (0, como razona el alumno). Con todos estos datos, calcula la probabilidad a partir del teorema de la probabilidad total como unión de dos probabilidades conjuntas a partir de dos sucesos dependientes.

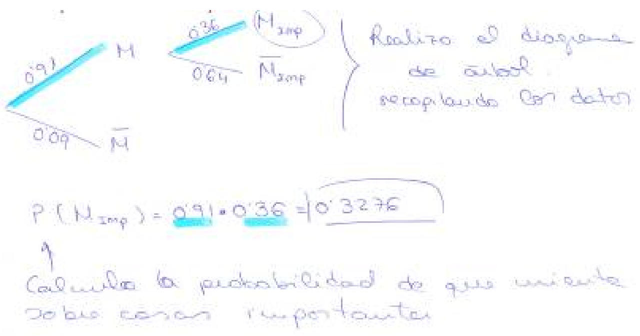
Llamo A al suceso "la persona miente usualmente"
y B al suceso "la persona miente sobre cosas importantes"
Tenemos que $P(A) = 0'91$ y $P(B/A) = 0'36$.
Nos piden calcular $P(B)$
Acudimos a la ley de probabilidad total con la partición formada por A y A^c.

$$P(B) = \underbrace{P(A)}_{0'91}\,\underbrace{P(B/A)}_{0'36} + \underbrace{P(A^c)}_{0'09}\,\underbrace{P(B/A^c)}_{0} = 0'91 \cdot 0'36 = \boxed{0'3276.}$$

Pues si la persona no miente, tampoco va a mentir sobre cosas importantes.

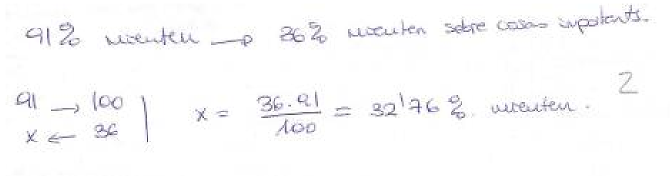
(Sujeto 4).

C2.4. *Resuelve correctamente interpretando el resultado como de probabilidad compuesta, utilizando un diagrama de árbol para dar la solución*. La solución es similar a otras anteriores, pero en lugar de usar la fórmula, el alumno la deduce a partir de un diagrama en árbol correctamente construido. Por ejemplo, el sujeto 13 dibuja el diagrama de árbol para identificar los sucesos que intervienen y la dependencia entre ellos, asignando adecuadamente probabilidades en cada rama para luego deducir la probabilidad conjunta multiplicando las probabilidades de las ramas que llevan al suceso de interés.



(Sujeto 13).

C2.5. *Resuelve correctamente utilizando una regla de tres*. Aunque no está claro que el estudiante identifique la probabilidad pedida, éste llega a una solución correcta por este método. Un ejemplo lo da el sujeto 141, quien interpreta que el 36% de los que mienten en cosas importantes los ha de calcular a partir del porcentaje de los que mienten, calculando tal proporción mediante una regla de tres.



(Sujeto 141).

Tabla 4.7.35. Frecuencia y porcentaje de respuestas en la definición

	Frecuencia	Porcentaje
C0	24	12,2
C1.1	5	2,6
C1.2	17	8,7
C1.3	7	3,6
C1.4	2	1,0
C1.5	6	3,1
C2.1	70	35,7
C2.2	21	10,7
C2.3	3	1,5
C2.4	29	14,8
C2.5	1	0,5
En blanco	11	5,6

Las distintas categorías de este ítem se han presentado muy repartidas y hacemos notar que una parte importante de los participantes no identifica correctamente los datos (C0). Otro grupo comete errores en la fórmula de la probabilidad condicional (C1.2). La categoría más representada (ver Tabla 4.7.35), ha sido la C2.1 (resuelve correctamente sin utilizar fórmulas) con un 35,7%. El resto de porcentajes ha estado muy repartido, destacando la categoría C2.4, resolver correctamente interpretando el resultado como de probabilidad compuesta, y utilizando un diagrama de árbol para dar la solución, con un 14,8% seguida de la C0, responde incorrectamente, no llegando a la identificación de todos los datos, con un 12,2% de respuestas.

Comparando nuestros resultados con los de Díaz vemos en la Tabla 4.7.36, que los porcentajes de respuestas correctas en este ítem en ambas muestras son muy similares, siendo éste ítem sencillo para las dos muestras. Sólo destaca el 4,3% más de alumnos de Psicología que responde incorrectamente no llegando a la identificación de todos los datos.

Tabla 4.7.36. Resultados en el ítem 12

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	35	17,9	92	22,2
C1	37	18,9	72	17,4
C2	124	63,2	250	60,4
Total	196	100,0	414	100,0

Por el contrario, en nuestros dos grupos sí encontramos diferencias, como se observa en la Tabla 4.7.37. Sorprendentemente, el porcentaje de respuestas correctas ha sido un 22,7% mayor en la muestra de alumnos del Máster, a pesar de la mejor preparación de los Licenciados en Matemáticas, posiblemente porque como hemos visto, para poder resolver el problema una primera dificultad es identificar los datos del enunciado y el porcentaje de alumnos de la licenciatura es mayor, un 7,3%, respecto a responder incorrectamente no identificando los datos.

Tabla 4.7.37. Resultados en el ítem 12 (Matemáticas y Máster de secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	22	23,2	13	12,9
C1	24	25,3	13	12,9
C2	49	51,5	75	74,2
Total	95	100,0	101	100,0

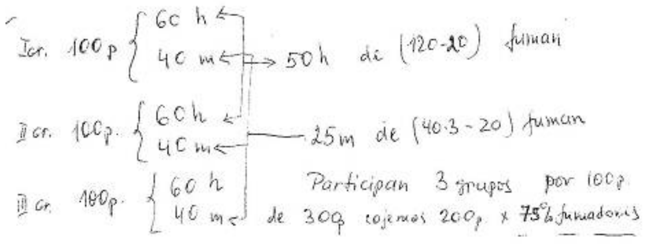
La categoría C1 está más representada en el grupo de Matemáticas con un 25,3% de respuestas y sólo se presenta en la muestra de alumnos del Máster en un 12,9%. Esta categoría, como hemos visto, incluye alumnos que no identifican la pregunta o bien confunden sucesos dependientes e independientes o hacen errores en la fórmula. En consecuencia parece que la mayor preparación formal no ha sido productiva para problemas sencillos como este, donde los alumnos del Máster han tenido más éxito en su resolución a partir del uso de recursos como el diagrama en árbol.

4.7.6. TEOREMA DE LA PROBABILIDAD TOTAL

Ítem 10. En una ciudad hay 60 hombres y 40 mujeres por cada 100 habitantes. 50 de cada 100 hombres y 25 de cada 100 mujeres fuman. Si cogemos 200 personas al azar, ¿cuántas de ellas fumarán?

Este ítem se incluye para comprobar si los estudiantes reconocen las situaciones en que puede aplicarse el teorema de la probabilidad total y son capaces de aplicarlo. Como en casos anteriores y siguiendo a Díaz (2007) se han definido tres categorías: (C0) No llega a identificar correctamente los datos del problema; (C1) Identifica los datos o realiza un diagrama en árbol correcto, pero no llega a resolver el problema y (C2) resuelve el problema correctamente. A continuación describimos estas categorías que hemos dividido en varios apartados:

C0. *No llega a identificar correctamente los datos del problema.* Díaz (2007) puntuó como cero estos casos pues el error en la respuesta es lo suficientemente serio para impedir llegar a una estimación de la probabilidad. No se han definido subcategorías en este apartado. Por ejemplo, el sujeto 69, no identifica los elementos correctamente, asignando al problema la existencia de tres grupos.



(Sujeto 69).

C1. *Identifica los datos o construye un diagrama de árbol correctamente, pero no llega a finalizar el problema.* Nosotros hemos desarrollado esta categoría en varias que se describen a continuación.

C1.1. *El alumno elabora un diagrama en árbol correcto, pero no avanza en la solución* como la siguiente solución, donde el sujeto 4 identifica correctamente los datos y realiza un diagrama, pero no identifica la pregunta:

200 < 120 hombres < 60 fuman / 60 no fuman
200 < 80 mujeres < 20 fuman / 60 no fuman

(Sujeto 4).

C1.2. *Cometer errores en las operaciones*. En la respuesta que reproducimos a continuación, el sujeto 41 identifica los datos y plantea bien el problema pero tiene errores al multiplicar los números decimales obteniendo un resultado incorrecto.

M: Ser mujer
H: Ser hombre
F: Fumar

$P(F|M)=0,25$ $\quad P(M)=0,6$

$P(F|H)=0,5$ $\quad P(H)=0,4$

$P(F)=?$ $\quad P(F|M)=\frac{P(M\cap F)}{P(M)}$

$200\cdot 0,35=70$

70 personas fumarán de las 200.

$0,5\cdot 0,4+0,25\cdot 0,6 = 0,35$

$P(F)=P(H\cap F)+P(M\cap F)=P(F|H)\cdot P(H)+P(F|M)\cdot P(M)$

(Sujeto 41).

C1.3. *Resuelve el problema sin tener en cuenta la probabilidad de ser hombre o mujer*. En estos futuros profesores se daría la falacia de las *tasas base*, descrita por Kahneman, Slovic y Tversy (1982), consistente en no utilizar una parte de los datos del enunciado, error observado en los ítems de opciones múltiples. En el siguiente ejemplo se calcula erróneamente la esperanza, ya que sólo se tiene en cuenta la probabilidad de fumar sin tener en cuenta la probabilidad de cada género, por lo que se llega a una solución incorrecta.

"$200\frac{50}{100}+200\frac{25}{100}=150$*; de las 200 personas fuman 150*" (Sujeto 52).

C1.4. *Error en el cálculo de probabilidades, al mezclar en las operaciones frecuencias absolutas y probabilidades*. Por ejemplo, el sujeto 73, en primer lugar, tiene un error en la fórmula de la probabilidad total (que sería el numerador de su fórmula). Además al calcular el numerador, en lugar de usar probabilidades usa frecuencias absolutas.

$$P(F)=\frac{P(M)\cdot P(F/M)+P(H)\cdot P(F/H)}{200}=\frac{80\cdot 50+120\cdot 100}{200}=$$

$$=\frac{4000+12000}{200}=\frac{1600\cancel{0}}{2\cancel{00}}=\boxed{80}$$

(Sujeto 73).

C2. *Llegar a dar el número esperado de personas que fuman.* Díaz (2007) dio una puntuación 2 a los alumnos que llegan a resolver correctamente el problema. Nosotros hemos subdividido en varias categorías, dando como correctas tanto a los alumnos que calculan correctamente la probabilidad total como los que calculan el valor esperado.

C2.1. *Calcular sólo la probabilidad total de personas que fuman, pero no calcula el número esperado*. Una respuesta de este tipo la da el sujeto 5 calcula la probabilidad utilizando la fórmula de la probabilidad total, pero no da el resultado esperado.

Llamo H al suceso "la persona es hombre" y M al suceso "la persona es mujer". (H y M forman una partición del espacio muestral)
Tenemos que $P(H)=0'6$, $P(M)=0'4$.
Si llamamos A al suceso "la persona fuma", por la ley de probabilidad total,

$$P(A)=\underbrace{P(H)}_{0'6}\ \underbrace{P(A/H)}_{0'5}+\underbrace{P(M)}_{0'4}\ \underbrace{P(A/M)}_{\substack{0'25\\ (25 \text{ de cada } 100)}}=0'3+0'1=\underline{0'4}$$ (40% de las personas fuman)

(Sujeto 5).

C2.2. *Resolver el problema correctamente para 100 personas y luego calcular el resultado mediante una regla de tres*. Se resuelve el problema usando porcentajes y, como se pide el total esperado sobre 200 personas, aplica una regla de tres para dar el total. Véase la siguiente solución.

100 habitantes → 60 HOMBRES
→ 40 MUJERES

50 HOMBRES FUMAN → 100
30 → 60

25 → 100
40 → 10

100 personas fuman 40
cada 200 personas fumarán 80

(Sujeto 31).

C2.3. *Calcular el resultado a partir del teorema de la probabilidad total*. El estudiante identifica correctamente los datos, plantea con notación adecuada la fórmula de la probabilidad total y la aplica correctamente. En la siguiente respuesta, se calcula la probabilidad de fumar para después multiplicarla por el número total de elementos obteniendo así el número de personas que fuman.

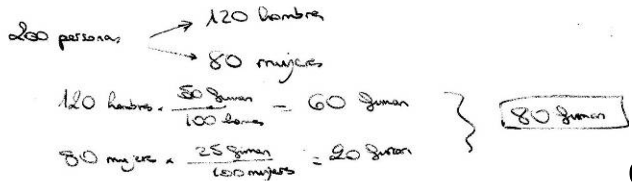
Vamos a calcular la probabilidad de que una persona fume.

$$p(\text{"fumar"}) = p(\text{"ser mujer"} \cap \text{"fumar"}) + p(\text{"ser hombre"} \cap \text{"fumar"}) =$$
$$= \frac{40}{100}\cdot\frac{25}{100}+\frac{60}{100}\cdot\frac{50}{100} = \frac{2}{5}\cdot\frac{1}{4}+\frac{3}{5}\cdot\frac{1}{2} =$$
$$= \frac{1}{10}+\frac{3}{10}=\frac{4}{10}=\frac{2}{5}$$

Fumarán $\frac{200\cdot 2}{5} = 80$ personas.

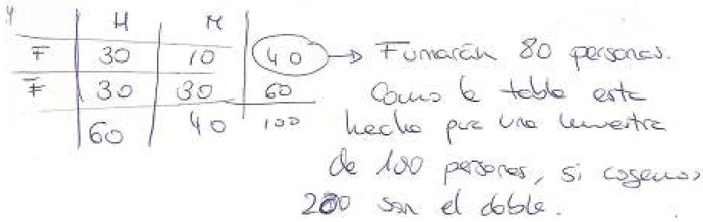
(Sujeto 4).

C2.4. *Calcula el valor esperado*. Así, el sujeto 1 calcula el número de hombres y mujeres por cada 200 habitantes y a partir de ellos calcula la esperanza como la suma del producto del número de elementos de cada género por la probabilidad de su ocurrencia.



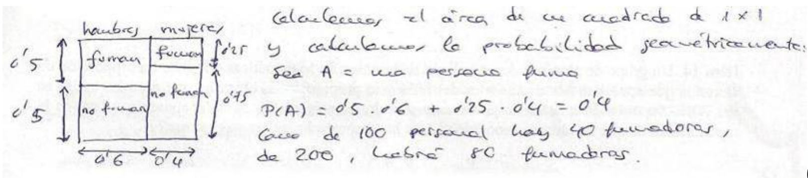
(Sujeto 1).

C2.5. *Calcula el valor esperado a partir de una tabla de contingencia*. El sujeto 18 se sirve de una tabla 2x2 para dar la solución correcta a partir de las frecuencias marginales.



(Sujeto 18).

C2.6. *Calcula la probabilidad total geométricamente a partir de un cuadrado 1x1*. En la siguiente respuesta, el sujeto 99 se sirve del área de un cuadrado 1x1 para dar la solución correcta de la probabilidad total.



(Sujeto 99).

Las categorías de este ítem están muy repartidas (ver Tabla 4.7.38), destacando principalmente la C2.3, consistente en calcular el resultado a partir del teorema de la probabilidad total, y la C2.2, calcular el resultado mediante una regla de tres, con un 28,1% y un 25% respectivamente. Del resto de categorías, destaca la C2.4, calcular el valor esperado, con un 13,8% de respuestas.

Un 74,5% de los futuros profesores de nuestra muestra fue capaz de responder correctamente a la solución esperada, aunque mayoritariamente utilizaban reglas de tres en lugar de calcular la probabilidad y luego la esperanza. Este resultado es similar con el resultado de Díaz en el que el porcentaje llegó al 64,5%, un 10% inferior. Esto puede ser debido a que el termino esperanza matemática está explícitamente incluido en la formación estadística de los futuros profesores de Matemáticas, pero no en el temario de Psicología. En cuanto a los alumnos que no identifican correctamente los resultados del problema, tenemos que el porcentaje de estudiantes de nuestra muestra es similar al encontrado en la muestra de estudiantes de Psicología.

Tabla 4.7.38. Frecuencia y porcentaje de respuestas en la definición

	Frecuencia	Porcentaje
C0	6	3,1
C1.1	2	1,0
C1.2	7	3,6
C1.3	11	5,6
C1.4	8	4,1
C2.1	9	4,6
C2.2	49	25,0
C2.3	55	28,1
C2.4	27	13,8
C2.5	4	2,0
C2.6	2	1,0
En blanco	16	8,1

Tabla 4.7.39. Resultados en el ítem 10

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	22	11,2	50	12,1
C1	28	14,3	97	23,4
C2	146	74,5	267	64,5
Total	196	100,0	414	100,0

En nuestros dos grupos encontramos diferencias (Tabla 4.7.40), ya que el porcentaje de alumnos que llega a dar el número esperado de personas que fuman o la probabilidad total de forma correcta ha sido un 11,8% mayor en la muestra de alumnos del Máster. Destaca el porcentaje de alumnos de la Licenciatura, que aunque identifican los datos, no finalizan el problema con un 13,2% más que en Máster.

Tabla 4.7.40. Resultados en el ítem 10 (Matemáticas y Máster de secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	10	10,5	12	11,9
C1	20	21,1	8	7,9
C2	65	68,4	81	80,2
Total	95	100,0	101	100,0

4.7.7. TEOREMA DE BAYES

Ítem 13. Una fábrica dispone de dos máquinas M1 y M2 que fabrican bolas. La máquina M1 fabrica el 40 % de las bolas y la M2 el 60%. El 5% de las bolas fabricadas por M1 y el 1% de las fabricadas por M2 son defectuosas. Tomamos una bola al azar que resulta ser defectuosa. ¿Cuál es la probabilidad de que haya sido fabricada por M1?

Este ítem se incluye para comprobar si los estudiantes reconocen las situaciones en que puede aplicarse el teorema de Bayes y son capaces de aplicarlo. Siguiendo a Díaz (2007), se han definido cinco categorías: (C0) No llega a identificar correctamente los datos del problema; (C1) identifica los datos o realiza un diagrama en árbol correcto pero no llega a resolver el problema; (C2) construye un diagrama en árbol adecuado o identifica los datos e identifica el problema como de probabilidad condicional; (C3) calcula correctamente la probabilidad total; y (C4) resuelve el problema correctamente. A continuación describimos estas categorías que hemos dividido en varios apartados:

C0. *No responde o responde incorrectamente, no llegando a la identificación correcta de los datos del problema.* Díaz (2007) puntuó como cero los casos en que el alumno no llega a identificar correctamente los datos. Describimos las categorías encontradas.

C0.1. C*onfunde una probabilidad condicional con su inversa,* por ejemplo el sujeto 23, que comete la falacia de la condicional transpuesta, ya señalada por Falk (1986). En vez de calcular la probabilidad pedida, calcula la probabilidad de ser defectuosa, habiéndose fabricado en la fábrica M1. Además no utiliza el resto de los datos del problema; por tanto consideramos que ni ha identificado los datos correctamente, ni tampoco la pregunta planteada. No visualiza las sucesivas particiones del espacio muestral, lo que bloquea la resolución del problema.

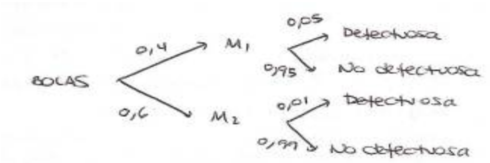
> *"La probabilidad de que haya sido fabricadas por M1 es* 5/40, *ya que como M1 fabrica el 40% de las bolas y el 5% de las bolas fabricadas por M1 son defectuosas"* (Sujeto 23).

C0.2. *Confundir probabilidad simple y condicional*. En la siguiente respuesta se entiende que como la bola es elegida al azar la probabilidad de ser defectuosa sólo depende de donde se fabricó.

"En este caso, es diferente si la bola es defectuosa o no ya que la bola es obtenida al azar. Por tanto la probabilidad será de un 40%" (Sujeto 97).

C1. *Identifica los datos o construye un diagrama en árbol adecuado, completando los datos faltantes,* pero no se identifica la probabilidad pedida y por tanto no se progresa en la solución del problema. Hemos considerado las siguientes categorías.

C1.1. *Construye correctamente el diagrama en árbol, pero no continúa el problema.* Según Díaz y de la Fuente (2007), el primer paso para resolver los problemas implica diferenciar entre probabilidad simple *P(M1), P(M2)* y probabilidad condicional *P(D/M1);* diferenciar una probabilidad condicional *P(D/M1)* y su inversa *P(M1/D);* y determinar las probabilidades de sucesos contrarios *P(C/M1),* etc. Por tanto, el estudiante debe discriminar todos estos conceptos, realizar correctamente las sucesivas particiones del espacio muestral, e identificar cuáles datos se refieren a cada uno de los conceptos anteriores en el enunciado del problema. Así, el sujeto 91 construye el diagrama de árbol del experimento correctamente pero no continúa el problema, pues, aparentemente no identifica la probabilidad pedida en el enunciado.

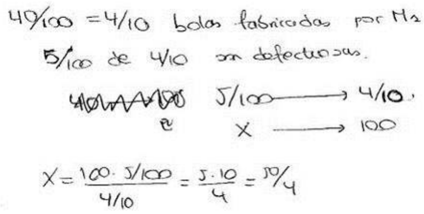


(Sujeto 91).

C1.2. *Identifica los datos, pero confunde la probabilidad pedida en el enunciado.* Por ejemplo, el sujeto 7 confunde probabilidad condicional y conjunta (Pollatsek, Well, Konold y Hardiman, 1987 y Ojeda, 1995). Por ello calcula el producto de la probabilidad de que haya sido la bola fabricada en la fábrica M1 por la probabilidad que sea defectuosa en la fábrica M1.

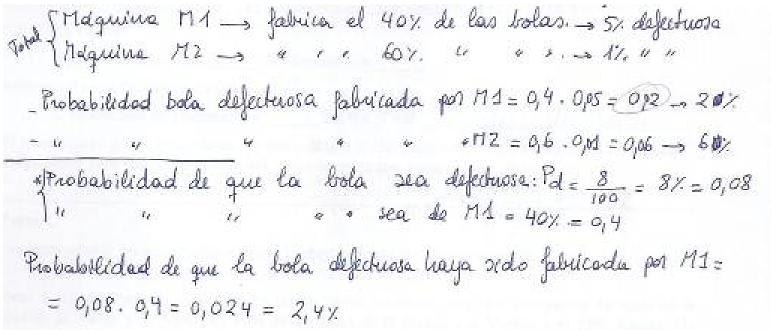
" $P = \frac{40}{100}\frac{5}{100}$ " (Sujeto 7).

C1.3. *Confunde la probabilidad pedida en el enunciado y además obtiene una probabilidad mayor que 1.* En la siguiente respuesta, el sujeto 78 ha interpretado la pregunta como de probabilidad conjunta, al igual que el ejemplo anterior. Pero, para resolver el problema, aplica una regla de tres, pues interpreta los sucesos como independientes y quiere calcular el 40% del 5%. Como consecuencia, de un error al aplicar la regla de tres, da como resultado la probabilidad 50/4 y no es consciente de que una probabilidad no puede ser mayor que la unidad.


40/100 = 4/10 bolas fabricadas por M1
5/100 de 4/10 son defectuosas.
~~40/100~~ 5/100 → 4/10.
x → 100
$x = \frac{100 \cdot 5/100}{4/10} = \frac{5 \cdot 10}{4} = 50/4$

(Sujeto 78).

C1.4. *Confunde la probabilidad pedida en el enunciado*, aunque llega a aplicar el teorema de la probabilidad total con algunos errores de cálculo. Un ejemplo, es el sujeto 25, que calcula primero la probabilidad total de ser defectuosa, pero se confunde en la probabilidad pedida, pues calcula finalmente la probabilidad conjunta de ser defectuosa y haberse fabricado en la fábrica M1. Además comete errores en las operaciones aritméticas, por lo que el valor de la probabilidad total es erróneo.


Total { Máquina M1 → fabrica el 40% de las bolas. → 5% defectuosa
Máquina M2 → " " " 60%. " " " → 1% " "
- Probabilidad bola defectuosa fabricada por M1 = 0,4 · 0,05 = 0,02 → 2%
- " " " " " " M2 = 0,6 · 0,01 = 0,06 → 6%
* Probabilidad de que la bola sea defectuosa: $P_d = \frac{8}{100}$ = 8% = 0,08
" " " " " sea de M1 = 40% = 0,4
Probabilidad de que la bola defectuosa haya sido fabricada por M1 =
= 0,08 · 0,4 = 0,024 = 2,4%

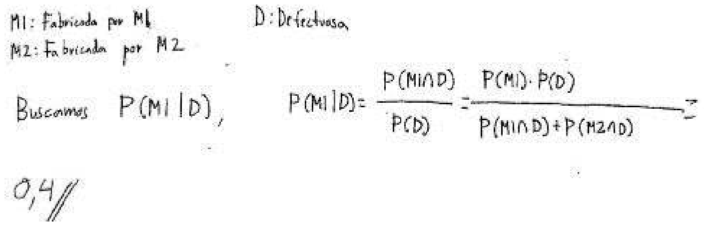
(Sujeto 25).

C2. *Diagrama correcto e identifica el problema como de probabilidad condicional, pero hace algún tipo de error, que le lleva a dar una solución incorrecta.* Para continuar, los futuros profesores deben identificar qué probabilidad se pide en el problema, y que ésta es una probabilidad condicional inversa. No es un paso sencillo, pues algunos autores indican que los estudiantes asocian el condicionamiento con el orden temporal de los sucesos y no encuentran natural que se condicione un suceso por otro que ocurre con posterioridad (falacia del eje de tiempos, descrita por Falk (1986) o concepción cronologista de la probabilidad condicional identificada por Gras y Totohasina (1995)). Hemos encontrado las siguientes categorías:

C2.1. *No tener en cuenta la proporción de bolas fabricadas por una y otra máquina*. En el siguiente caso, se interpreta correctamente los datos e identifica el problema como de la probabilidad condicionada pero a la hora de calcular las probabilidades, se comete un error en el cálculo de la probabilidad total, pues no tiene en cuenta la proporción de bolas producidas en cada fábrica.

"$P\left(M1/\,defectuosa\right)=\dfrac{P\left(M1\cap defec\right)}{P(defec)}=\dfrac{0,5\cdot 100}{0,5\cdot 100+0,1\cdot 100}=\dfrac{0,5}{0,6}\cong 1$" (Sujeto 63).

C2.2. *Tomar los sucesos como independientes*. Aunque se identifica los datos y el problema como de probabilidad condicional, aplica la regla del producto para sucesos independientes. Por ejemplo, el sujeto 41, multiplica la probabilidad de ser defectuoso por la probabilidad de que sea fabricado en la fábrica M1. Por otro lado, no queda claro el desarrollo del denominador.



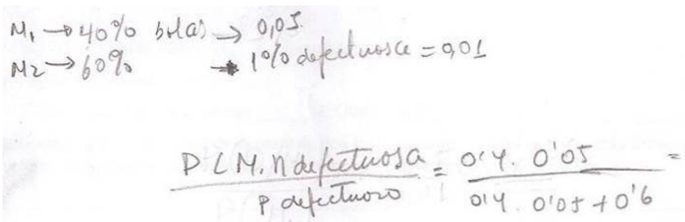
M1: Fabricada por M1 D: Defectuosa
M2: Fabricada por M2

Buscamos $P(M1|D)$, $P(M1|D)=\dfrac{P(M1\cap D)}{P(D)}=\dfrac{P(M1)\cdot P(D)}{P(M1\cap D)+P(M2\cap D)}=$

0,4//

(Sujeto 41).

C2.3. *Errar en la fórmula de Bayes*. En la siguiente respuesta, el sujeto 7 identifica correctamente el problema pero confunde las probabilidades que aparecen en la probabilidad total a la hora de definir el denominador del teorema de Bayes, por lo cual ni siquiera lleva al cálculo correcto de la probabilidad total.***



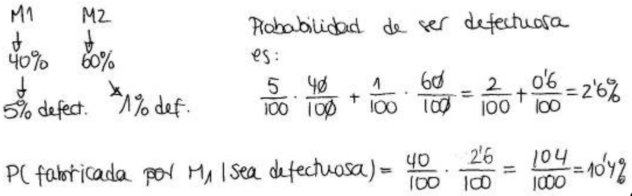
$M_1 \to 40\%$ bolas $\to 0{,}05$
$M_2 \to 60\%$ $\to$ 1% defectuosa = 0,01

$\dfrac{P(M_1 \cap \text{defectuosa}}{P \text{ defectuoso}} = \dfrac{0'4 \cdot 0'05}{0'4 \cdot 0'05 + 0'6} =$

(Sujeto 7).

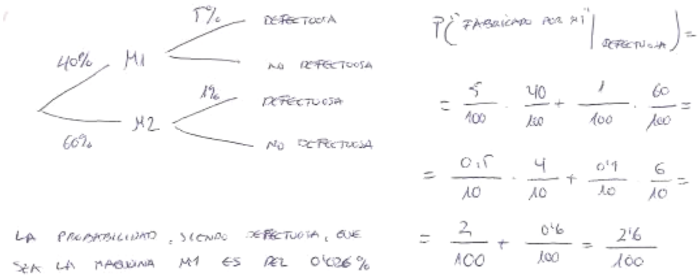
C3. *Calcular correctamente la probabilidad total*. Después de identificar el problema como el cálculo de una probabilidad condicional y recordar la fórmula de Bayes, el futuro profesor debe calcular el numerador y denominador. El denominador debe ser calculado con la regla de la probabilidad total, esto es, multiplicando las probabilidades de cada rama del árbol y sumando cada una de esas probabilidades conjuntas. Debe entender que se trata de sucesos dependientes, para aplicar correctamente la regla del producto en este caso. Hemos encontrado varios casos:

C3.1. *El alumno identifica los datos y la pregunta del enunciado y llega a aplicar el teorema de la probabilidad total correctamente*, pero al aplicar la fórmula de la probabilidad condicional hace algún tipo de error. Por ejemplo, el sujeto 80 calcula correctamente la probabilidad total de ser defectuoso pero erra en el cálculo de la probabilidad condicionada que calcula como si fuese una probabilidad conjunta en caso de sucesos dependientes. Suponemos entonces que se trata de un error de fórmula, más que de un error conceptual.



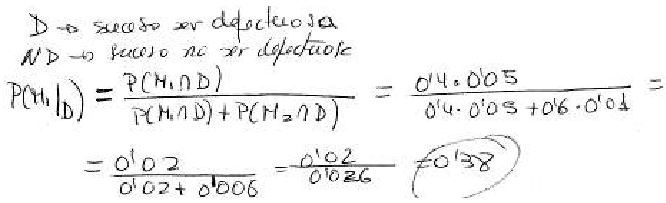
(Sujeto 80).

C3.2. *El alumno aplica correctamente el teorema de la probabilidad total y no continúa con la solución.* Por ejemplo, el sujeto 57 en lugar de calcular la probabilidad condicionada, calcula la probabilidad de que una bola sea defectuosa. Ha identificado correctamente los datos y representado un diagrama en árbol adecuado. Calcula la probabilidad total, identificando la dependencia de sucesos al aplicar la regla del producto. No queda claro si cometió un error en la probabilidad pedida pero no continúa el problema.



(Sujeto 57).

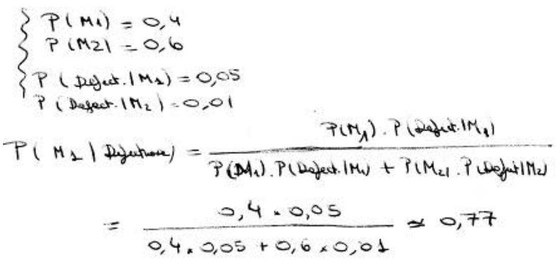
C3.3. *El alumno identifica correctamente los datos y la probabilidad pedida, incluida la fórmula de Bayes, pero tiene fallos en los cálculos realizados.* Algunos futuros profesores, como en el siguiente ejemplo, cometen fallos aritméticos, como este caso, que falla a la hora de dividir la probabilidad conjunta entre la probabilidad de la condición.



(Sujeto 22).

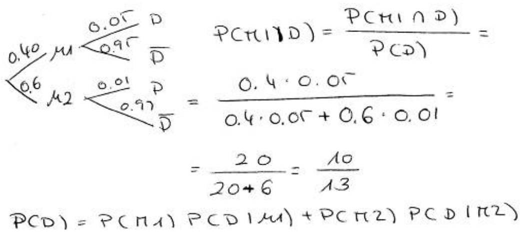
C4. *Resuelve correctamente.* Díaz (2007) puntúa con 4 puntos los casos de los estudiantes que, habiendo completado todos los pasos anteriores, finalmente, los sintetiza y calcula el numerador (probabilidad conjunta) y denominador (probabilidad total) para obtener la probabilidad inversa, es decir, aplica el teorema de Bayes. Nosotros hemos desglosado esta categoría en varias que describimos a continuación:

C4.1. *Resuelve correctamente el problema utilizando el teorema de Bayes*. Serían los futuros profesores que identifican los datos y la pregunta del problema, aplicando correctamente el teorema de Bayes, calculando la probabilidad conjunta para sucesos dependientes y la probabilidad total. Por ejemplo el sujeto 1 utiliza el teorema de Bayes, calculando anteriormente todas las probabilidades que proporciona el enunciado.



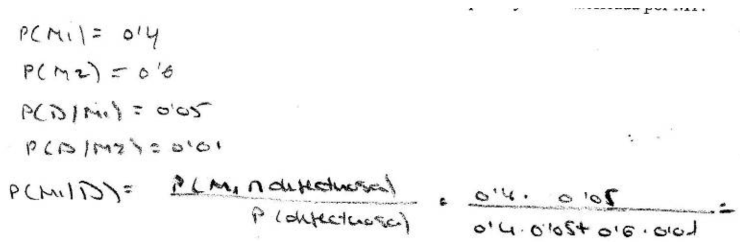
(Sujeto 1).

C4.2. *Resuelve correctamente sirviéndose del diagrama de árbol.* El alumno haría los mismos pasos que en el caso anterior, pero además se ayuda de un diagrama en árbol. En el siguiente ejemplo, se interpreta las probabilidades para construir el teorema de Bayes a partir de un diagrama de árbol.



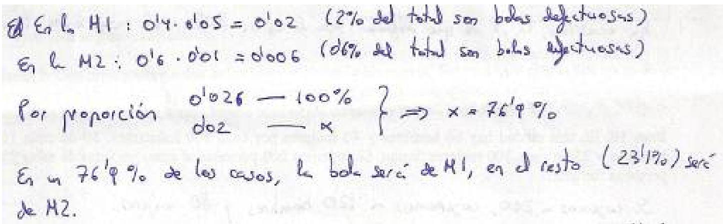
(Sujeto 18).

C4.3. *Resuelve correctamente pero deja indicada la fórmula sin calcular la probabilidad*. En la siguiente respuesta, el sujeto 19 construye el teorema de Bayes pero no realiza la operación.



(Sujeto 19).

C4.4. *Resuelve correctamente mediante una regla de tres.* En la siguiente solución, el sujeto 108 calcula la probabilidad total y mediante una regla de tres calcula la probabilidad condicional.


En la H1 : 0'4 · 0'05 = 0'02 (2% del total son bolas defectuosas)
En la H2 : 0'6 · 0'01 = 0'006 (0'6% del total son bolas defectuosas)
Por proporción 0'026 — 100% ; 0'02 — x } ⇒ x = 76'9 %
En un 76'9 % de los casos, la bola será de H1, en el resto (23'1%) será de H2.

(Sujeto 108).

Las distintas categorías de este ítem se han presentado muy repartidas. La categoría más representada, ver Tabla 4.7.41, ha sido la C4.1, resolver correctamente utilizando el teorema de Bayes, con un 26% de respuestas. El resto de porcentajes ha estado muy repartido, destacando la categoría C1.2, identificar los datos confundiendo la probabilidad pedida en el enunciado, con un 14,9% seguida de la C4.3, resolver correctamente pero dejar indicada la fórmula sin calcular la probabilidad, con un 8,7%.

Tabla 4.7.41. Frecuencia y porcentaje de respuestas en la definición

	Frecuencia	Porcentaje
C0.1	7	3,6
C0.2	2	1,0
C1.1	3	1,5
C1.2	29	14,9
C1.3	1	0,5
C1.4	3	1,5
C2.1	13	6,6
C2.2	4	2,0
C2.3	9	4,6
C3.1	9	4,6
C3.2	5	2,6
C3.3	15	7,7
C4.1	51	26,0
C4.2	4	2,0
C4.3	17	8,7
C4.4	2	1,0
En blanco	22	11,2

Comparando nuestros resultados con los de Díaz (Tabla 4.7.42) vemos que los resultados han sido similares. Aunque para la categoría C4, resolver correctamente, los resultados de la muestra de Díaz es superior en un 7,5%, el resto de resultados son muy distintos. El porcentaje de alumnos que no responden o responden incorrectamente no identificando los datos del problema fue superior en la muestra de Matemáticas pero sólo en un 2%. Ocurre algo similar para las categorías C1 y C3, donde el porcentaje es un poco superior para nuestra muestra en un 6,3% y 4,7% respectivamente.

Tabla 4.7.42. Resultados en el ítem 13

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	31	15,8	57	13,8
C1	36	18,4	50	12,1
C2	26	13,3	78	18,8
C3	29	14,8	42	10,1
C4	74	37,7	187	45,2
Total	196	100,0	414	100,0

Comparando nuestros dos grupos, Tabla 4.7.43, vemos que el porcentaje de aciertos, categoría C4, fue mayor en los alumnos del Máster de secundaria en un 24,2%. El porcentaje de las restantes categorías es muy similar en ambos grupos, tanto en aquellos que no identifican o construyen solamente el diagrama de árbol, excepto en la categoría C2 donde el porcentaje de alumnos que identifican el problema y construyen el diagrama de árbol pero cometen algún error, es superior en la muestra de alumnos de la Licenciatura de Matemáticas en un 17,1%.

Tabla 4.7.43. Resultados en el ítem 13 (Matemáticas y Máster de secundaria)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
C0	15	15,8	16	15,8
C1	18	18,9	18	17,8
C2	21	22,1	5	5,0
C3	17	17,9	12	11,9
C4	24	25,3	50	49,5
Total	95	100,0	101	100,0

4.7.8. SÍNTESIS DE RESULTADOS EN ÍTEMS ABIERTOS

Una vez analizados individualmente los ítems abiertos, en esta sección estudiaremos algunos indicadores de los conocimientos mostrados por los participantes en estos ítems, que, como se indicó, evalúan el conocimiento formal sobre probabilidad condicional.

En la Tabla 4.7.44 hemos comparado la frecuencia y porcentaje del número de alumnos que responden correctamente a cada ítem, tanto para nuestra muestra, como para la de Díaz. En ella podemos observar que las mayores dificultades, en las dos muestras, se han dado para el ítem 8, definición de la probabilidad condicional, por lo que podemos decir que la mayoría de los futuros profesores, aunque comprenden el significado de la probabilidad condicional, no saben definirla correctamente o cometen sesgos en tal definición. Otra conclusión es que los participantes en nuestra muestra presentan un porcentaje de aciertos inferior para el ítem relacionado con el teorema de Bayes que los de la muestra de Díaz (2007). Esto puede ser debido a que el temario de Psicología hace más hincapié en tales problemas, por lo que los alumnos están más familiarizados con ellos. El resto de ítems ha tenido un porcentaje mayor de respuestas correctas para la muestra de futuros profesores que para los alumnos de Psicología. Los porcentajes de aciertos han sido altos en la mayoría de ellos, algo más bajo para el ítem relacionado con la probabilidad condicional (ítem 9) en el que aproximadamente uno de cada dos alumnos respondió correctamente.

Tabla 4.7.44. Síntesis de resultados en items de conocimiento formal

	Futuros profesores (n=196)		Psicología (n=414)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Ítem 8. Definición	31	15,8	127	30,7
Ítem 9. Probabilidad condicional	112	57,1	142	34,3
Ítem 10. Probabilidad total	146	74,5	267	64,5
Ítem 11. Independencia	153	78,1	287	69,3
Ítem 12. P. compuesta, dependencia	124	63,3	250	60,4
Ítem 13. Teorema de Bayes	74	37,8	187	45,2
Ítem 14. P. compuesta, independencia	154	78,6	218	52,7

En resumen, los futuros profesores muestran un conocimiento formal alto de la probabilidad condicional, en lo que se refiere a resolución de problemas y algo menor a la hora de verbalizar una definición del concepto.

Tabla 4.7.45. Síntesis de resultados en ítems de conocimiento formal (Matemáticas y Máster)

	Matemáticas (n=95)		Máster (n=101)	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Ítem 8. Definición	19	20,0	12	11,9
Ítem 9. Probabilidad condicional	44	46,3	68	67,3
Ítem 10. Probabilidad total	65	68,4	81	80,2
Ítem 11. Independencia	75	78,9	78	77,2
Ítem 12. P. compuesta, dependencia	49	51,6	75	74,3
Ítem 13. Teorema de Bayes	24	25,3	50	49,5
Ítem 14. P. compuesta, independencia	78	82,1	76	75,2

Respecto a nuestras dos submuestras (Tabla 4.7.45), destaca el bajo porcentaje de respuestas correctas por parte de los alumnos de la licenciatura para los ítems relacionados con la definición y teorema de Bayes, un 20% y el 25,26%

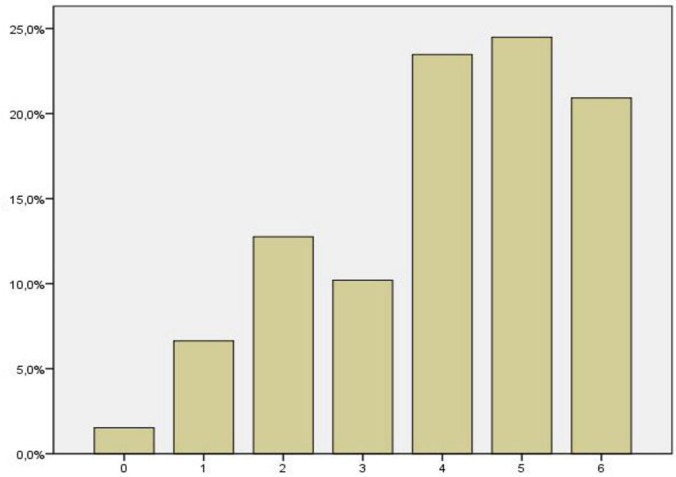
respectivamente. También los alumnos del Máster de secundaria presentan pobres resultados en la definición, donde sólo respondieron correctamente un 11,8% de respuestas correctas, pero los correspondientes al teorema de Bayes son mejores. Destacan los dos ítems relacionados con la independencia, en los que se alcanzaron los mejores resultados para ambos grupos. Estos resultados implican que estos alumnos están familiarizados con este concepto, y saben aplicarlo en el contexto planteado.

En la Tabla 4.7.46 hemos analizado el número de respuestas correctas por parte de los alumnos para los ítems abiertos. Observamos que el número de respuestas correctas ha fluctuado entre cero y seis, para ambas muestras. Ningún alumno es capaz de resolver correctamente todas las preguntas y sólo un 13,7% en el grupo de matemáticas y un 20% en los alumnos del Máster responden 6 de las 7.

Tabla 4.7.46. Número de ítems correctos por alumno, ítems 8 al 14

	Lic. Matemáticas		Máster Secundaria		Muestra Global	
Nº resp. correctas	Frecuencia	Porcentaje	Frecuencia	Frecuencia	Frecuencia	Frecuencia
0	1	1,1	2	2,0	3	1,5
1	9	9,5	4	4,0	13	6,6
2	15	15,8	10	9,9	25	12,8
3	13	13,7	7	6,9	20	10,2
4	22	23,1	24	23,8	46	23,5
5	22	23,1	26	25,7	48	24,5
6	13	13,7	28	27,7	41	20,9
7	0	0	0	0	0	0
Total	95	100,0	101	100,0	196	100,0

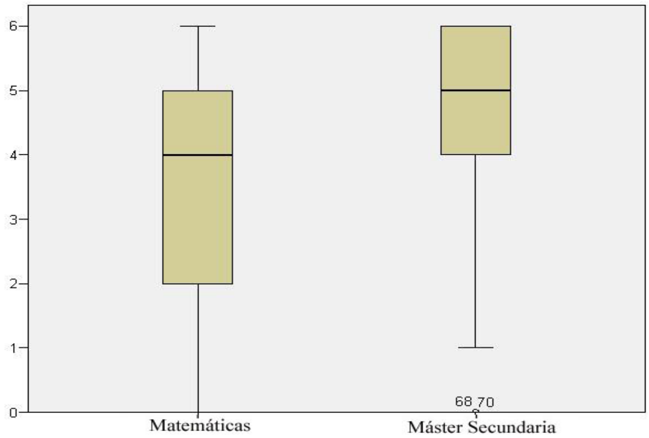
Figura 4.7.1. Diagrama de barras en ítems correctos de respuesta abierta



Como se muestra en la Figura 4.7.1, los resultados más frecuentes se dan para aquellos alumnos que contestan a cuatro, cinco o seis ítems abiertos correctamente, un

68,9% de los alumnos. Encontramos diferencias entre los dos grupos que componen nuestra muestra ya que, como se observa en la Figura 4.7.2, los alumnos del Máster de secundaria responden a más ítems abiertos correctamente, con el 75% de los alumnos por encima de cuatro ítems correctos. Por el contrario, en la muestra de alumnos de la licenciatura sólo el 50% contestó a cuatro o más ítems correctamente.

Figura 4.7.2. Diagrama de cajas, comparativa puntuación ítems correctos de respuesta abierta



Los valores descriptivos de ambas muestras (Tabla 4.7.47) indican que el valor medio de respuestas correctas fue inferior para el grupo de la Licenciatura, con una media de 3,73 respuestas correctas, comparado al 4,35 de media de respuestas correctas en la muestra del Máster. En resumen, la probabilidad condicional sigue siendo un concepto difícil de aplicar para los futuros profesores, por lo que sería necesario en incidir en su formación al respecto.

Tabla 4.7.47. Análisis descriptivo de ítems correctos por alumno, ítems 8 al 14

	Lic. Matemáticas	Máster Secundaria	Muestra Global
Media	3,73	4,35	4,05
Desv. Típica	1,59	1,55	1,60
Mínimo	0	0	0
Máximo	6	6	6

Seguidamente tomamos en cuenta las distintas puntuaciones posibles de cada ítem. Los valores descriptivos de ambas muestras (Tabla 4.7.48) indican valores medios y la dispersión aproximada, un poco inferior para la muestra de la licenciatura con 12,13 puntos de media en este grupo, aunque con resultados menos dispersos, respecto a los 12,85 puntos de media en la muestra del máster.

Como se muestra en el histograma (Figura 4.7.3), los resultados más frecuentes se presentan para los valores comprendidos entre 10 y 17 puntos, obtenidos por el 83,2% del total de la muestra. Sólo el 3% de los alumnos tuvo un valor total menor de 5 puntos y un 1% mayor de 17 puntos (un alumno). Encontramos pocas diferencias entre los dos grupos que componen nuestra muestra ya que, como se observa en la Figura 4.7.4, el 75% de los alumnos de ambos grupos tuvo un resultado superior a 10. Sólo en el caso de los alumnos que componen el 25% con menores resultados se observan diferencias ya que los alumnos del Máster presentan una mayor dispersión que los de la Licenciatura.

Tabla 4.7.48. Análisis descriptivo de ítems según puntuación, ítems 8 al 14

	Lic. Matemáticas	Máster Secundaria	Muestra Global
Media	12,13	12,85	12,5
Desv. Típica	3,39	3,68	3,55
Mínimo	4	1	1
Máximo	18	18	18

Figura 4.7.3. Puntuación total en ítems de respuesta abierta

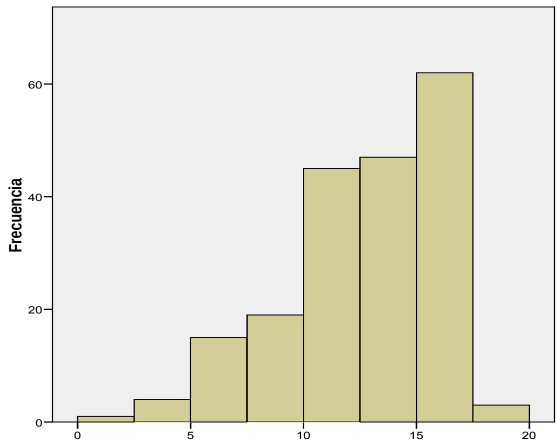
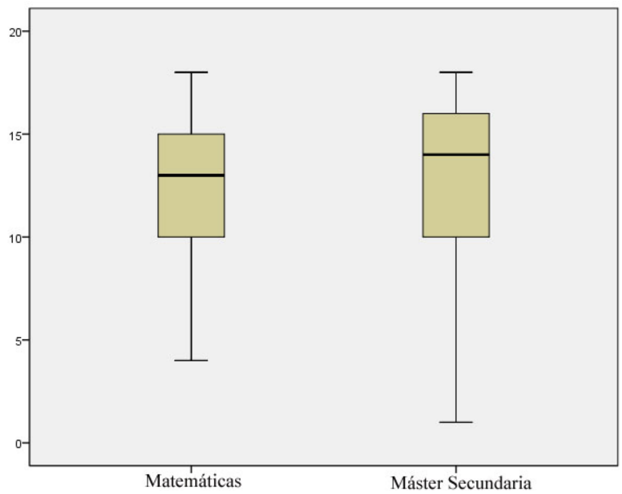


Figura 4.7.4. Diagrama de cajas, comparativa puntuación total ítem abiertos



4.7.9. RELACIÓN ENTRE CAPACIDAD DE RESOLUCIÓN DE PROBLEMAS Y SESGOS

Una vez estudiadas separadamente las respuestas correspondientes a los sesgos y las relacionadas con el conocimiento formal, en este apartado trataremos de analizar la relación existente entre las dos partes de las que se compone el cuestionario.

Correlaciones

Para estudiar la relación entre número de sesgos y capacidad de resolución de problemas se calculó, en primer lugar los coeficientes de correlación de Pearson entre las variables "número de ítems de opción múltiple correctos", "puntuación en cada ítem abierto", "puntuación total en ítems abiertos" y "número total de ítems abiertos correctos". En la Tabla 4.7.49 presentamos estas correlaciones, en las que se ha marcado con asteriscos las que son estadísticamente significativas, teniendo en cuenta las comparaciones múltiples, es decir, dividiendo el nivel de significación α=0,05 por el número total de correlaciones (45), con lo que se obtiene un nivel corregido de significación α=0,001. Como se observa, sólo se dan correlaciones estadísticamente significativas para las variables "Puntuación total ítems abiertos" y "Número ítems abiertos correctos" en relación con la variable "Número ítem de opción múltiple correctos" ya que en el resto de combinaciones el coeficiente de correlación es muy bajo.

Tabla 4.7.49. Correlaciones entre variables relacionadas con el conocimiento formal y el número ítems de opción múltiple correctos

	N. ítems de opción múltiple correctos
Ítem 8	0,142
Ítem 9	0,089
Ítem 10	0,188
Ítem 11	0,072
Ítem 12	0,166
Ítem 13	0,348
Ítem 14	0,184
Puntuación total ítems abiertos	0,350*
Número ítems abiertos correctos	0,372*

El coeficiente de correlación de Pearson para las variables "número de respuestas correctas en los ítems de opción múltiple" y "número de respuestas correctas en los ítem de respuesta abierta" da un valor de r=0,372 (p=0,000). Por tanto están relacionadas linealmente el número de respuestas correctas en los primeros ítems con el número de respuestas correctas en los restantes, aunque la relación entre ambas variables es baja, pues sólo explicaría el 13,8% de la varianza.

De igual forma, hemos calculado el coeficiente de correlación de Pearson para las variables "número de sesgos en los ítems de opción múltiple" y "número de respuestas correctas en los ítem de respuesta abierta" para las dos submuestras. Para los alumnos de la Licenciatura de Matemáticas tenemos un valor de r=0,022 (p=0,829); es decir existe una correlación casi nula entre ambas variables. Por tanto no están relacionadas linealmente el número de sesgos en los primeros ítems con el número de respuestas correctas en los restantes para el grupo de alumnos de la licenciatura. Para la submuestra de alumnos del Máster de secundaria tenemos un valor de r=-0,241 (p =0,015); es decir existe una correlación negativa y baja entre ambas variables y es significativa a un nivel de confianza del 95%. Por tanto están relacionadas linealmente el número de sesgos en los primeros ítems con el número de respuestas correctas en los restantes para el grupo de alumnos del Máster, pero como en el caso general es una relación muy débil.

El coeficiente de correlación de Pearson para las variables "número de respuestas correctas en los ítems de respuesta múltiple" y cada uno de los ítem de respuesta abierta, del ítem 8 al ítem 14 de la muestra conjunta da valores bajos del coeficiente de correlación. El test de hipótesis para todas las correlaciones da valores *p-valor* mayores que 0,05. Podemos concluir que no están relacionadas linealmente el número de respuestas correctas en los primeros ítems con cada uno de los ítems abiertos.

Análisis factorial

En segundo lugar, la relación entre sesgos en el razonamiento condicional y conocimiento formal se analiza a través del análisis factorial del conjunto de respuestas a todos los ítems y en el total de la muestra. Los ítems de opción múltiple se puntúan simplemente como 1=correcto, 0= incorrecto, mientras que los abiertos se puntúan de 0 a 2 o de 0 a 4 como se explicó en los apartados anteriores. El análisis se ha realizado con variables tipificadas, para que todas ellas aporten la misma contribución a la inercia total del conjunto de datos.

La técnica de análisis factorial fue utilizada por Díaz (2007) en la validación de constructo de su cuestionario. Según Muñiz (1994) el análisis factorial permite comprobar las hipótesis teóricas sobre la estructura del constructo. En nuestro caso, la principal hipótesis es la falta de relación entre los ítems relacionados con los sesgos y los relacionados con el conocimiento formal, por lo que esperamos que al realizar el análisis aparezcan en factores separados, al menos en los primeros factores que explican el mayor porcentaje de inercia.

Esperamos que los factores en nuestro análisis reproduzcan en parte los resultados de Díaz (2007), aunque es de esperar algunas diferencias: Por un lado, el tamaño de su muestra fue mucho mayor; por otro, la autora incluyó en el cuestionario varios ítems no incluidos en el estudio actual, que pueden interrelacionar con el resto e influenciar la matriz de correlaciones. Finalmente, la diferente preparación y conocimientos de los estudiantes hace razonable esperar alguna variación. Según Martínez-Arias (1995) la estructura factorial de las respuestas a un cuestionario permite, además, comprobar la dimensionalidad latente e interpretar los factores en términos de las dimensiones supuestas en el mismo.

Siguiendo el mismo procedimiento de Díaz (2007) para comparar mejor los resultados, la extracción de factores se llevó a cabo mediante el método de componentes principales; con objeto de obtener factores estadísticamente independientes y de máxima variabilidad, al tiempo que no se deforma la estructura de los datos (Cuadras, 1981). Este método parte de una estimación inicial más alta de las comunalidades (Martínez-Arias, 1995). Como método de rotación se usó la rotación Varimax, método ortogonal, que conserva las comunalidades y la suma de porcentajes de varianza explicados por los factores (Afifi y Clark, 1990). Está orientado a maximizar la varianza de los factores.

Antes de aplicar el método, se comprobó la unidad experimental de las variables (todos son ítems referidos al mismo constructo) y número de casos necesarios (se recomiendan al menos 10 por variable); en nuestro caso tenemos 14 variables y 196 sujetos, luego cumplimos esta condición (Cuadras, 1981).

Tabla 4.7.50. KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin		0,61
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	275,95
	gl.	105
	Sig.	0,00

Se contrastó la hipótesis de que los elementos de la matriz de correlaciones fuera de la diagonal principal son diferentes de cero, mediante la prueba de esfericidad de

Barlett (Tabla 4.7.50). Se obtuvo un valor altamente significativo, lo que indica la existencia de correlaciones suficientemente altas entre las variables para poder realizar el análisis. La prueba de Kaiser-Meyer-Olkin dio un valor mayor que 0,6 recomendado para poder llevar a cabo un análisis factorial con garantías. Todo ello garantizó la aplicabilidad del método. En la Tabla 4.7.51 se presentan las comunalidades obtenidas que oscilan entre 0,414 (ítem 10, probabilidad total) y 0,674 (falacia de las tasas base) lo que indica que cada ítem tiene una parte específica fuerte, en especial alguno de ellos (Afifi y Clark, 1990).

Tabla 4.7.51. Comunalidades mediante análisis de componentes principales

	Inicial	Extracción
1. Tasas base	1,000	0,674
2. Indep/exclusividad	1,000	0,606
3. Confundir probabilidades	1,000	0,681
4. Falacia conjunción	1,000	0,523
5. Condicional transpuesta	1,000	0,640
6. Eje temporal/canales	1,000	0,662
7a. Reposición	1,000	0,616
7b. Eje temporal	1,000	0,549
8. Definición	1,000	0,636
9. P. condicional	1,000	0,553
10. P. total	1,000	0,414
11. Independencia	1,000	0,633
12. P. compuesta, dep.	1,000	0,468
13. Bayes	1,000	0,419
14. P. compuesta, indep.	1,000	0,494

Tabla 4.7.52. Varianza total explicada mediante análisis de componentes principales

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	2,481	16,540	16,540	2,481	16,540	16,540	1,687	11,247	11,247
2	1,565	10,436	26,976	1,565	10,436	26,976	1,526	10,173	21,420
3	1,196	7,975	34,951	1,196	7,975	34,951	1,510	10,069	31,489
4	1,164	7,760	42,711	1,164	7,760	42,711	1,402	9,348	40,837
5	1,114	7,427	50,138	1,114	7,427	50,138	1,280	8,532	49,370
6	1,047	6,977	57,115	1,047	6,977	57,115	1,162	7,745	57,115
7	0,984	6,561	63,676						
8	0,921	6,140	69,816						
9	0,819	5,461	75,277						
10	0,754	5,028	80,305						
11	0,744	4,960	85,266						
12	0,632	4,216	89,481						
13	0,594	3,957	93,439						
14	0,572	3,813	97,252						
15	0,412	2,748	100,000						

Se usaron múltiples métodos para determinar cuántos factores extraer. La extracción inicial obtuvo 6 factores con autovalor mayor que 1, que explicaron el 57,1% de la varianza total (Tabla 4.7.52). Este número de factores es menor a la mitad del número inicial de variables menos 1.

El primer factor explica un 16,5% de la varianza, y el segundo un 10,4% mientras que los siguientes alrededor cada uno, lo que indica la importancia relativa del primer y segundo factor.

En el gráfico de sedimentación (scree plot) (Figura 4.7.5) se observa de nuevo que la mayor varianza es debida al primer y segundo factor.

Figura 4.7.5. Gráfico de sedimentación

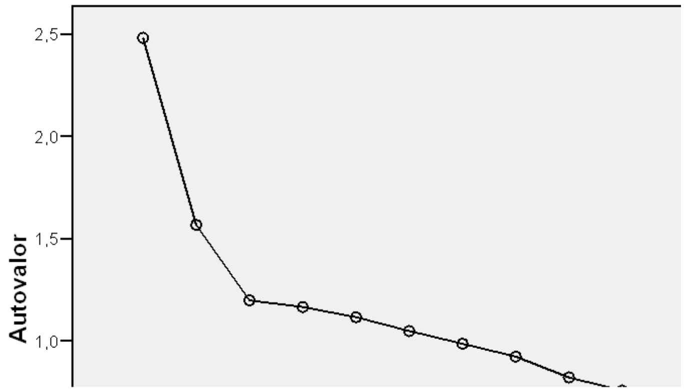


Tabla 4.7.53. Matriz de componentes no rotada mediante análisis de componentes principales

	Componente					
	1	2	3	4	5	6
13. Bayes	0,642					
12. P. compuesta, dep.	0,557	-0,361				
14. P. compuesta, indep.	0,517		-0,452			
10. P. total	0,464	-0,322				
4. Falacia conjunción	0,449	-0,240			0,328	0,378
7b. Eje temporal	0,416		0,310	0,416	-0,292	
1. Tasas base		0,605	0,363			
6. Eje temporal/canales	0,380	0,594		-0,371		
3. Confundir probabilidades		0,574	-0,335	0,386		
2. Indep/ exclusividad	0,395		0,490			-0,371
8. Definición	0,279		-0,252	0,498	-0,320	0,366
11. Independencia	0,450		-0,326	-0,468		
5. Condicional transpuesta				0,311	0,703	
9. P. condicional	0,296	-0,331	0,261			0,475
7a. Reposición	0,358		-0,359		0,402	-0,416

Incluimos la matriz no rotada de componentes (Tabla 4.7.53), donde las variables se presentan ordenadas según la importancia relativa de su contribución al primer factor.

Hemos suprimido los coeficientes factoriales menores de 0,25 para facilitar la interpretación. Incluso antes de la rotación se observa que diferentes variables contribuyen a diversos factores (altas y bajas correlaciones con ellos). Pero también se observa que la mayoría de los ítems contribuyen al primer factor y un número apreciable de ellos tiene un peso importante en el mismo, lo cual contribuye una nueva evidencia de existencia del constructo subyacente (Morales, 1988). No obstante, puesto que el método de extracción trata de maximizar la varianza de los primeros factores, puede considerarse que la matriz factorial no rotada contiene factores no puros, en el sentido de que cada factor explica su varianza y parte del siguiente, por lo que se aconseja hacer una rotación para facilitar la interpretación (Tabla 4.7.54). Hemos representado en color grisáceo los ítems relacionados con los sesgos para ayudar en la visualización. Se cumplen las siguientes condiciones (Afifi y Clark, 1990):

- Tomar sólo factores que sean interpretables para el investigador.
- Cada fila de la matriz rotada tenga al menos un cero; es decir para cada variable debe haber al menos un factor que no contribuya a su varianza.
- Para cada factor, habrá un conjunto de variables cuyas saturaciones se aproximen a cero.
- Cada factor debe tener peso importante de al menos dos variables, pues de otro modo sería un factor específico.

Los factores se interpretan a continuación.

Tabla 4.7.54. Método de rotación: Normalización Varimax con Kaiser

	Componente					
	1	2	3	4	5	6
11. Independencia	0,776					
14. P. compuesta, indep.	0,597					
12. P. compuesta, dep.	0,461		0,309	0,381		
13. Bayes	0,407		0,334			
1. Tasas base		0,786				
6. Eje temporal/canales		0,781				
2. Indep/ exclusividad		0,261	0,719			
7b. Eje temporal			0,700			
10. P. total	0,290		0,507			
9. P. condicional				0,732		
4. Falacia conjunción				0,683		
8. Definición					0,747	
3. Confundir probabilidades		0,298			0,716	
5. Condicional transpuesta						0,690
7a. Reposición	0,346					0,672

Factor 1. Resolución de problemas. El primer factor explica el 11,2% de la varianza. Este alto porcentaje de varianza explicado refleja la importancia del primer factor, al cual contribuyen la mayoría de los ítems de respuesta abierta. En particular, el problema relacionado con la independencia, que presenta la contribución mayor, seguido de los problemas de probabilidad compuesta y el teorema de Bayes. Hay una contribución menor del problema de cálculo de probabilidad condicional en situación de muestreo con reposición y el problema de probabilidad total. Todos estos problemas (excepto el 7) requieren un proceso de resolución de al menos dos pasos, de los que el primero es el cálculo de una probabilidad condicional, que se utiliza en los siguientes pasos (ej. regla del producto). Podemos interpretar este factor como la *habilidad de resolver problemas de probabilidad condicional complejos*. Estos resultados replican los de Díaz (2007) cuyo primer factor también incluyó estos ítems, aunque con diferente peso y orden de importancia.

Factor 2. Sesgos en el razonamiento condicional. El segundo factor explica el 10,2% de la varianza, lo cual es también una alta proporción. Incluye los ítems de la falacia de las tasas base, falacia del eje temporal, confusión entre independencia y mutua exclusividad y confusión entre diferentes probabilidades; los dos primeros con mucha mayor saturación. Por tanto, este conjunto de sesgos aparecen separados de los ítems relacionados con la resolución de problemas, lo que confirma nuestra hipótesis de que un estudiante con altos conocimientos matemáticos, todavía podría caer en los anteriores sesgos. En el trabajo de Díaz (2007) los ítems correspondientes a sesgos aparecieron en su mayoría en factores separados del primer factor, aunque la relación entre diferentes sesgos no se mostró tan clara como en nuestro estudio ocurre con este segundo factor.

Factor 3. Sesgos que afectan la resolución de algunos problemas. Los ítems que más contribuyen a este factor (que explica un 10% de inercia) son dos sesgos: La falacia del eje temporal y la confusión entre independencia y mutua exclusividad. Por otro lado, aparece una contribución moderada de algunos ítems de resolución de problemas: Probabilidad total, Bayes y probabilidad compuesta para caso de sucesos dependientes. Es claro que, tanto el problema de probabilidad total como de Bayes, requiere comprender bien las particiones sucesivas del espacio muestral y, en consecuencia, no confundir la mutua exclusividad de los conjuntos de la partición con independencia. Por otro lado, la correcta solución del problema de Bayes requiere superar la falacia del eje de tiempos. Es menos clara la relación con el caso de la probabilidad compuesta para sucesos dependientes, pero esta es necesaria para resolver tanto el problema de probabilidad total, como del de Bayes; posiblemente esto explica su presencia en este factor.

Factor 4. Falacia de la conjunción. (9,3% de varianza explicada) Los dos ítems que más contribuyen a este factor son la falacia de la conjunción y el cálculo de la probabilidad condicional a partir de la definición; este factor reproduce casi exactamente el quinto factor en el trabajo de Díaz; pero en nuestro caso aparece una relación con el cálculo de la probabilidad compuesta, aunque con poco peso.

Factor 5. Definición de la probabilidad condicional. (8,5% de varianza explicada) Únicamente dos ítems contribuyen a este factor: la definición de la probabilidad condicional y la confusión entre diferentes probabilidades. Recordemos que en el ítem de definición se preguntaba a los participantes por la diferencia entre probabilidad simple y compuesta, por lo que es razonable que estos dos ítems aparezcan asociados. Por otro lado, al ser los únicos en este factor indica la especificidad de la competencia necesaria para dar una definición correcta, no relacionada ni con la capacidad de resolver problemas ni con la superación de los sesgos de razonamiento. Una explicación es la baja frecuencia de alumnos con una definición correcta en nuestra muestra. En el trabajo de Díaz, la definición también apareció separada de la resolución de problemas y de los ítems relacionados con sesgos, aunque en su caso se relacionaba con algunos pocos ítems de resolución de problemas (los más sencillos, algunos de los cuáles no han sido incluidos en este estudio) y con un problema de enumeración del espacio muestral tampoco incluido en este caso.

Factor 6. Falacia de la condicional transpuesta. (7,7% de varianza explicada) El último factor sólo agrupa dos ítems ambos con alta saturación. La falacia de la condicional transpuesta sólo aparece en este, por tanto hemos dado su nombre a este factor. Se asocia al ítem 7a, un ítem sencillo de cálculo de probabilidad condicional en muestreo con reposición. Ambos ítems fueron sencillos, y esto explica la posible asociación. También en este caso, este factor reproduce exactamente el factor 6 en el estudio de Díaz (2007).

4.8. CONCLUSIONES DEL ESTUDIO 2

El Estudio 2, que se ha presentado en este capítulo, estuvo orientado *a evaluar la presencia de los sesgos más comunes sobre el razonamiento en probabilidad condicional*, que se describieron en el capítulo 2, en una muestra de futuros profesores de Educación Secundaria. Al mismo tiempo *se deseaba evaluar su conocimiento común de dicho contenido*. Un segundo objetivo fue *comparar los resultados de la evaluación en futuros profesores de secundaria, con los obtenidos por Díaz (2007),* con el mismo cuestionario en estudiantes de Psicología. Como tercer objetivo del estudio 2 *se deseaba evaluar la competencia en resolución de problemas de probabilidad condicional por los futuros profesores y su interrelación con los sesgos* anteriormente descritos.

Las hipótesis que planteamos respecto a los anteriores objetivos fueron las siguientes:

- *Esperábamos, en primer lugar, encontrar una proporción alta de futuros profesores en la muestra que presenten los sesgos evaluados en el estudio*.
- Por el contrario*, esperábamos que los resultados en los ítems relacionados con la resolución de problemas en nuestro estudio fuesen mejores que los encontrados por Díaz (2007),* debido a la mejor preparación en estadística y probabilidad de los participantes en nuestro estudio.
- Finalmente, *la tercera hipótesis de este estudio era la falta de relación entre la competencia matemática de los estudiantes para resolver problemas de probabilidad condicional y los sesgos observados en su razonamiento probabilístico*.

Para conseguir los objetivos logrados y analizar las hipótesis planteadas, a lo largo del capítulo se ha llevado a cabo un estudio pormenorizado de las respuestas de la muestra de futuros profesores a una parte de los ítems del cuestionario de Díaz (2007). Se ha seguido el método de análisis de la citada autora, aunque en los ítems de respuesta abierta se ha ampliado su categorización de respuestas. Se han comparado nuestros resultados con los de Díaz y además se han comparado los resultados en las dos submuestras utilizadas en el estudio. En lo que sigue resumimos nuestras conclusiones sobre los citados objetivos e hipótesis.

Sesgos en el razonamiento sobre probabilidad condicional

Al analizar el porcentaje de futuros profesores que incurre en las diferentes falacias o sesgos, hemos mostrado que dicho porcentaje es alto. Tanto en nuestra muestra, como en la de Díaz, se ha confirmado nuestra primera hipótesis, pues hay una fuerte presencia de los sesgos analizados.

Destaca por su incidencia la *falacia del eje de tiempos,* que se dio en alrededor de dos de cada tres participantes en nuestro estudio en el ítem 6 y uno de cada tres en el ítem 7b. En menor medida, aunque también con alto porcentaje, aparece la *falacia de las tasas base*, que en nuestra muestra se dio en 39,8%, pero que en la muestra de Díaz sólo alcanzó el 15%. El siguiente sesgo por orden de frecuencia fue la *confusión entre probabilidad conjunta y condicional,* con uno de cada tres encuestados. El resto de sesgos, *falacia de la conjunción* y de la *condicional transpuesta,* se dieron en un porcentaje menor, alrededor del 10% en nuestra muestra y en menor porcentaje en la de Díaz.

Deducimos que la formación matemática recibida por los futuros profesores no es suficiente para hacerles conscientes de la existencia de estos sesgos y que, por lo tanto, los podrían transmitir a sus estudiantes. Aunque los ítems propuestos estaban diseñados para mostrar distintas falacias y sesgos comunes en la población, esto no justifica su presencia tan extendida en la muestra, pues hay que tener en cuenta que muchos de estos problemas ocurren fácilmente en la vida cotidiana o profesional. Por otro lado, los resultados en algunos de los ítems de opción múltiple han sido peores que los obtenidos por Díaz (2007) en estudiantes de Psicología, a pesar de la mayor preparación matemática de los alumnos de nuestra muestra.

Para comprobar si existe relación entre la frecuencia de aparición de los diferentes sesgos estudiados y el tipo de estudiante, hemos realizado un test Chi-cuadrado de contraste de hipótesis, obteniendo resultados estadísticamente muy significativos. En consecuencia, los diferentes sesgos encontrados en cada ítem tienen relación, con la muestra de estudiantes, siendo mayor la falacia de las tasas base, confusión entre sucesos independientes y mutuamente excluyentes en los futuros profesores de Matemáticas que estudiantes de Psicología, mientras que la falacia del eje de tiempo es ligeramente mayor en el segundo grupo. El resto de sesgos tiene poca incidencia y es similar en ambos grupos.

Al comparar los estudiantes del Máster con los de la Licenciatura de Matemáticas, hemos realizado un test Chi-cuadrado, con resultados no estadísticamente significativos. En consecuencia, no observamos diferencias en la presencia de los diferentes sesgos en los dos grupos que componen nuestra muestra. Será por tanto necesario que los formadores de profesores, tanto en la Licenciatura de Matemáticas, como en el Máster tengan en cuenta estos sesgos y se organicen actividades formativas que los ayuden a superarlos, ya que, como indican Serradó, Azcárate y Cardeñoso (2006), estos segos podrían incidir en la confianza que muestren en la enseñanza de este tema en el futuro.

Resolución de problemas

Al contrario de lo encontrado con los ítems que miden los sesgos, en los ítems abiertos, el porcentaje de alumnos que cometen errores es mucho menor que los de la muestra del estudio de Díaz. Por tanto, también se confirma nuestra hipótesis al respecto. Solamente en el caso del ítem relacionado con el teorema de Bayes se dio un porcentaje de errores mayor en la muestra de nuestro estudio que en la de Psicología, siendo el porcentaje muy bajo. Por otro lado, en los dos estudios coinciden los resultados en el ítem 9, relacionado con el cálculo de la probabilidad condicional, que fue un ítem con dificultades por parte de ambas muestras.

Destacamos que nuestros análisis de las respuestas en estos ítems han permitido extender las categorías cuantitativas propuestas por Díaz (2007) para puntuar la mayor o menor corrección de la respuestas. Este análisis pormenorizado nos ha servido para detectar diferentes sesgos y conflictos semióticos que también se translucen en las respuestas abiertas, con lo que se confirma la presencia de sesgos detectada en los ítems de opciones múltiples:

- La *confusión entre condicionamiento y causalidad* se presenta en un 32% de las respuestas del ítem 8 (definición), donde los alumnos expresan la idea de que el suceso condicionado ha de depender del condicionante.
- La *falacia del eje temporal* se presenta en un 17,3% de las respuestas la ítem 8, dando a entender en la definición que el suceso condicionante ha de ser previo al condicionado.
- Un 16 % de estudiantes tienen *dificultad en determinar el espacio muestral* en un problema sencillo consistente en imponer una condición a la suma de dos dados y otro 5% no diferencian el orden de los dados en este contexto (ítem 9).
- La *falacia del jugador* se presenta en un 6,1% de estudiantes en el ítem 11

(independencia) al estimar la probabilidad de que se produzca un nuevo resultado en función de los resultados anteriores.

- Unos pocos alumnos en el ítem 11 *confunden los sucesos independientes con sucesos excluyentes,* error señalado por Sánchez (1996) y Kelly y Zwiers (1986).

Relación entre sesgos y conocimiento formal

Los resultados sobre este punto confirman nuestra hipótesis sobre la escasa relación entre sesgos en el razonamiento condicional y conocimiento formal del tema, apoyando de este modo los resultados de Díaz (2007) con estudiantes de Psicología.

Por un lado, al correlacionar el número total de respuestas correctas en los ítems de opción múltiple (que miden diferentes tipos de sesgos) con la puntuación obtenida en los ítems abiertos y otros indicadores del conocimiento formal del tema, se obtienen correlaciones muy pequeñas. Aunque en algún caso hayan sido estadísticamente significativas, no se traduce este resultado en una significación práctica, pues la baja correlación no permite predecir, en función del conocimiento del tema, los sesgos que cometerá el futuro profesor.

Por otro lado, el análisis factorial realizado del conjunto de respuestas a todos los ítems reproduce de una manera bastante fiel los resultados obtenidos por Díaz (2007). Hay alguna variación en la composición de los diferentes factores, pues, por un lado, se suprimieron varios ítems de nuestro estudio, por lo cual la correlación de dichos ítems influyó en la estructura factorial del análisis de Díaz, pero no en el nuestro.

Por otro, nuestra muestra, aunque importante y válida para el análisis, no es tan amplia como la de la citada autora, cuyos resultados, por dicho motivo serán más fiables. Finalmente, al cambiar la preparación previa del sujeto que respondió al cuestionario, es de esperar variaciones. Aun así, los resultados permiten corroborar la falta de relación entre ítems de uno y otro tipo de una forma general, aunque las conclusiones respecto a la interpretación de cada factor se detallaron en el análisis. Estos resultados también contribuyen a reforzar, con una nueva muestra, la validez de constructo del cuestionario RPC, que ya fue comprobada por la autora durante su construcción.

En resumen, los resultados del Estudio 2, muestran por un lado, un conocimiento formal general bueno de la probabilidad condicional y capacidad para resolver problemas relacionados en los futuros profesores, unido a una fuerte presencia de sesgos de razonamiento relacionados con la probabilidad condicional, que no depende de la alta preparación matemática.

La consecuencia práctica de este resultado es que si centramos la preparación de los futuros profesores únicamente en el componente formal del conocimiento de la probabilidad condicional, no podremos asegurar que estén libres de sesgos, si no se les confronta con los mismos. Como consecuencia, al no ser conscientes de estos sesgos, no serán capaces de diagnosticar estos errores en sus estudiantes cuando los presenten, pues ellos mismos los comparten.

Será necesario organizar situaciones didácticas para los profesores que les permita concienciarse de sus propios sesgos y les ayuden a superarlos. Para contribuir a lograr este nuevo objetivo, en el capítulo 5 se analizan algunos recursos que podrían ser útiles para diseñar este tipo de actividades (Estudios 3 y 4) y en el capítulo 6 se estudiará una experiencia formativa para profesores basada en uno de estos recursos (Estudio 5). De este modo, con el conjunto de estudios que componen esta información completaremos el ciclo evaluación, búsqueda de recursos formativos, diseño de actividades formativas, evaluación de su efectividad.

CAPÍTULO 5.
ANALISIS DE RECURSOS PARA LA FORMACIÓN DE PROFESORES

5.1. INTRODUCCIÓN

En los estudios 1 y 2 se puso de manifiesto las necesidades formativas de los futuros profesores sobre algunos contenidos matemáticos relacionados con la probabilidad condicional. Además, en el capítulo 2, se sugiere que los conocimientos estadísticos en sí mismos no son suficientes para que los docentes puedan enseñar probabilidad de una manera efectiva y desarrollar en sus estudiantes un adecuado razonamiento probabilístico. Algunos profesores pudieran desconocer la metodología propuesta en los nuevos currículos (basada en experimentos y simulaciones) o no ser conscientes de algunas de las dificultades y sesgos probabilísticos de sus alumnos (Stohl, 2005). Si su formación inicial se centró en las competencias matemáticas, y recibieron una enseñanza de la probabilidad muy clásica, basada en la axiomática de Kolomogorov y en el razonamiento deductivo, éstos profesores pueden sentirse inseguros con enfoques más informales, como los sugeridos hoy día en el currículo de la educación primaria y secundaria (Serradó, Azcárate y Cardeñoso, 2006).

Una consecuencia del análisis realizado en los Estudios 1 y 2 y en el capítulo 2 es la necesidad de desarrollar y evaluar los conocimientos de los docentes teniendo en cuenta los diferentes componentes que se describieron en dicho capítulo para el conocimiento del profesor. Es importante proporcionar al profesor actividades que les sirvan para conectar los aspectos conceptuales y didácticos de cada tema matemático (Ball, 2000). Las actividades presentadas a los profesores también deben basarse en el enfoque constructivista y social del aprendizaje (Jaworski, 2001).

Los Estudios 3 y 4 presentados en este capítulo están dirigidos a analizar dos tipos de recursos en los que podamos basarnos para llevar a cabo esta formación: recursos en Internet (Estudio 3) y paradojas clásicas en la historia de la probabilidad (Estudio 4). A continuación se describen cada uno de estos estudios.

5.2. RECURSOS PARA LA ENSEÑANZA DE LA PROBABILIDAD CONDICIONAL EN INTERNET

5.2.1. INTRODUCCIÓN

La estadística es una de las materias que ha tenido mayor influencia de la tecnología, y en particular de Internet (Galmacci, 2001), como se observa en la página de la International Association for Statistical Education, (www.stat.auckland.ac.nz/~iase/) que recoge vínculos a otros servidores y publicaciones, incluyendo tesis doctorales, actas de conferencias y revistas electrónicas, como *Statistics Education Research Journal.*

Entre otras posibilidades, los libros de texto se empiezan a transformar en ediciones electrónicas, libremente accesibles a la consulta, modificación y sugerencias a través de Internet. Es también sencillo obtener datos de todo tipo para que los estudiantes puedan realizar investigaciones sobre casi cualquier tema. También pueden combinar diferentes conjuntos de datos en un mismo proyecto o “enviar” a la red sus propias colecciones de datos para que sean usadas por nuevos estudiantes en cualquier rincón del planeta. Las listas de discusión entre profesores o entre alumnos, la "tutoría" de alumnos a distancia, cuando no es posible la comunicación directa con el profesor, ya es habitual en escuelas y universidades. A continuación se describen los objetivos e hipótesis del Estudio 3, orientado a la evaluación de la utilidad de algunos de estos recursos en la formación del profesorado.

5.2.2. OBJETIVOS E HIPÓTESIS DEL ESTUDIO 3

El objetivo general del Estudio 3 es localizar, clasificar y analizar algunos recursos didácticos en Internet que sean útiles en el estudio de la probabilidad condicional e independencia, en relación a las carencias mostradas en los Estudios 1 y 2, tanto respecto al conocimiento matemático, como respecto al conocimiento didáctico, en las diferentes facetas consideradas en este trabajo. Más concretamente, perseguimos dos objetivos específicos:

1. El primer objetivo es *seleccionar algunos recursos relacionados con la probabilidad condicional e independencia y realizar una clasificación de los mismos*. Han de tratarse de recursos interesantes para los profesores que complementen los libros de texto. Deben requerir conocimientos de probabilidad condicional o independencia. Al mismo tiempo han de tener potencial para organizar un debate en torno a las cuestiones matemáticas y didácticas, que contribuyan a incrementar el conocimiento del profesor.
2. El segundo objetivo es *realizar un análisis de los objetos matemáticos puestos en juego en el uso didáctico* de algunos de los recursos seleccionados e *identificar algunos posibles conflictos semióticos* que puedan surgir durante su solución.
3. El tercer objetivo es *valorar los procesos matemáticos utilizados en el trabajo con los recursos* seleccionados y la *idoneidad didáctica de los mismos* en los cursos de formación de profesores.

Nuestra hipótesis sobre este estudio es que *podremos identificar una variedad de recursos que muestren la riqueza de objetos matemáticos relacionados con la probabilidad condicional.* Asimismo *pensamos mostrar la idoneidad didáctica de estos recursos* para la formación de profesores en el campo de la probabilidad condicional.

La hipótesis se fundamenta, tanto en el análisis de significado de referencia de la probabilidad condicional, llevado a cabo en la sección 1.4, como en nuestro marco teórico, descrito en la sección 1.3, el cuál específica la complejidad del conocimiento matemático. Las practicas matemáticas necesarias para resolver problemas relacionados con el uso de estos recursos, permitirán, de acuerdo a Godino, Batanero y Font (2007), hacer aflorar *configuraciones*, objetos intervinientes y emergentes, de los sistemas de prácticas, en las categorías descritas por los autores. Nos basamos, asimismo, en el estudio previo de recursos en Internet presentado en Contreras (2009) y Contreras, Díaz y Batanero (2009), donde se mostró la alta idoneidad afectiva de los recursos, así como la idoneidad adecuada en el resto de los componentes de este concepto.

5.2.3. MATERIAL Y MÉTODO

La localización de los recursos analizados en este capítulo se ha realizado a través de varios sistemas:

- Explorando algunos servidores de educación estadística, que incluyen listados de recursos en Internet. Hemos explorado los servidores de "Biblioteca virtual de recursos manipulativos" (nlvm.usu.edu/es/nav/vlibrary.html), la plataforma "Mundo matemático" (www.planetamatematico.com/); NCTM (www.nctm.org/), IASE (www.stat.auckland.ac.nz/~iase/) "Proyecto Descartes" del Ministerio de Educación (descartes.cnice.mec.es/) y el servidor de nuestro grupo de investigación (www.ugr.es/~batanero/).
- A partir de algunos artículos que describen recursos en Internet, como los trabajos de Mills (2002) y de Díaz y de la Fuente (2005).
- Mediante búsqueda directa en buscadores de Internet, utilizando palabras claves como "Applet" y alguna de las siguientes "probabilidad condicional" "probabilidad condicionada", "independencia", "conditional probability", "Bayes", etc.

Una vez localizado un recurso, se clasificó en algunos de los apartados que se describen en este capítulo. Localizados los principales recursos dentro de cada una de las categorías, se eligió uno interesante desde el punto de vista didáctico. En este capítulo llevamos a cabo un análisis completo de dichos ejemplos. Asimismo, se seleccionaron otros dos recursos de interés para hacer una descripción resumida y el resto simplemente se listó en una tabla. No pretendemos ser exhaustivos, sino sólo mostrar algunos recursos útiles para nuestros objetivos.

El análisis del ejemplo elegido se realiza desde diferentes puntos de vista. En primer lugar, describimos el recurso, proporcionando datos de su ubicación en Internet y su autor, así como los objetivos que se persiguen con él, en el caso de que se hayan explicitado o puedan deducirse de su análisis. En segundo lugar se lleva a cabo un análisis matemático de las posibles soluciones correctas (en caso de que se trate de un problema o un juego cuya estrategia se justifique mediante la solución de un problema). Si se trata de un recurso de exploración, analizamos los objetos matemáticos en juego. En algún caso se describe la literatura o historia relacionada y las variantes del mismo encontradas en otros sitios de Internet.

De acuerdo a Godino, Font y Wilhelmi (2008) una tarea importante para el profesor es valorar su práctica docente con la finalidad de favorecer el aprendizaje de los estudiantes. Para facilitar ésta valoración, los autores describen diversos niveles de análisis, algunos de los cuáles creemos que también pueden aplicarse al estudio de los recursos didácticos, entre ellos los recursos en Internet. En concreto, para la descripción de los recursos realizaremos dos tipos de análisis:

- Los sistemas de prácticas y objetos matemáticos implícitos en el trabajo con el recurso. Según los autores, el análisis de objetos ligados a las prácticas matemáticas se aplica durante la planificación de la enseñanza. Permite descomponer el proceso de estudio previsto en una secuencia de episodios y, para cada uno de ellos, describir las prácticas supuestas. En nuestro caso consideramos un único episodio. Como consecuencia, trataremos de identificar las configuraciones y objetos latentes en el trabajo de los estudiantes al resolver problemas con los recursos analizados.
- Los procesos matemáticos requeridos y posibles conflictos semióticos de los estudiantes en el uso del recurso, describiendo con detalle las que se hayan identificado en las investigaciones previas, como por ejemplo en Díaz (2005) y Díaz y de la Fuente (2005).

5.2.4. ANÁLISIS DE RESULTADOS

5.2.4.1. JUEGOS

En esta categoría se incluyen diversos juegos encontrados en Internet, donde interviene la probabilidad condicional. Su utilidad en la enseñanza es que hacen aflorar los errores y las concepciones erróneas de los estudiantes. En la formación de profesores, estos juegos permiten contextualizar la reflexión epistemológica sobre la probabilidad condicional y otras ideas estocásticas fundamentales, así como analizar las posibles dificultades y obstáculos de los alumnos en el tema (Godino, Batanero y Flores, 1999).

El problema de Monty Hall

En primer lugar analizamos un recurso (cuya pantalla principal se muestra en la Figura 5.2.1) que simula el problema conocido como “Problema de Monty Hall”, y que incluye también algunas actividades en relación al juego.

Dirección en Internet: www.ugr.es/~jmcontreras/1

Figura 5.2.1. Pantalla principal del Juego 1

Descripción

Este juego lo encontramos dentro de los Applets incluidos en el “National Library of Virtual Manipulatives” (nlvm.usu.edu/) en relación con los temas de tratamiento de datos y probabilidad. Permite experimentar una de las versiones del Problema de Monty Hall. Está inspirado en el concurso televisivo Let's Make a Deal (Hagamos un trato), emitido entre 1963 y 1986 en la televisión americana y su nombre proviene del presentador del concurso, Monty Hall. El concurso generó bastante polémica en relación a las posibles soluciones del problema matemático latente y muestra las intuiciones incorrectas en relación a la probabilidad condicional. La formulación más conocida de dicho problema proviene de una carta a la columna de Marilyn vos Savant en *Parade Magazine* (Bohl, Liberatore, y Nydick, 1995) y que reproducimos a continuación.

Supón que estás en un concurso, y se te ofrece escoger entre tres puertas: Detrás de una de ellas hay un coche, y detrás de las otras, cabras. Escoges una puerta, digamos la nº1, y el presentador, que sabe lo que hay detrás de las puertas, abre otra, digamos la nº3, que contiene una cabra. Entonces te pregunta: "¿No prefieres escoger la nº2?". ¿Es mejor para ti cambiar tu elección?

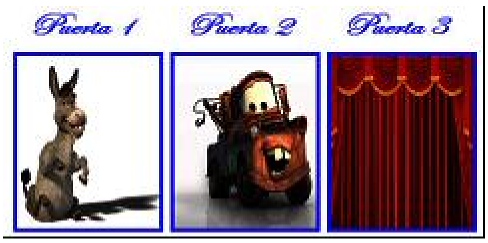
El problema original fue planteado por Selvin (1975), quien hace la primera mención del término “problema de Monty Hall”. Un problema análogo denominado “problema de los tres prisioneros”, fue publicado por Gardner (1959a), aunque su versión hace el proceso de elección explícito, evitando las suposiciones de la versión original. En lo que sigue, primero analizamos las posibles soluciones correctas al problema de búsqueda de la mejor estrategia en el juego y, a continuación, realizamos los dos tipos de análisis indicados.

Solución matemática del juego

Cuando se trabaja con el problema de Monty Hall en un curso de probabilidad, podemos hacer a los estudiantes alguna pregunta del tipo: *¿Debe el concursante mantener su elección original o escoger la otra puerta? ¿Hay alguna diferencia entre cambiar o no?* Les pediremos además que justifiquen su decisión.

Para simplificar el problema, asumimos que solamente hay dos tipos de jugador, los que nunca cambian de puerta y los que cambian siempre. La pregunta se limita a ver qué tipo de jugador tiene la mayor probabilidad de ganar el coche. En caso de que los estudiantes no logren dar la solución o den una solución errónea (lo cual es lo más frecuente), se puede dar la oportunidad de simular el juego usando el recurso (ver Figura 5.2.2) y obtener datos experimentales que les ayuden a intuir (y posteriormente demostrar) la solución correcta.

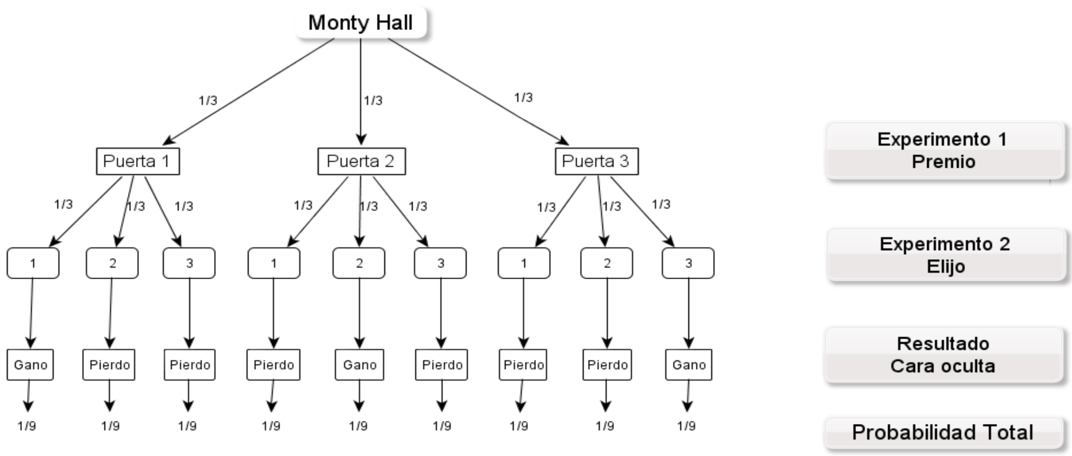
Figura 5.2.2. Resultado de una simulación



Solución intuitiva 1

La solución correcta se basa en tres suposiciones básicas: (a) Que el presentador siempre abre una puerta; (b) que la escoge después de que el concursante escoja la suya, y (c) que tras ella siempre hay una cabra. Un diagrama en árbol (Figura 5.2.3) visualiza las distintas posibilidades que podemos encontrarnos. Hay dos puertas sin premio y una con premio; por tanto la posibilidad de elegir la puerta premiada es 1/3. Si elegida una puerta no se cambia, sólo hay 1/3 de posibilidades de ganar y 2/3 de perder. Si, por el contrario, se cambia de puerta, la probabilidad de ganar será la misma que elegir inicialmente la puerta sin premio, es decir 2/3.

Figura 5.2.3. Espacio muestral cuando no se cambia de puerta

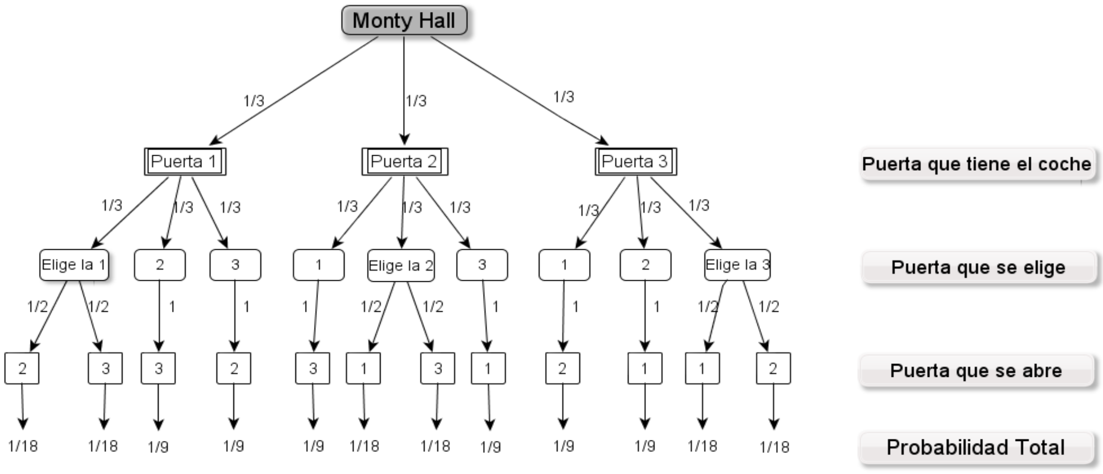


Solución intuitiva 2

Otro razonamiento es el siguiente: Consideramos, en primer lugar, el experimento “puerta que tiene el premio” (cada puerta tiene una probabilidad de tener el premio de 1/3). A continuación, consideramos la puerta que se elige (1/3 cada puerta). Estos dos primeros experimentos son independientes. El tercer experimento es la puerta que abre el locutor, que es dependiente de los anteriores, como se muestra en el diagrama en árbol (Figura 5.2.4). Si cambiamos de puerta, las posibilidades de ganar son de 1/3, sumando las probabilidades de todas las ramas del árbol. Sin embargo si cambiamos de puerta tenemos una probabilidad de 2/3 porque:

- Si escogemos una puerta con una cabra, entonces el presentador muestra la otra cabra. Nosotros cambiamos (a la puerta que tiene el coche) y ganamos;
- Si escogemos puerta con coche, entonces, el presentador muestra la otra cabra, cambiamos (a la puerta con la segunda cabra) y perdemos.

Figura 5.2.4. Espacio muestral en el experimento compuesto



Solución experimental

El trabajo de los alumnos con el Applet, experimentando con el juego y reflexionando sobre los resultados, proporciona una experiencia intuitiva sobre los resultados que se obtienen en este juego con cada una de las dos estrategias (cambiar o no cambiar de puerta). Partiendo de la evidencia de estos resultados (claramente se observa experimentalmente que las posibilidades de ganar el juego son el doble al cambiar la puerta), el alumno ve sus intuiciones contradichas. Es decir, se produce un conflicto cognitivo y al tratar de resolverlo, eventualmente puede llegar a uno de los razonamientos intuitivos mostrados anteriormente. En el transcurso de la experimentación con el Applet, debemos elegir una de las tres puertas al azar (este es el primer experimento aleatorio). La experiencia con el juego muestra al alumno la siguiente secuencia:

- Una vez hemos elegido una de las puertas, el programa nos abre una puerta donde no hay dicho premio.
- A continuación, el programa nos da la opción de quedarnos con la puerta elegida o cambiar a la que queda sin descubrir.

Cuando realizamos el juego, el Applet nos proporciona una tabla similar a la Tabla 5.2.1. A partir de los resultados proporcionados en esta tabla podemos comparar la frecuencia relativa de ganar en cada una de las dos estrategias e intuir qué estrategia es mejor. Teniendo en cuenta que los resultados son aleatorios, deberíamos realizar el juego un número de veces considerable para que los resultados se ajusten a la solución del problema, pero el ordenador permite un gran número de simulaciones rápidamente.

Tabla 5.2.1. Tabla de datos proporcionada por el Applet

% de ganar cuando no cambias (número de aciertos de las veces que no cambiamos).
% de ganar cuando cambias (número de aciertos de las veces que cambiamos).

En conclusión, el Applet nos proporciona una solución experimental de cuál es la estrategia ganadora. Pero no nos explica la razón de por qué una estrategia es preferible a la otra. Será necesario que el profesor trate de reconducir al estudiante a una de las soluciones intuitivas 1 o 2, para que comprenda el comportamiento del juego.

Solución formal 1

La solución formal de este problema utiliza las propiedades de la probabilidad condicionada, que es un objeto cuya definición es sencilla de entender pero difícil de aplicar. Para llegar a la solución definimos los siguientes sucesos:

- *A:* El jugador selecciona la puerta que contiene el coche en su selección inicial.
- *B:* El jugador selecciona una puerta que contiene una cabra en su selección inicial.
- *G:* El jugador gana el coche.

Para calcular $P(G)$, basta con notar que $G=(G \cap A) \cup (G \cap B)$, ya que $A \cap B = \emptyset$ y $A \cup B = \Omega$. Esto equivale a decir que $\{A, B\}$ es una partición de Ω, siendo Ω el espacio muestral del experimento; por tanto, aplicando el axioma de la unión de probabilidades:

$$P(G)=P((G \cap A) \cup (G \cap B)) = P(G \cap A) + P(G \cap B) = P(G/A)P(A) + P(G/B)P(B)$$

En cualquier caso, dado que no tenemos ninguna razón para pensar lo contrario, podemos aplicar el principio de indiferencia, suponiendo que las puertas son indistinguibles y diremos que $P(A)=1/3$ y $P(B)=2/3$ pues hay un coche y dos cabras. Para ello, aplicamos simplemente la regla de Laplace. Ahora debemos definir qué tipo de jugador estamos estudiando:

- *Jugador que nunca se cambia:* En este caso $P(G/A)=1$ y $P(G/B)=0$ pues el jugador se queda con su selección inicial. Por lo tanto $P(G)=1/3$.
- *Jugador que siempre se cambia:* En este caso $P(G/A)=0$ y $P(G/B)=1$ pues el jugador se cambia a la única puerta cerrada que queda (y sabemos que como el presentador sabe dónde está el coche, siempre mostrará una cabra). Por lo tanto $P(G)=2/3$.

En resumen, si mantiene su elección original gana si escogió originalmente el coche (con probabilidad de 1/3), mientras que si cambia, gana si escogió originalmente una de las dos cabras (con probabilidad de 2/3). Por lo tanto, el concursante debe cambiar su elección si quiere maximizar la probabilidad de ganar el coche.

Solución formal 2

Sea ξ: $(\Omega, P) \Rightarrow \{1,2,3\}$ la variable aleatoria que asigna un número de puerta (aquella detrás de la cual se encuentra el coche). Esta variable aleatoria tiene distribución discreta uniforme (es decir todos los valores son equiprobables) y son estocásticamente independientes. Sea ϕ: $(\Omega'', P'') \Rightarrow \{\eta\}$ la variable aleatoria número de la puerta que abre el moderador y que dependerá de las anteriores.

- Si $\eta=\xi$ (el concursante elige el coche), entonces hay dos posibles valores con probabilidad 1/2 (los números de las dos puertas no elegidas por el concursante).
- En caso contrario, sólo hay un valor, con probabilidad 1 (el número de la puerta sin coche). La probabilidad que el candidato se lleve el coche bajo el supuesto que él no cambia de puerta es entonces $P(\eta=\xi)=1/3$. La probabilidad que el candidato se lleve el coche bajo el supuesto que él cambia de puerta es entonces $P(\eta\neq\xi)=2/3$.

Objetos matemáticos puestos en juego

Al resolver matemáticamente el juego mediante alguna de las soluciones anteriores se utilizan implícitamente los objetos matemáticos que se muestran en la Tabla 5.2.2. Observamos que, dependiendo de la solución, se pueden usar una configuración diferente de objetos matemáticos, siendo más complejas las soluciones formales, especialmente la segunda que involucra la idea de variable aleatoria. Por tanto, el recurso que hemos construido en sí mismo no determina el trabajo matemático que se hace, sino que éste viene determinado por el tipo de solución obtenida. Ello hace que con este Applet se pueda trabajar a diversos niveles de profundidad, dependiendo del tipo de estudiante.

Tabla 5.2.2. Configuraciones epistémicas en las soluciones correctas

Tipo	Objetos matemáticos	Significado en la situación	Int. 1	Int. 2	Exp.	Form. 1	Form. 2
Problema	- Elección de la puerta	- Determinar la mejor estrategia	X	X	X	X	X
Lenguajes	- Verbal	- Explicación de la situación	X	X	X	X	X
	- Gráfico	- Diagrama en árbol - Representación icónica del juego	X	X			
	- Simbólico	- Expresar sucesos, probabilidades				X	X
	- Numérico: Porcentajes	- Probabilidades en porcentaje			X		
	- Numérico: Frecuencias	- Resultados del experimento			X		
	- Icónico	- Iconos que representan los sucesos y resultados			X		
Conceptos	- Experimento aleatorio	- Elegir una puerta - Puerta que abre el locutor - Ganar el premio	X X X	X X X	X X X	X X X	X X X
	- Sucesos; espacio muestral	- Puertas 1, 2, 3 - Ganar/no ganar	X X	X X	X X	X X	X X
	- Experimento compuesto	- Composición de los experimentos anteriores	X	X	X	X	X
	- Sucesos en el experimento compuesto	- Producto cartesiano de los espacios muestrales anteriores	X	X	X	X	X
	- Frecuencia relativa	- Éxitos / número experimentos			X		
	- Convergencia	- Tendencia de la frecuencia a un valor fijo			X		
	- Intersección de sucesos	- Conjunto común de sucesos.				X	
	- Unión de sucesos	- Sucesos en uno u otro conjunto				X	
	- Suceso imposible	- Intersección de un suceso y su complementario				X	
	- Probabilidad clásica	- Proporción de casos favorables a posibles.	X	X		X	X
	- Prob. frecuencial	- Límite de la frecuencia			X		
	- Axiomas probabilidad	- Explicitación de los axiomas				X	X
	- Probabilidad condicional	- Proporción de ocurrencia suceso respecto a otro	X	X	X	X	X
	- Regla de la suma	- Probabilidad de ganar	X	X		X	X
	- Regla del producto	- Probabilidad conjunta; dependencia					X
	- Variable aleatoria	- Número de puerta con premio - Número de puerta elegida					X
	- Igualdad de v. aleatorias	- Coincidencia de valores; acierto					X
	- Distribución de probabilidad	- Conjunto de valores con sus probabilidades					X
	- Distribución discreta uniforme	- Conjunto finito de valores equiprobables					X
	- Variables aleatorias independientes	- La distribución de una no depende de la de la otra					X
Procedimientos	- Cálculo probabilístico intuitivo	- Aplicar reglas intuitivas	X	X			

	- Cálculo probabilístico formal	- Aplicar reglas de cálculo formal				X	X
	- Cálculo de probabilidad frecuencial	- Estimar la probabilidad mediante la frecuencia			X		
	- Representación gráfica	- Diagrama; esquema	X	X			
Propiedades	- Diferencia probabilidad condicionada y simple	- Restricción del espacio muestral	X	X		X	X
	- La frecuencia converge a la probabilidad	- Ley empírica de los grandes números			X		
	- Teorema Prob. total	- Aplicar a la situación				X	
	- Axioma de unión	- La probabilidad de la suma es suma de probabilidades	X	X		X	X
Argumentos	- Razonamiento deductivo	- Demostración de la solución	X	X		X	X
	- Razonamiento empírico	- Comparar aciertos con distintas estrategias			X		

Dificultades posibles de los estudiantes

Aunque el problema es aparentemente simple, su complejidad se muestra en el análisis realizado de los objetos matemáticos y de los procesos que se analizarán en la sección 5.2.5. También en la literatura relacionada con este problema se han descrito varias soluciones erróneas que comentamos a continuación. Estas soluciones pueden ser debidas a errores en el proceso de representación-interpretación (conflictos semióticos) o bien a la atribución de propiedades que no tienen a ciertos objetos o situaciones, como vemos en los casos que siguen.

Razonamiento erróneo 1. Percepción de la independencia

Un primer problema se produce porque *no se percibe la dependencia de los sucesivos experimentos* (elegir una puerta inicialmente) y (puerta que abre el locutor). Es decir, o bien no se visualiza la estructura del experimento compuesto o se suponen los sucesivos experimentos como independientes, atribuyendo una propiedad (independencia) que no tienen los experimentos. Pensamos que esto se produce debido a un conflicto semiótico, pues no se ha interpretado correctamente la descripción verbal del experimento (ha habido un fallo de interpretación de esta descripción verbal, que no es más que la representación del experimento real).

A primera vista parece obvio que da igual cambiar de puerta o no, pues no se visualiza la forma en que la información proporcionada por el locutor afecta a la probabilidad inicial de obtener un premio que, sin esta información, es 1/3. De nuevo hay un fallo en percibir una propiedad: Se puede condicionar un suceso por otro que aparece antes o después de él y este condicionamiento puede cambiar la probabilidad inicial del suceso. Este error de razonamiento es explicado por Falk (1986), mediante la "falacia del eje temporal" que consiste en que las personas creen erróneamente que una información actual (la puerta mostrada por el locutor) no puede afectar a un suceso que ocurrió con anterioridad a la misma (en qué puerta estaba el premio). Esta falacia, que apareció en el Estudio 2, puede estar causada, en parte, por la confusión entre condicionamiento y causalidad (conflicto semiótico producido al confundir entre sí dos conceptos diferentes).

Desde el punto de vista de la probabilidad, si un suceso *A* es la causa estricta de un suceso *B*, siempre que suceda *A*, sucederá *B*, por lo que $P(B/A)=1$, es decir, si un suceso *A* es causa de otro suceso *B*, entonces *A* es dependiente de *B*. Pero el contrario no siempre se cumple según Falk (1986): Un suceso *A* puede ser dependiente de otro suceso *B* sin que uno sea la causa del otro. Por ejemplo, se sabe que el cáncer de pulmón depende del hábito de fumar; pero fumar en sí mismo no es siempre la causa del cáncer.

Razonamiento erróneo 2. Incorrecta percepción del espacio muestral

Otra posibilidad de error en este problema es una incorrecta enumeración del espacio muestral en uno o varios de los experimentos que intervienen. Habría un fallo en pasar de la idea de espacio muestral (intensivo) al espacio muestral concreto (extensivo) o lo que es lo mismo, fallo en la particularización del espacio muestral en este experimento. La intuición nos dice que, una vez elegida la puerta, y quitando la puerta que abre el locutor, que nunca tiene premio, sólo quedan dos posibilidades equiprobables. Por tanto, la puerta que nosotros escogimos tiene un 50 % de tener una cabra y por tanto da igual cambiar qué no hacerlo. En este razonamiento se está realizando una incorrecta enumeración del espacio muestral al calcular la probabilidad condicionada, otro sesgo descrito por Gras y Totohasina (1995) y que apareció en algunos futuros profesores en el Estudio 2. El problema radica en que no consideramos la información disponible de que "el presentador conoce donde está el premio". Ya que el presentador abre la puerta *después* de la elección de jugador, el espacio muestral en el segundo experimento depende del resultado del primero:

- Si el jugador escoge en su primera opción la puerta que contiene el coche (con una probabilidad de 1/3), entonces el presentador puede abrir cualquiera de las dos puertas restantes. El espacio muestral tiene dos posibilidades con probabilidad 1/2 Además, el jugador pierde el coche si cambia cuando se le ofrece la oportunidad.

- Pero, si el jugador escoge una cabra en su primera opción (con probabilidad 2/3), el presentador *sólo* tiene la opción de abrir una puerta, y esta es la única puerta restante que contiene una cabra. En ese caso, el espacio muestral tiene un solo elemento, la puerta restante *tiene* que contener el coche, por lo que cambiando lo gana.

Razonamiento erróneo 3. Incorrecta asignación inicial de probabilidades

Una variante de la anterior interpretación incorrecta es la siguiente: Si el presentador escoge de manera aleatoria entre las puertas que aún no se han abierto, entonces la probabilidad que el candidato se lleve el coche (en el caso de no cambiar de puerta) es 1/2 pues el coche ha de estar en una de las puertas no abiertas. El problema ahora es que la asignación de probabilidades a las puertas que abre el locutor es incorrecta. Este fallo se debe a incorrecta aplicación de la regla de la suma de probabilidades porque hay una errónea descomposición-composición de los sucesivos espacios muéstrales en los diferentes experimentos.

Esto sucede porque lo que muestra el presentador no afecta la elección original, sino sólo a las otras dos puertas no escogidas. Una vez se abre una puerta y se muestra la cabra, esa puerta tiene una probabilidad nula de contener un coche, por lo que deja de tenerse en cuenta. Si el conjunto de esta puerta más la elegida tenían una probabilidad de contener el coche de 2/3 en el experimento inicial (elegir la puerta), entonces, si una tiene una de ellas (la abierta) tiene probabilidad de 0 en el segundo experimento (que la puerta tenga el coche), la tercera puerta (no elegida ni abierta) debe tener una probabilidad de 2/3. Es decir, la probabilidad de 2/3 se traspasa entera a la puerta no escogida ni abierta por el locutor (en lugar de dividirse entre las dos puertas sin abrir), porque en ningún caso puede el presentador abrir la puerta escogida inicialmente.

Razonamiento erróneo 4. Interpretación incorrecta de la convergencia

Podría originarse una reafirmación en la creencia de que es indiferente cambiar o no de puerta si, al experimentar con el Applet, el alumno obtiene (debido a la aleatoriedad) un resultado parecido con las dos estrategias. Esta posibilidad es mayor cuando el número de experimentos que se hagan con el Applet es pequeño, pues la convergencia de las frecuencias relativas a la probabilidad se cumple a largo plazo, pero no en pequeñas series de ensayos. Si el alumno obtiene este resultado, podría llegar a admitir que su suposición inicial era correcta. Habría acá el peligro de que se reafirme en la *"creencia en la ley de los pequeños números"*, que consiste en esperar convergencia incluso en pequeñas series de experimentos (Kahneman, Slovic y Tversky, 1982). Errores relacionados con la heurística de la representatividad aparecieron en algunos futuros profesores en el Estudio 2.

Variantes y otros juegos

El juego de Monty Hall está basado en una paradoja clásica de la teoría de las probabilidades, que fue planteada por Joseph Bertrand (1822-1900), matemático francés cuyas principales áreas de trabajo fueron la Teoría de Números, la Geometría Diferencial y la Teoría de las Probabilidades. En 1888 publicó el libro *Calcul des probabilitiés,* el cual, contiene numerosos ejemplos de problemas de probabilidades en los cuales el resultado depende del método de resolución del problema, entre ellos el siguiente problema original:

Tenemos tres cajas: una caja que contiene dos monedas de oro, una caja con dos monedas de plata, y una caja con una de cada tipo. Después de elegir una caja al azar se toma una moneda al azar, por ejemplo una moneda de oro. ¿Cuál es la probabilidad de que la otra también sea de oro?

Puede parecer que la probabilidad de que vuelva a salir otra moneda de oro sea de 1/2, pero de hecho, la probabilidad es 2/3. Las soluciones correctas e incorrectas descritas se aplican ahora a esta variante. Encontramos otra versión con cartas de colores. El problema en este caso es el siguiente:

Supongamos que se tienen tres cartas: una carta de color negro en ambos lados, una carta blanca en ambos lados y una carta mixta que es de color negro en un lado y blanca por el otro. Ponemos todas las cartas en un sombrero, cogemos una al azar, y la colocamos sobre una mesa. El lado de la carta hacia arriba es de color negro. ¿Cuáles son las probabilidades de que el otro lado también sea negro?

El problema con las 100 puertas: Una versión algo más elaborada es replantear el problema con 100 puertas. Tras la elección original el presentador abre 98 de las restantes para mostrar que tras de ellas hay cabras; el concursante si no cambiase su elección ganaría el coche sólo si lo ha escogido originalmente (1 de cada 100 veces). Pero si la cambia, ganaría si no lo ha escogido originalmente (y por tanto es lo que resta tras abrir las 98 puertas), 99 de cada 100 veces.

Urna con fichas de dos colores (two colours). Esta actividad (Figura 3.3.5) permite al usuario simular la extracción bolas de color rojo y verde que se encuentran en tres cajas. Las cajas están predispuestas de manera que hay dos bolas rojas en una caja, dos bolas de color verde en otra, y una verde y una bola roja en el tercero. El usuario puede mezclar el orden de las cajas y el orden en el que se extraen las bolas de las cajas. Para ejecutar el modo de prueba múltiple, introducimos el número de ensayos que deseamos en el cuadro y hacemos clic en el botón de ejecutar múltiples ensayos. El programa permite al usuario manipular el número de pruebas para experimentar con la probabilidad condicional. Para calcular la probabilidad de que el evento se produzca, debe tener en cuenta qué efecto tiene el primer evento en el segundo. En este Applet la condición es que la primera bola debe ser verde. En el podemos comparar el número de veces que la primera bola sea verde y el número de veces que ambas son de color verde para conocer la probabilidad condicional. En la Tabla 5.2.3 se presentan las variantes encontradas.

Figura 5.2.5. Two colours

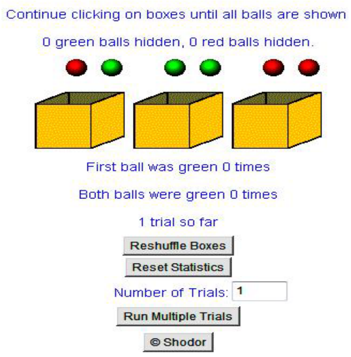


Tabla 5.2.3. Algunos juegos sobre probabilidad condicional

Nombre	Dirección
Cartas	web.educastur.princast.es/ies/iesreype/Departamentos/Mates/paradojas.htm
Las dos monedas	www.betweenwaters.com/probab/coingame/coinmainD.html
Let’s make a deal	www.stat.sc.edu/~west/javahtml/LetsMakeaDeal.html
Monty Hall	nlvm.usu.edu/en/nav/frames_asid_117_g_3_t_5.html?from=topic_t_5.html
Two colours	www.shodor.org/interactivate/activities/twocolors/
100 puertas	estadisticaparatodos.es/taller/montyhall/montyhall.html

5.2.4.2. EXPLORACIÓN DE CONCEPTOS

Incluimos en este apartado los recursos que pueden servir para visualizar algunos de los objetos matemáticos que se relacionan con la probabilidad condicional, o algunas de las propiedades o teoremas de los mismos. Permiten al estudiante variar diferentes datos, tales como el número de sucesos o las probabilidades de los mismos y ver el efecto de dicho cambio sobre otros sucesos y probabilidades. A continuación describimos algunos recursos en esta categoría.

Exploración de las operaciones con sucesos y sus probabilidades

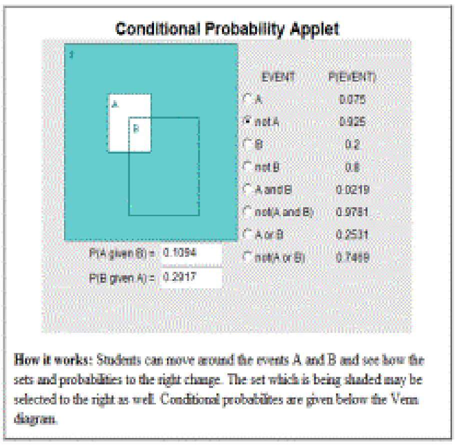
En primer lugar analizamos un recurso (cuya pantalla principal se muestra en la Figura 5.2.6 que permite explorar las operaciones entre dos sucesos.

Dirección en Internet: www.stat.tamu.edu/~west/applets/Venn1.html

Descripción

El recurso muestra un diagrama de rectángulo con una partición del espacio muestral en un suceso *A* y su contrario *no A* y otro suceso *B* y su contrario. Las probabilidades de *A* y *B* y de sus contrarios están fijadas, mientras que las de las operaciones binarias con estos cuatro sucesos van a depender de su situación relativa en el espacio muestral. Pinchando con el ratón, se puede también colorear diferentes sucesos, $A, B, \overline{A}, \overline{B}, A \cap B, \overline{A \cap B}, A \cup B$ y $\overline{A \cup B}$. La posición relativa de los sucesos *A* y *B* se pueden modificar moviendo el cursor, visualizando las probabilidades $P(A \cap B)$, $P(\overline{A \cap B})$, $P(A \cup B)$ y $P(\overline{A \cup B})$. Este programa calcula automáticamente las probabilidades condicionales $P(B/A)$ y $P(A/B)$, aunque no muestra cómo se hace el cálculo sino tan sólo el resultado.

Figura 5.2.6. Pantalla del Conditional ProbabilityApplet



Un primer objetivo es que el alumno perciba el significado de la intersección, la unión, los sucesos complementarios y cómo cambian las probabilidades según la posición relativa de los sucesos. También permite observar la diferencia entre *P(A/B)* y *P(B/A)*, ya que muchos estudiantes confunden estas dos probabilidades o las consideran iguales, según Einhorn y Hogarth (1986). Falk (1986) denomina *falacia de la condicional transpuesta* a este error. Otra posible aplicación de este recurso sería comprobar que independencia no es lo mismo que mutua exclusividad, confusión que tienen algunos estudiantes según Heitele (1975) y Kelly y Zwiers (1986). Moviendo los sucesos A y B hasta que no tengan intersección común (es decir sean mutuamente excluyentes) se observa claramente que tanto la probabilidad condicional $P(A/B)$ como la de la intersección $P(A \cap B)$ son nulas, aunque el producto de las probabilidades de los sucesos A y B no lo sea. Por tanto no se cumple la definición de independencia, aunque los sucesos sean excluyentes.

También se puede trabajar el error denominado *falacia de la conjunción* (Tversky y Kahneman, 1982b) o creencia de que es más probable la intersección de dos sucesos que la de uno de ellos por separado o la de su unión. Los estudiantes pueden experimentar que la intersección siempre tiene una probabilidad igual o menor que la de cualquiera de los sucesos que la componen.

Análisis matemático del recurso

El recurso es básicamente una visualización de las diferentes operaciones que se pueden formar con dos sucesos, y sus respectivas probabilidades. A partir de los sucesos $A, B,\ \overline{A}$ y $\overline{B}$ con $P(A),\ P(B),\ P(\overline{A})$ y $P(\overline{B})$ fijos, el Applet nos muestra directamente las siguientes propiedades:

- Probabilidad de la intersección: $P(A \cap B) = P(A) + P(B) - P(A \cup B)$
- Probabilidad de la unión: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Probabilidad del complementario: $P(\overline{A}) = 1 - P(A)$
- Leyes de Morgan: $P(\overline{A \cap B}) = P(\overline{A} \cup \overline{B})$ *o* $P(\overline{A \cup B}) = P(\overline{A} \cap \overline{B})$
- Cálculo de la probabilidad condicional a partir de la intersección $P(A/B) = \dfrac{P(A \cap B)}{P(B)}$; $P(B/A) = \dfrac{P(A \cap B)}{P(A)}$.

Objetos matemáticos puestos en juego

A continuación incluimos la tabla de análisis de objetos matemáticos y significados implícitos en el recurso.

Tabla 5.2.4. Objetos matemáticos implícitos en el recurso

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Exploración de operaciones entre sucesos y sus probabilidades	- Experimentación del cambio de probabilidades al cambiar la posición relativa de dos sucesos
Lenguajes	- Gráfico: Visualización mediante rectángulos	- Partición de un espacio muestral (*A*, no *A*; *B*, no *B*) - Intersección y unión de *B* con el suceso *A* - El área de rectángulo total sería la probabilidad 1
	- *A*, *B*	- Sucesos
	- *A* and *B*, not *A* and *B*	- Intersección de sucesos e intersección de complementarios
	- Not *A*, not *B*	- Complementarios de los sucesos *A* y *B*
	- *A* or *B*, Not *A* or *B*	- Unión de los sucesos *A* y *B*
	- *A* given *B*, *B* given *A*	- Condición; hay una incorrección pues lo que se condiciona es la probabilidad no los sucesos
	- *P*(*A*), *P*(*B*), …	- Probabilidad de los sucesos
	- *P*(*Event*)	- Se refiere a las probabilidades de los sucesos listados en la pantalla
	- Verbal	- Explicación de la situación
Conceptos	- Experimento aleatorio	- Experimento abstracto, no se concreta
	- Sucesos	- Cuatro partes en el espacio muestral, se trata de sucesos disjuntos dos a dos
	- Espacio muestral	- Conjunto de posibilidades
	- Complementarios	- El espacio muestral menos el suceso
	- Unión	- Suceso formado por el conjunto de todos los elementos que forman parte de cada suceso por separado
	- Intersección	- Suceso formado por el conjunto de todos los elementos comunes a los todos los sucesos
	- Partición	- El área de cada rectángulo y su complementario es igual al del rectángulo mayor
	- Probabilidad	- Medida relativa del área de cada parte respecto al total
	- Probabilidad condicional *P*(*B/A*)	- Medida relativa del área de $B \cap A$ respecto a cada parte *A*
Procedimientos	- Cambio de posición relativa con el cursor	- Se modifica el área de algunas regiones y su medida relativa respecto al total
	- Cálculo de probabilidades condicionadas	- Se aplicaría la fórmula; automático
	- Comparación de probabilidades	- Representación de las distintas probabilidades simples y condicionadas; automático; visualmente las puede comparar el alumno
	- Representación gráfica	- Diagrama
Propiedades	- La probabilidad condicional *P*(*A*/*B)* puede ser diferente de la probabilidad condicional *P(B/A)*	- La medida relativa del área de $B \cap A$ respecto a cada parte *A*, puede ser diferente a la medida relativa del área de $B \cap A$ respecto a *B*

	- Independencia	- Cuando dos sucesos son independientes se muestra que la probabilidad condicional de *B* condicionado a *A* es igual a la probabilidad de *B*
	- Incompatibilidad	- Cuando dos sucesos son incompatibles se muestra que la probabilidad de *B* condicionado a *A* es igual a cero
Argumentos	- Visualizaciones	- Visualización de las distintos sucesos y su relación

Dificultades posibles de los estudiantes

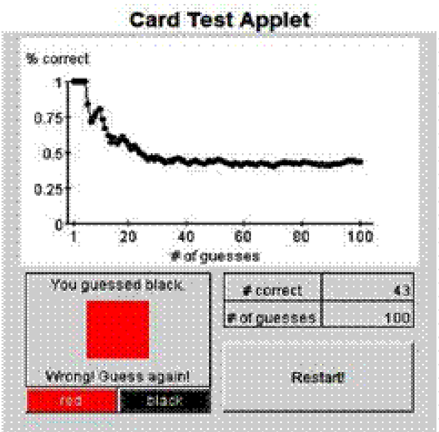
Una de las principales dificultades que pueden encontrar los estudiantes con este recurso es la interpretación del lenguaje del Applet. Por ejemplo, se puede interpretar que la probabilidad sólo depende de la posición relativa de los sucesos, aunque en realidad también dependería del tamaño de los sucesos en relación al espacio muestral.

Por otro lado, en la columna derecha sólo aparece mención a *P(event)*, pero luego en cada fila no vuelve a aparecer la notación de probabilidad. La notación coloquial para los sucesos intersección y la probabilidad condicional puede ocasionar errores. Einhorn y Hogarth (1986) sugieren que los enunciados que usan la conjunción "*y*" pueden llevar a confundir la probabilidad conjunta y la probabilidad condicional, confusión que apareció en algunos casos en los Estudios 1 y 2.

Variantes y otros recursos de exploración

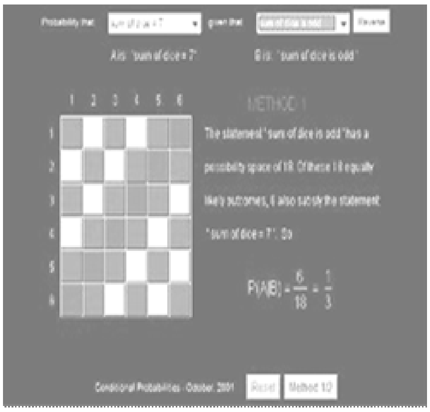
En la Figura 5.2.7 mostramos un recurso que puede servir para explorar la idea de independencia. Los estudiantes deben imaginar una baraja de cartas que contienen tarjetas de color rojo y negro y hacer predicciones sobre la ocurrencia de diferentes sucesos. El porcentaje de tarjetas rojas puede ser modificado.

Figura 5.2.7. Card test Applet



Se trata de hacer concienciar a los estudiantes de que la probabilidad de cada suceso no varía en función de los resultados obtenidos. Se les debe alentar a jugar el juego de varias maneras. En primer lugar, hacemos un ejercicio en el que la carta adivinada es la roja. De esta manera se puede estimar la proporción de tarjetas rojas encubiertas. El mismo ejercicio puede ser realizado por adivinar el porcentaje de cartas de color negro. Al final del ejercicio, los estudiantes deben ser animados a reflexionar sobre la idea de independencia y sobre la existencia de sesgos tales como la falacia del jugador. También se puede hacer observar la estabilización de las frecuencias relativas a la larga, pero hacer notar las fluctuaciones en las series cortas de ensayos.

Figura 5.2.8. Conditional probability Applet



El recurso de la Figura 5.2.8 consta de un cuadrado de seis por seis que representa las 36 posibilidades a la hora de tirar 2 dados de seis caras y nos permite investigar cómo se comportan las probabilidades condicionales. Hay 2 listas desplegables en la parte superior de la pantalla. Podemos elegir una lista (blancas, azul, probabilidad de ser blanca, probabilidad de ser azul, etc.) para ver los diferentes sucesos que se presentan en el Applet y decidir cuál es el suceso condicionante. Cuando se haya elegido el suceso de cada lista, algunas de las celdas del cuadrado se colorearan de rojo o amarillo.

Tabla 5.2.5. Algunos recursos para visualización de la probabilidad condicional

Nombre	Dirección
Condicional probability demo	onlinestatbook.com/chapter5/conditional_demo.html
Conditional probability	www.rfbarrow.btinternet.co.uk/htmasa2/Prob2.htm
Conditional probability and independent events	www.cut-the-knot.org/Curriculum/Probability/ConditionalProbability.shtml
Conditional Probability and Multiplication Rule	www.spsu.edu/math/deng/m2260/stat/cond/cond.html
Conditional probability Applet	www.stat.tamu.edu/~west/Applets/Venn1.html
Probability by Surprise	www-stat.stanford.edu/~susan/surprise/

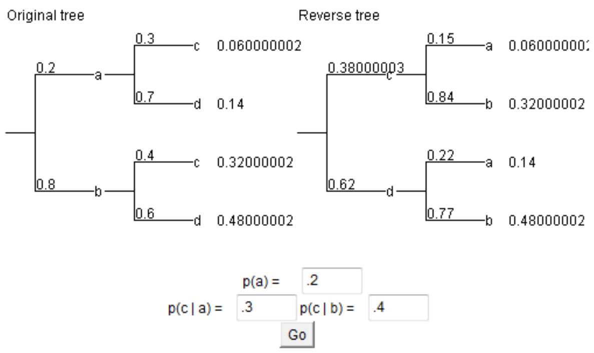
Los cuadrados de color representan el número de combinaciones de los dados que satisfagan la condición *B* (la segunda condición de la derecha de la lista desplegable). De estas celdas de color, el rojo representa las combinaciones, que también cumplen la primera condición (*A*). Existen dos métodos para calcular $P(A/B)$, uno de ellos implica contar los cuadrados de colores, el otro utiliza una fórmula. El Applet nos proporciona métodos para ver cómo se relacionan entre sí. Si pulsamos "Reverse", se intercambian las declaraciones *A* y *B*. Con lo que debemos detectar rápidamente que $P(A/B)$ no es, en general, igual a $P(B/A)$. En la Tabla 5.2.5 presentamos las direcciones de estos y otros recursos de exploración.

5.2.4.3. RECURSOS PARA RESOLVER PROBLEMAS

En este apartado analizamos algunos recursos que pueden ayudar al alumno en su resolución de problemas de probabilidad compuesta y condicionada. En el Departamento de Ingeniería Matemática, Universidad de Chile encontramos el recurso mostrado en la Figura 3.5.1, dentro de un Curso de probabilidades y procesos estocásticos.

Dirección en internet: www.dim.uchile.cl/~mkiwi/ma34a/libro/chapter4/Tree/Tree.html

Figura 5.2.9. Pantalla principal del recurso para construcción de diagramas en árbol



Descripción

Según Fischbein (1987), el diagrama en árbol facilita la resolución de problemas de probabilidad. Fischbein comparte con Bruner la hipótesis de que una estructura matemática puede manifestarse de una manera enactiva, icónica y simbólica sin cambiar sus características y de que el uso de métodos adecuados de representación facilita la resolución de problemas de probabilidad. Entre las representaciones gráficas, destaca el diagrama de árbol, que son uno de los recursos más útiles para visualizar tanto situaciones de combinatoria como situaciones probabilísticas. En la terminología de Fischbein (1987) pertenecen a los "modelos esquemáticos" y presenta importantes características intuitivas. Ofrecen una representación de la estructura de la situación, lo que contribuye a la inmediatez de la comprensión y a la búsqueda de la solución del problema. A pesar de esta importancia, Pesci (1994) demostró que algunos estudiantes tenían dificultades para construir diagramas de árbol adecuados para representar situaciones problemáticas y, en ocasiones, el mismo gráfico es la causa de muchos errores. Estos errores también aparecieron en las investigaciones de Navarro-Pelayo (1994) y Roa (2000) y, en algún caso, en el Estudio 2.

El recurso Bayes Tree está pensado para enseñar esta tarea. La pantalla presenta un diagrama de árbol con dos ramas, cada una con dos bifurcaciones, por lo cual puede ser útil para visualizar la solución de problemas de Bayes en caso de sólo dos sucesos en la partición del espacio muestral. Presenta unas casillas donde se pueden introducir las probabilidades a priori *P(A)* y las probabilidades condicionales *P(C/A)* y *P(C/B)*. Una vez introducidos los datos, se presiona el botón "go" y el programa calcula las probabilidades conjuntas $P(A \cap C) = P(A)\, P(C/A)$ etc. También incluye el árbol inverso que permite calcular la probabilidad condicional inversa $P(A/C)$, por lo cual, de nuevo sirve para visualizar la diferencia entre estas dos probabilidades condicionales.

Análisis matemático del recurso

El recurso resume básicamente el teorema de Bayes y las propiedades de la probabilidad condicionada:

- Teorema de Bayes: $P(B_i / A) = \dfrac{P(A/B_i)P(B_i)}{\sum_i P(B_i)P(A/B_i)} = \dfrac{P(A/B_i)P(B_i)}{P(A)}$
- Cálculo de la probabilidad condicional: $P(A/B) = \dfrac{P(A \cap B)}{P(B)}$
- Probabilidad de la intersección: $P(A/B)P(B) = P(A \cap B)$ $P(B/A)P(A) = P(A \cap B)$
- Cálculo de la probabilidad simple a partir de la intersección y de la condicional (variante) $P(B) = \dfrac{P(A \cap B)}{P(A/B)}$ ó $P(A) = \dfrac{P(A \cap B)}{P(B/A)}$.

Objetos matemáticos puestos en juego

El análisis de este recurso muestra los siguientes objetos matemáticos implícitos:

Tabla 5.2.6. Objetos matemáticos implícitos en el recurso

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Exploración del teorema de Bayes	- Experimentación de la descomposición de los sucesos al cambiar las probabilidades condicionadas y simples
Lenguajes	- Gráfico: Visualización mediante diagramas de árbol	- Partición de un espacio muestral en el experimento simple - Partición de un espacio muestral en el experimento compuesto - Descomposición de un experimento en probabilidades condicionadas - Posibles sucesos que originan un suceso dado
	- *a, b, c*	- Sucesos de la partición
	- $C\|B$, $C\|A$	- Condicionamiento de sucesos
	- $P(A)$, $P(C\|B)$, $P(C\|B)$	- Probabilidad simple y probabilidades condicionadas iniciales
	- Icónico	- Iconos que representan los sucesos y resultados
Conceptos	- Experimento aleatorio	- Experimento abstracto, no se concreta
	- Sucesos - Sucesos complementarios	- Cuatro sucesos en el espacio muestral - Los sucesos *a* y *b* son complementarios - Los sucesos *c* y *d* son complementarios
	- Probabilidad simple	- Probabilidad de cada elemento de forma independiente
	- Probabilidad condicional	- Probabilidad de los elementos compuestos que forman el diagrama de árbol
	- Intersección de sucesos - Probabilidad conjunta	- Conjunto común de sucesos - Probabilidad de la intersección
Procedimientos	- Cambio de probabilidades iniciales con el cursor	- Se modifica el diagrama de árbol y las respectivas probabilidades simples y condicionadas
	- Cálculo de probabilidades condicionadas	- Se aplicaría la fórmula; automático
	- Cálculo de probabilidades inversas	- Se aplicaría la fórmula; automático
	- Comparación de probabilidades	- Representación de las distintas probabilidades simples y condicionadas; automático
	- Representación gráfica	- Diagrama de árbol; esquema
Propiedades	- La probabilidad condicional $P(A/B)$ puede ser diferente de la probabilidad condicional $P(B/A)$	- Se observa numéricamente
	- Teorema probabilidad total	- A partir de las distintas probabilidades representadas podemos calcular la probabilidad total
	- Teorema de Bayes	- A partir de las distintas probabilidades representadas podemos calcular la probabilidad condicional de cualquier suceso condicionado a otro
	- Independencia	- Cuando dos sucesos son independientes se muestra que la probabilidad condicional de un suceso condicionado a otro es igual a la probabilidad del primero
	- Incompatibilidad	- Cuando dos sucesos son incompatibles se muestra que la probabilidad condicional de un suceso condicionado a otro es igual a cero
Argumentos	- Visualizaciones	- Definición visual de las distintos sucesos y de la relación entre ellos

Dificultades posibles de los estudiantes

La posible complejidad radica en la falta de claridad del Applet, ya que no queda claro qué probabilidad calcula a lo largo de cada rama, pues no se explicita mediante notación simbólica. Los errores que pueden cometer los estudiantes a la hora de interpretar el Applet van desde identificar *a* y *b* como dos sucesos que no tienen nada en común (siendo uno el complementario del otro) o creer que el Applet representa cuatro sucesos distintos (siendo en realidad dos y sus complementarios). Otro problema que se puede dar a la hora de interpretar el diagrama es que en ningún momento aparece el significado de los valores que se calculan cuando se ingresan los resultados. Lo que lleva que se confundan probabilidades simples, condicionales o conjuntas por ejemplo *P(c)* con *P(c/a).*

Navarro-Pelayo (1994) indica que los estudiantes tienen dificultades para realizar un diagrama en árbol completo y correcto, algo que puede venir ligado a una incorrecta enumeración de los sucesos. Hay una relación entre el espacio muestral de un experimento compuesto y las operaciones combinatorias. El inventario de todos los posibles sucesos en dicho espacio muestral requiere un proceso de construcción combinatorio, a partir de los sucesos elementales en los experimentos simples.

Variantes

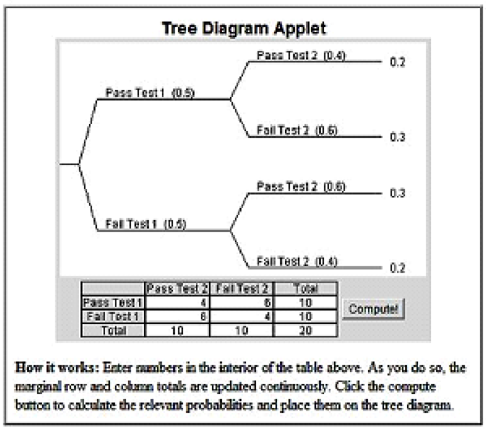
En la Figura 5.2.10 se incluye otro recurso que se presenta conjuntamente el diagrama en árbol y la tabla de doble entrada correspondiente. Escribimos los números en el interior de la tabla. Como se ve, el marginal de fila y columna de los totales se actualizan continuamente. Hacemos clic en el botón “calcular” y aparecerán las probabilidades en su lugar dentro del diagrama de árbol.

Se han encontrado otros ejemplos, tales como la Figura 5.2.11 que crea automáticamente diagramas en árbol que pueden ayudar a la interpretación de cómo se distribuyen los sucesos. Otros ejemplos los reproducimos en la Tabla 5.2.7.

Tabla 5.2.7. Algunas variantes del Bayes Tree

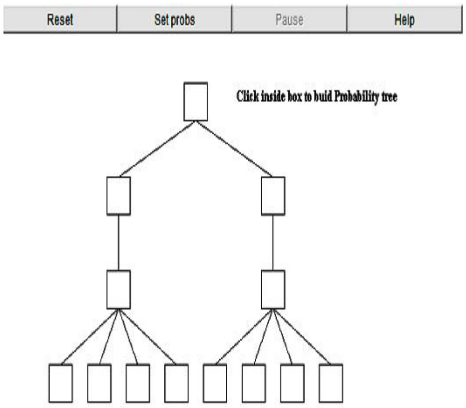
Nombre	Dirección
Building probability trees	www-stat.stanford.edu/~susan/surprise/ProbabilityTree.html
Libro de Probabilidades Universidad de Chile	www.dim.uchile.cl/~mkiwi/ma34a/libro/chapter4/Tree/Tree.html
Tree Diagram Applet	http://www.stat.tamu.edu/~west/applets/tree.html
Probability trees	www.geocadabra.nl/geocadabra_english_probabilitytree.htm

Figura 5.2.10 Tree Diagram Applet



	Pass Test 2	Fail Test 2	Total
Pass Test 1	4	6	10
Fail Test 1	6	4	10
Total	10	10	20

Figura 5.2.11. Building a probability tree



5.2.4.4. LECCIONES O LIBROS DE TEXTO

Son muchos los profesores o investigadores que ponen libros de texto en Internet, entre ellos, textos de estadística que contienen una lección sobre la probabilidad condicionada. La variedad de libros de texto es muy grande, desde simples trabajos en formato pdf, que no se diferencian de un libro tradicional, hasta páginas de Internet que, además de definiciones o explicaciones, tienen recursos interactivos, calculadoras, ejemplos, tablas o juegos. Analizamos a continuación uno de ellos.

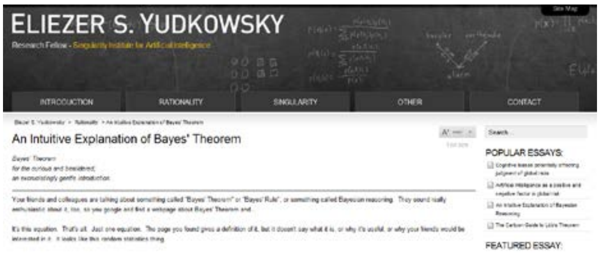
Dirección en Internet: yudkowsky.net/rational/bayes

Descripción

En la web de Eliezer Yudkowsky (Figura 5.2.12), podemos encontrar 'An Intuitive Explanation of Bayes Theorem', un tratado sobre el teorema de Bayes dentro del contexto científico médico (con ejemplos aplicados al cáncer de mama), donde se explica en qué consiste la probabilidad condicional, el teorema de Bayes y las diferencias entre el razonamiento estadístico tradicional y el bayesiano. Un aspecto importante es el de las tablas 2x2 que aparecen en la mayoría de los ejemplos. Con ellas podemos comprobar todas las propiedades de la probabilidad condicionada tales como la intersección, teorema de Bayes, etc.

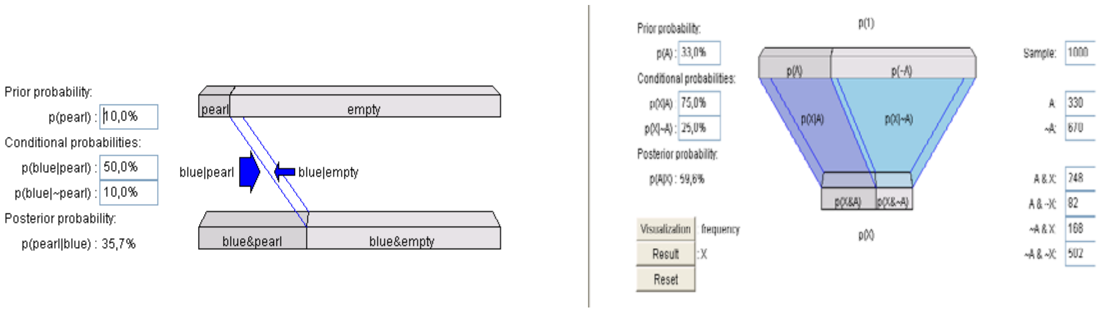
El autor plantea cuestiones médicas tales como: *"El 1% de las mujeres a la edad de cuarenta años que participan en una mamografía tienen cáncer de mama. El 80% de las mujeres con cáncer de mama positivo dan positiva en una mamografía. El 9,6% de las mujeres sin cáncer de mama también tendrá positivas mamografías. Una mujer en* *este grupo de edad dio positiva la mamografía. ¿Cuál es la probabilidad de que ella tenga realmente el cáncer de mama?”*. Proporciona recursos para realizar o visualizar el cálculo y variar los datos, donde se pueden observar elementos correspondientes a la probabilidad condicionada y a las relaciones que se dan entre las probabilidades.

Figura 5.2.12. An Intuitive Explanation of Bayes Theorem



El autor también destaca los distintos errores que se pueden dar en este tipo de problemas, tanto a nivel de estudiantes como a nivel profesional, indicando, por ejemplo que *“sólo alrededor del 15% de los médicos son capaces de hacer el problema de manera correcta”*. Además podemos encontrar calculadores para ayudarnos a comprobar cómo funcionan este tipo de problemas y explicaciones de cómo descomponer el problema de manera que la resolución de este sea lo más fácil posible.

Figura 5.2.13. Dos recursos de visualización en la lección



Otro aspecto importante que podemos observar es la descripción de los distintos errores que se pueden dar en este tipo de problemas, tanto a nivel estudiantes como a nivel profesional, tales como *“sólo alrededor del 15% de los médicos son capaces de resolver el problema de manera correcta”*. Además podemos encontrar calculadores para ayudarnos a comprobar cómo funcionan estos problemas y explicaciones de cómo descomponer el problema de manera que la resolución de este sea lo más fácil posible. También se introducen recursos de visualización (Ver ejemplos en la Figura 5.2.13).

Análisis matemático del recurso

El recurso resume el teorema de Bayes y el funcionamiento de la Estadística Bayesiana, sobre todo aplicadas a tablas 2x2. A partir de los sucesos A y B_i con $P(A)$ y $P(B_i)$, el autor nos muestra cómo funcionan las diferentes propiedades de la probabilidad condicionada, aplicando estas propiedades a ejemplos médicos, sobretodo sobre cáncer.

Objetos matemáticos puestos en juego

A continuación incluimos la tabla de análisis de objetos matemáticos y significados implícitos en el recurso. Nótese que la notación para el suceso intersección es incorrecta y no hay una notación específica para la probabilidad condicional.

Tabla 5.2.8. Objetos matemáticos implícitos en el recurso

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Exploración del teorema de Bayes	- Experimentación del cambio de probabilidades inversas al cambiar los datos
Lenguajes	- *A*, *B*, *x*	- Sucesos de la partición
	- *A* y *B*	- Intersección de sucesos
	- *A* |*B*, *B* |*A*, *A*|*x*, *x*|*A*	- Condición
	- *P*(*A*), *P*(*B*), …	- Probabilidad de los sucesos
	- Icónico	- Iconos que representan los sucesos y resultados
	- Verbal	- Explicación de la situación
Conceptos	- Experimento aleatorio	- Experimento abstracto, no se concreta
	- Sucesos	- Cuatro partes en el espacio muestral, se trata de sucesos disjuntos
	- Tablas 2x2	- Situación espacial de los sucesos
	- Complementarios	- El total menos el suceso
	- Unión	- Suceso formado por el conjunto de todos los elementos que forman parte de cada suceso por separado
	- Intersección	- Suceso formado por el conjunto de todos los elementos comunes a los todos los sucesos
	- Partición	- El área de los cuatro rectángulos es igual al del rectángulo mayor
	- Probabilidad	- Medida relativa de cada parte respecto al total
	- Probabilidad condicional	- Medida relativa de $B\cap A$ respecto a cada parte *A* o de *B*
	- Intersección de sucesos	- Conjunto común de sucesos
	- Cálculo de probabilidades condicionadas	- Se aplicaría la fórmula en los calculadores
	- Cálculo de probabilidades simples	- Se aplicaría la fórmula en los calculadores
	- Comparación de probabilidades	- Representación de las distintas probabilidades simples y condicionadas
Propiedades	- La probabilidad condicional *P*(*A/B*) puede ser diferente de la probabilidad condicional *P*(*B/A*)	- La medida relativa del área de $B\cap A$ respecto a cada parte *A* puede ser diferente a la medida relativa del área de $B\cap A$ respecto a *B*

	- Teorema probabilidad total	- A partir de las distintas probabilidades representadas podemos calcular la probabilidad total
	- Teorema de Bayes	- A partir de las distintas probabilidades representadas podemos calcular la probabilidad condicional de cualquier *A* condicionado a *B*
	- Independencia	- Cuando dos sucesos son independientes se muestra que la probabilidad condicional de *B* condicionado a *A* es igual a la probabilidad de *B*
	- Incompatibilidad	- Cuando dos sucesos son incompatibles se muestra que la probabilidad condicional de *B* condicionado a *A* es igual a cero
Argumentos	- Visualizaciones	- Definición visual de las distintos sucesos y de la relación entre ellos

Dificultades posibles de los estudiantes

Una de las principales dificultades que pueden encontrar los estudiantes es que la notación y las definiciones son poco explicitas y, a veces, la forma en la que el autor se expresa es informal. Aparte de ello, el autor intenta explicar el comportamiento del teorema de Bayes sin dar nociones básicas de la probabilidad, sus propiedades y sus definiciones. Aunque posiblemente el recurso es útil para médicos y otros profesionales de ciencias de la salud, quienes han de usar el teorema de Bayes en el diagnóstico, puede que los estudiantes no transfieran lo aprendido a otras situaciones de uso del teorema de Bayes. También podrían presentarse algunas de las dificultades ya descritas en relación a los recursos anteriores y que aparecieron en el Estudio 2.

Otros ejemplos

En la Figura 5.2.14 mostramos una unidad de introducción a la probabilidad condicionada, el teorema de la probabilidad total y el teorema de Bayes, obtenida de la página Descartes del Ministerio de Educación. En ella se supone el conocimiento de los conceptos elementales de la teoría de la probabilidad (suceso, operaciones con sucesos, etc.). La unidad tiene también una práctica para recordar las propiedades de la probabilidad y la relación con la unión e intersección de sucesos.

El recurso de la Figura 5.2.15 es una lección basada en varios problemas de probabilidad condicional, en particular el trabajo con experimentos compuestos. Incluye situaciones que a muchas personas les resulta difícil de comprender ya que sus resultados experimentales pueden parecer contrarios a la intuición, con lo los podemos definir como paradójicos. Esta lección pretende conducir a consideraciones más profundas sobre las relaciones matemáticas y a la adquisición de herramientas para la resolución de problemas relacionados con probabilidad condicional y probabilidad de eventos simultáneos. Al terminar esta lección, se intenta que los estudiantes hayan aprendido la fórmula para la probabilidad de eventos simultáneos independientes. Se han encontrado otros ejemplos, que reproducimos en la Tabla 5.2.9.

Figura 5.2.14. Probabilidad Condicional

Figura 5.2.15. Probabilidad condicional y probabilidad de eventos simultáneos

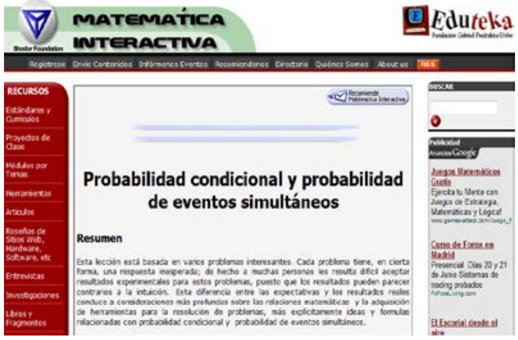


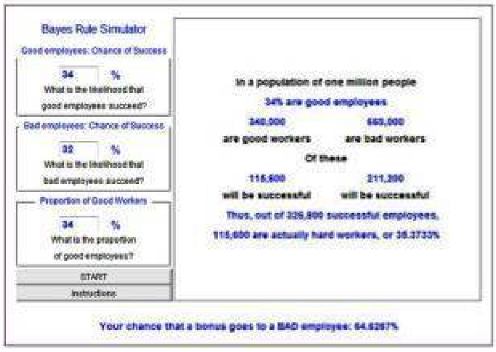
Tabla 5.2.9. Algunos textos sobre probabilidad condicional en Internet

Nombre	Dirección
An Intuitive Explanation of Bayes' Theorem	yudkowsky.net/rational/bayes
Conditional Probability; Definitions and Interpretations	www.math.uah.edu/stat/prob/Conditional.xhtml
Contingency Tables	onlinestatbook.com/stat_sim/contingency/index.html
Probabilidad condicional y ejemplos	www.hrc.es/bioest/Probabilidad_15.html
Probabilidad condicional y probabilidad de eventos simultáneos	www.eduteka.org/MI/master/interactivate/lessons/pm4.html
Statmedia	www.ub.edu/stat/GrupsInnovacio/Statmedia/demo/Statmedia.htm
The Beginnings of Probability...	mathforum.org/isaac/problems/prob1.html
Probability of Simultaneous Events Discussion	www.shodor.org/interactivate/discussions/ProbabilityOfSimulta/
Base Rates	onlinestatbook.com/chapter5/base_rates.html
The Probability of Penalizing the Innocent Due to Bad Test Results	www.intuitor.com/statistics/BadTestResults.html
Java Applets: Bayes	www.dim.uchile.cl/~mkiwi/ma34a/libro/chapter4/Bayes/Bayes.html
Conditional probability module	http://www.dim.uchile.cl/~mkiwi/ma34a/
Conditional Probability	www.cut-the-knot.org/Probability/ConditionalProbability.shtml
Conditional Probability	www.math.uah.edu/stat/prob/Conditional.xhtml
Conditional Probability and Independence	people.hofstra.edu/Stefan_Waner/tutorialsf3/unit6_5A.html
Conditional Probability	www.mathgoodies.com/lessons/vol6/conditional.html
Conditional Probability Investigation	www.saskschools.ca/curr_content/mathb30/prob/les5/invest3.htm

5.2.4.5. CALCULADORES

Algunos recursos son simplemente una ayuda en los cálculos. Entre ellos encontramos el Calculador de Bayes, que calcula la probabilidad condicionada utilizando el teorema de Bayes (Figura 5.2.16).

Figura 5.2.16. Calculador de Bayes



El Applet consiste en un simulador de probabilidad condicionada a partir de un ejemplo sobre una empresa. Se desea recompensar los buenos empleados con gran capacidad para el trabajo, con unos bonos. ¿Qué probabilidades hay de premiar a los buenos y no a los malos? A la izquierda, se introduce la posibilidad de "éxito" para cada tipo de empleado. Por ejemplo, el 80% en el primer cuadro significa que 20 veces de cada 100, una no gran capacidad para trabajo no es reconocida, o el 25% significa que 25 de cada 100 personas trabajan duro.

Análisis matemático del recurso

El recurso resume el teorema de Bayes a partir de un calculador que simula un ejemplo concreto sobre una empresa. A partir de dos sucesos, el Applet nos muestra los resultados del teorema de Bayes. Para ello utiliza los siguientes elementos y propiedades matemáticas:

- Probabilidad inicial, verosimilitud, probabilidad final;
- Teorema de Bayes: $P(B_i / A) = \dfrac{P(A/B_i)P(B_i)}{\sum_i P(B_i)P(A/B_i)} = \dfrac{P(A/B_i)P(B_i)}{P(A)}$.

Objetos matemáticos puestos en juego

A continuación incluimos la tabla de análisis de objetos matemáticos y significados implícitos en el recurso. Nótese que la notación para los sucesos es incorrecta y no hay una notación para la probabilidad condicional.

Tabla 5.2.10. Objetos matemáticos implícitos en el recurso

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones problemas	- Exploración del teorema de Bayes	- Calcular la probabilidad condicionada a partir de sucesos simples
Lenguajes	- Buen empleado, mal empleado, proporción de buenos empleados	- Sucesos
	- Icónico	- Iconos que representan los sucesos y resultados
	- Verbal	- Explicación de la situación
	- Sucesos	- Conjunto del espacio muestral
	- Complementarios	- El total menos el suceso
	- Intersección	- Suceso formado por el conjunto de todos los elementos comunes a los todos los sucesos
	- Probabilidad	- Medida relativa de cada parte respecto al total
	- Probabilidad condicional	- Medida relativa de $B \cap A$ respecto a cada parte *A* o de *B*
	- Cálculo de probabilidades condicionadas	- Se aplicaría la fórmula en los calculadores
	- Cálculo de probabilidades simples	- Se aplicaría la fórmula en los calculadores
	- Comparación de probabilidades	- Representación de las distintas probabilidades simples y condicionadas
Propiedades	- La probabilidad condicional $P(A/B)$ puede ser diferente de la probabilidad condicional $P(B/A)$	- La medida relativa del área de $B \cap A$ respecto a cada parte *A* puede ser diferente a la medida relativa del área de $B \cap$A respecto a *B*
	- Teorema probabilidad total	- A partir de las distintas probabilidades representadas podemos calcular la probabilidad total
	- Teorema de Bayes	- A partir de las distintas probabilidades representadas podemos calcular la probabilidad condicional
Argumentos	- Visualizaciones	- Definición visual de las distintos sucesos y de la relación entre ellos

Dificultades posibles de los estudiantes

Una de las principales dificultades que pueden encontrar los estudiantes es la notación, ya que en el ejemplo las definiciones de los sucesos son poco explicitas. Además, el Applet intenta explicar el comportamiento del teorema de Bayes sin dar nociones básicas ninguna del comportamiento de la probabilidad condicionada.

Otros ejemplos

En la Figura 5.2.17 mostramos un calculador aplicado al juego de dados y a la probabilidad condicional. Se selecciona dos sucesos de entre todos los posibles que podrían ocurrir al tirar un par de dados y que se despliegan en el menú al seleccionar el color de cada dado. En los sucesos del espacio producto se pueden seleccionar los distintos sucesos del experimento, como puede ser que el color blanco sea 4 y la suma es 7. Podemos ver que hay 11 pares de valores con el elemento 4 y para las que 2 dan como suma 7. Por lo tanto, la probabilidad de que la suma 7, dado que aparece un 4, es

2 /11. A partir de ejemplos de este tipo podemos comprobar cuál es el funcionamiento y la relación de los sucesos en la probabilidad condicionada.

En el recurso de la Figura 5.2.17 encontramos un calculador de probabilidad simple y condicionada aplicado a un ejemplo en el que se supone que el espacio muestral muestra la población de adultos en una pequeña ciudad que hayan cumplido los requisitos para un título universitario. Se clasifican los datos según el sexo y el tipo de empleo. Eligiendo una persona que se elegirán al azar podemos encontrar *P(E), P(M) y P(M/E).*

Se han encontrado otros ejemplos, que reproducimos en la Tabla 5.2.11.

Figura 5.2.16. Probabilidad Condicionada

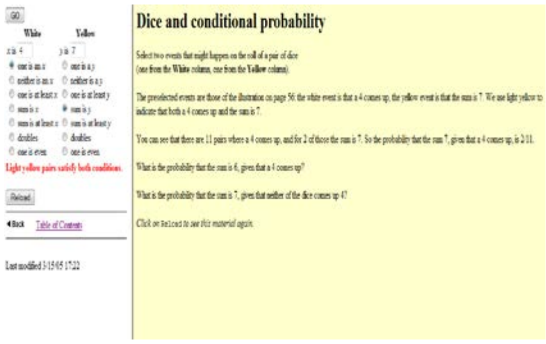


Figura 5.2.17. Conditional Probability and Multiplication Rule

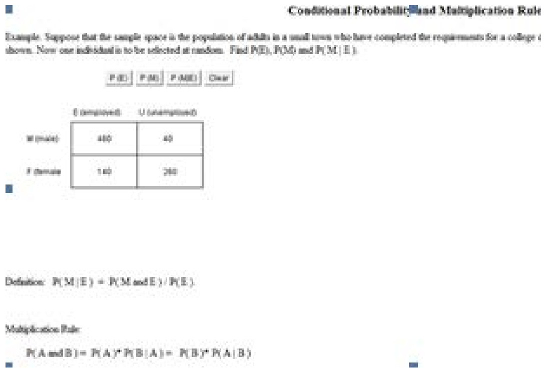


Tabla 5.2.11. Algunos calculadores sobre probabilidad condicional en Internet

Nombre	Dirección
Probability Calculator: Solve Basic Probability Problems	stattrek.com/Tools/ProbabilityCalculator.aspx
Bayes' Rule Calculator	stattrek.com/Tools/BayesRuleCalculator.aspx
Bayes Demonstration Applet	www.stat.sc.edu/~west/applets/bayesdemo.html
Basic Probability Calculator	www.easycalculation.com/statistics/probability.php
Bayes Rule Applet: Who to Reward?	www.gametheory.net/Mike/Applets/Bayes/WhoReward.html

5.2.5. PROCESOS MATEMÁTICOS

De acuerdo a Godino, Font y Wilhelmi (2008), un segundo tipo de análisis es el de los *procesos matemáticos y conflictos semióticos*. En toda práctica interviene un *sujeto* (en nuestro caso los posibles alumnos que resuelven el problema). Este nivel de análisis se centra en los procesos que, efectivamente, intervienen en la realización de las prácticas previstas y los conflictos de los estudiantes.

Además de los objetos matemáticos ya descritos, los autores consideran procesos de *materialización – idealización, particularización – generalización, descomposición – reificación, representación – significación, personalización – institucionalización*. La finalidad es describir la complejidad de las prácticas matemáticas como factor explicativo de los conflictos semióticos que se podrían producir en su realización. En nuestro caso, puesto que el análisis de los recursos se hace a priori, el estudio tendrá un carácter hipotético.

Podemos observar los siguientes procesos de *materialización-idealización* (es decir pasar de algo que se percibe a algo que no se percibe):

- Los objetos gráficos de los Applet (puertas o premio en el problema de Monty Hall, rectángulos en el Applet de visualización) e incluso las acciones que realizamos con él (pinchar en una puerta; mover un rectángulo) son objetos visibles u ostensivos, pero representan a otros objetos u acciones que son imaginarios (puerta en el concurso, coche o cabra reales; apuesta que se hace; sucesos en una partición).
- Las operaciones aritméticas se representan simbólicamente; por ejemplo *1/3* representa la operación de dividir la unidad en tres partes y también el resultado de la división, que es un número real que no percibimos directamente.
- Los diferentes símbolos representan propiedades, conceptos u operaciones. Por ejemplo, $G=(G\cap A)\mathrm{U}(G\cap B)$ representa en forma visible una propiedad matemática.
- El diagrama en árbol de las soluciones intuitivas representa, por un lado la estructura del experimento (por pasos) y en cada paso, los resultados del experimento y sus probabilidades.

Respecto a los *procesos de particularización – generalización (dualidad extensivo – intensivo)*, es decir, cuando pasamos de un caso particular, generalizando a una propiedad de un conjunto o viceversa, cuando una propiedad que sabemos es general, la aplicamos a un caso particular, observamos los siguientes:

- Sabemos que la suma total de todas las probabilidades de los sucesos en un experimento es la unidad. En cada ejemplo, particularizando llegamos a las probabilidades de los sucesos dados. Por ejemplo, en el problema de Monty Hall deducimos que la probabilidad inicial de elegir una de las tres puertas es *1/3*.
- Al trabajar con el Applet en el problema de Monty Hall, vemos que si cambiamos de puerta estamos obteniendo el doble de aciertos que si no cambiamos. Aunque no

sepamos por qué, aceptamos (generalizando) que la probabilidad de acierto cambiando de puerta será el doble que si no se cambia. En los diferentes diagramas en árbol, observamos que la suma de probabilidades en cada rama es la unidad. Generalizando, asumimos que la suma de todos los sucesos elementales en cada experimento ha de ser igual a 1.

- Conocemos el axioma de la unión que se cumple en forma general. Para cada problema concreto, lo aplicamos para calcular la probabilidad total o la probabilidad de la unión de dos o más sucesos (aunque se cumple para cualquier número de sucesos, acá particularizamos para cada caso).

Los *procesos de representación y significación* aparecen continuamente en el trabajo matemático, pues, como no podemos operar directamente con objetos ideales, representamos las operaciones sobre los mismos por medio de símbolos o por medio de otros objetos. Es decir, o bien representamos un objeto abstracto de cierta manera o bien tenemos que interpretar (dándole un significado) una representación matemática:

- El juego real de Monty Hall o la construcción real de un diagrama en árbol lo representamos en el Applet mediante gráficos dinámicos; la elección de una puerta la representamos pinchando en el gráfico que representa esta puerta; la construcción de una nueva rama, la representamos mediante el dibujo que produce el Applet.
- El objeto "probabilidad" lo representamos por la letra *P*; la probabilidad de un suceso que denominamos *A* lo representamos mediante *P(A). A* (el suceso) a su vez representa un resultado; por ejemplo, ganar en un juego.

El alumno que trata de resolver el problema tiene que pasar constantemente de considerar objetos elementales (unitarios) a considerar objetos compuestos de varios objetos elementales (sistémico). Encontramos los siguientes *procesos de descomposición – reificación (dualidad sistémico – unitario)*:

- Cada suceso de un experimento aleatorio es elemental. Pero el espacio muestral del experimento es sistémico.
- Cada resultado de un experimento es elemental. Pero la frecuencia relativa o el porcentaje de veces que ocurre un resultado se calcula como una operación sobre el total de los resultados al realizar *n* veces el experimento.
- Cada rama del diagrama en árbol es elemental, mientras que todo el diagrama en árbol es sistémico.

- En la solución formal 2 al problema de Monty Hall, cada valor de la variable aleatoria es elemental, mientras que la variable aleatoria en sí misma es sistémica, lo mismo que su distribución.

Finalmente, tenemos los *procesos de personalización - institucionalización (dualidad personal – institucional).* Los autores indican que al comenzar un proceso de estudio será necesario lograr que los estudiantes asuman el problema y se involucren en su solución. Es lo que Brousseau (1986) denomina la *devolución* del problema. Pensamos que esto se consigue totalmente con estos recursos pues aumenta el interés del alumno y olvida que se trata de una clase de matemáticas. Ha personalizado la situación. El proceso inverso es pasar de lo personal a lo institucional. Esto lo logra el maestro cuando discute colectivamente con los estudiantes para decidir cuáles de las soluciones son correctas y por qué.

5.2.6. IDONEIDAD DIDÁCTICA DEL TRABAJO CON LOS RECURSOS

A continuación aplicamos la noción de idoneidad didáctica, descrita en el marco teórico, al caso de la solución del problema de Monty Hall. En el trabajo real en el aula, se plantearía el problema a los estudiantes, dejando un tiempo para la posible solución. Seguidamente se discutirían con los estudiantes las soluciones correctas e incorrectas encontradas por los mismos, hasta lograr que se acepte alguna de las soluciones correctas. El profesor ayudaría a analizar las causas de los errores y haría un resumen de lo aprendido. En caso de resistencia a la solución, se dejaría experimentar con la solución para que, al ver sus intuiciones refutadas por la evidencia empírica, los estudiantes tratasen de revisar la solución errónea y sus causas.

La idoneidad didáctica se define como la articulación de seis componentes (Godino, Wilhelmi y Bencomo, 2005; Godino, Contreras y Font, 2006) cada uno de los cuales puede presentarse en mayor o menor grado. Para analizar si se verifican los diferentes tipos de idoneidad descritos en el punto anterior, los autores proporcionan unas guías de descriptores para poder valorar cada uno de los tipos de idoneidad. En nuestro caso, la intención sería utilizar el Applet con el proceso de estudio descrito en cursos de formación de profesores de secundaria. A continuación describimos y valoramos resumidamente estos componentes de la idoneidad:

- *Idoneidad epistémica o matemática:* Se trata de ver si hay representatividad de los significados institucionales implementados (o pretendidos), respecto de un significado de referencia. Pensamos que el Applet, junto con el proceso de estudio descrito podría tener una idoneidad matemática en el estudio de los conceptos de: probabilidad condicional, experimento compuesto, dependencia e independencia y experimentos dependientes e independientes. Como vimos en la Tabla 5.210, esta idoneidad podría ser más o menos grande, dependiendo del tipo de solución encontrada. En general las soluciones formales tienen mayor idoneidad en un curso universitario y de formación de profesores, pero en un curso con alumnos de Secundaria, las soluciones intuitivas podrían ser suficientes. La solución empírica, tiene en general, baja idoneidad matemática, a menos que se complemente con una solución intuitiva o formal.
- *Idoneidad cognitiva:* Mide el grado en que los significados pretendidos/ implementados son asequibles a los alumnos, así como si los significados personales logrados por los alumnos son los significados pretendidos por el profesor. Pensamos que la situación planteada tiene suficiente idoneidad en cursos de formación de profesores, sobre todo de profesores de Secundaria que son licenciados en matemáticas. Asimismo podría tener idoneidad suficiente en los últimos cursos de Secundaria o Bachillerato, pues los razonamientos descritos están al alcance de los alumnos.
- *Idoneidad interaccional*: Grado en que la organización de la enseñanza permite identificar conflictos semióticos y resolverlos durante el proceso de instrucción. Este tipo de idoneidad dependerá de cómo organiza el educador el trabajo en el aula. Será importante que los estudiantes trabajen en grupos para que surja el conflicto y se explicite. Será importante también organizar una solución colectiva de las soluciones para que los mismos alumnos ayuden a sus compañeros a detectar los puntos equivocados.
- *Idoneidad mediacional:* Disponibilidad y adecuación de los recursos necesarios para el desarrollo del proceso de enseñanza-aprendizaje. No se precisa de muchos recursos, pues incluso podría hacerse una simulación con objetos físicos. Pero, con un solo ordenador en el aula, donde los alumnos pueden jugar colectivamente se puede trabajar esta situación. La idoneidad aumentaría si se trabaja en el aula de informática donde cada alumno o profesor puede experimentar individualmente o

por parejas con el recurso.

- *Idoneidad emocional:* Interés y motivación del alumnado en el proceso de estudio. Pensamos que esta es más alta con este juego que sin duda intriga e interesa a todo el que trata de resolverlo.

5.2.7. CONCLUSIONES DEL ESTUDIO 3

En el Estudio 3 descrito en las anteriores secciones se plantearon los siguientes objetivos:

1. *Seleccionar algunos recursos relacionados con la probabilidad condicional e independencia y realizar una clasificación de los mismos*.
2. *Realizar un análisis de los objetos matemáticos puestos en juego en el uso didáctico* de algunos de los recursos seleccionados e *identificar algunos posibles conflictos semióticos* que puedan surgir durante su solución.
4. *Valorar los procesos matemáticos utilizados en el trabajo con los recursos* seleccionados y la *idoneidad didáctica de los mismos* en los cursos de formación de profesores.

Nuestra hipótesis sobre este estudio fue que *podríamos identificar una variedad de recursos que mostrasen la riqueza de objetos matemáticos relacionados con la probabilidad condicional.* Asimismo *pensábamos mostrar la idoneidad didáctica de estos recursos* para la formación de profesores en el campo de la probabilidad condicional.

Se han tratado de cumplir los objetivos mediante una búsqueda, clasificación y análisis de los distintos recursos. El resultado ha sido un índice bastante amplio de recursos que proporcionan ayudas de aprendizaje, plantean una visión diferente del término probabilidad condicionada, plantean problemas paradójicos, permiten la graficación, simulación y experimentación y proporcionan al estudiante un aspecto visual del que carecen los libros de texto.

Pensamos que un uso adecuado de estas ayudas, debidamente insertado en el proceso de aprendizaje, ayudará al futuro profesor o profesor en ejercicio a superar algunos de los errores y sesgos descritos en los Estudios 1 y 2. Se han cumplido asimismo los objetivos de valoración de los objetos y procesos matemáticos implícitos en el uso de los recursos, así como nuestra hipótesis inicial sobre las dificultades potenciales del trabajo con los mismos y su idoneidad didáctica. Todos estos resultados se han publicado en Batanero, Fernándes y Contreras (2009), Contreras, Díaz y Batanero (2009) y Contreras, Batanero y Fernándes (2010).

La disponibilidad de recursos libremente accesibles en Internet hace que la cultura y la ciencia se estén democratizando cada vez más, por lo que el aprendizaje se lleva a cabo no sólo en el aula tradicional (Galmacci, 2001). El uso de este tipo de recursos aumenta la motivación de los alumnos por el tema, ya que se presentan los conceptos de una forma más llamativa y permite al alumno adoptar un papel activo en su aprendizaje. Es por ello importante que el profesor tenga en cuenta estos recursos y los incorpore a su enseñanza. Sin embargo, un recurso didáctico por sí sólo no resuelve todos los problemas. Se plantea, así el reto de diseñar unidades didácticas para la enseñanza de la probabilidad condicional y la formación de profesores que incorpore estos recursos.

El análisis de las posibles soluciones (como hemos hecho) y los objetos matemáticos usados es útil en la formación de profesores, para mejorar su capacidad de análisis del componente didáctico y hacerles reflexionar sobre la complejidad del campo de la probabilidad.

5.3. PARADOJAS DE PROBABILIDAD COMO RECURSO DIDÁCTICO

5.3.1. INTRODUCCIÓN

Otra herramienta didáctica posible en la formación de profesores son las paradojas aparecidas en la historia de la probabilidad (Batanero, Contreras y Arteaga, 2009; Fernándes, Batanero, Contreras y Díaz, 2009; Batanero, Contreras, Fernándes y Ojeda, 2010). Puesto que el profesor tiene unos conocimientos sólidos en probabilidad elemental, para provocar su reflexión, hemos de buscar problemas, que, siendo aparentemente sencillos, puedan tener soluciones contraintuitiva. Es fácil encontrar estos problemas, pues en la historia de la probabilidad y la estadística muchos problemas sencillos resultaron difíciles de resolver, incluso para matemáticos famosos (Batanero, Henry y Parzysz, 2005).

Lesser (1998) indica que el uso inteligente de paradojas en la clase de matemáticas apoya una pedagogía constructivista, promoviendo un aprendizaje profundo a partir de las creencias previas y dando al profesor el papel de facilitador del aprendizaje. Falk y Konold (1992) afirman que el análisis de estas paradojas requiere, una conciencia de sus propios pensamientos, lo que es tan importante como el aprendizaje de la solución correcta y es un paso vital para alcanzar la capacidad matemática abstracta. Konold (1994) destaca el efecto motivador de los resultados sorprendentes, que anima a los estudiantes a explorar el problema más formalmente.

A continuación se describe el Estudio 4, orientado a seleccionar algunas paradojas relacionadas con la probabilidad condicional e independencia que permitan afrontar a los profesores con los sesgos encontrados en los Estudios 1 y 2. A continuación describimos los objetivos e hipótesis, metodología y resultados de este estudio.

5.3.2. OBJETIVOS E HIPÓTESIS DEL ESTUDIO 4

El objetivo general del Estudio 4 es recopilar y analizar algunas paradojas clásicas ligadas a la probabilidad condicional e independencia, que puedan servir para organizar situaciones didácticas dirigidas a la formación de profesores, tanto respecto al conocimiento matemático, como respecto al conocimiento didáctico en las facetas consideradas en este trabajo. Los objetivos específicos son los siguientes:

1. El primer objetivo es *seleccionar algunas paradojas cuya solución pueda abordarse con conocimientos elementales de probabilidad condicional e independencia*. Han de tratarse de problemas curiosos, que inciten el interés de los profesores y cuya solución no sea evidente. Deben requerir conocimientos de probabilidad condicional o independencia y tener potencial para organizar un debate en torno a las cuestiones matemáticas y didácticas, que contribuyan a aumentar el conocimiento del profesor.
2. El segundo objetivo es *realizar un análisis de los objetos matemáticos puestos en juego en las posibles soluciones* de las paradojas seleccionadas, similar al realizado en el Estudio 3 con los recursos en Internet e *identificar algunos posibles conflictos semióticos* que puedan surgir durante su solución.
3. El tercer objetivo es *valorar los procesos matemáticos utilizados en el trabajo con las paradojas* seleccionadas y la *idoneidad didáctica de este recurso* en los cursos de formación de profesores.

Hipótesis

La hipótesis que nos planteamos respecto a las paradojas seleccionadas es que su *análisis nos permitirá mostrar una variedad de objetos matemáticos involucrados en sus posibles soluciones, así como dificultades y conflictos potenciales en los estudiantes, que están relacionados con las dificultades descritas en las investigaciones previas que se analizaron en el capítulo 2 y los encontrados en los Estudios 1 y 2.*

Asimismo esperamos *mostrar la idoneidad didáctica de este recurso para la formación de profesores.* Las hipótesis se fundamentan en el marco teórico descrito en el capítulo 1, en el que se específica la complejidad del conocimiento matemático, y también en el análisis del objeto matemático "probabilidad condicional", llevado a cabo en el mismo capítulo.

Se espera que la riqueza de objetos matemáticos revelada en dicho análisis aparezca en las prácticas matemáticas necesarias en la solución de las paradojas por parte de posibles alumnos o profesores. Por otro lado, tanto en el capítulo 2, como en los Estudios 1 y 2, se describieron numerosos sesgos de razonamiento en relación a la probabilidad condicional e independencia. Pensamos que estos sesgos son los que explican muchas de las paradojas de probabilidad condicional y por ello aparecerán en el trabajo de los futuros profesores o profesores en ejercicio, precisamente debido al carácter contraintuitivo de las mismas.

5.3.3. MATERIAL Y MÉTODO

En el Estudio 4 hemos localizado, y posteriormente analizado, algunas paradojas de probabilidad condicional e independencia que cumplen las condiciones descritas en los anteriores apartados. La localización de estas paradojas se ha realizado a través de varios sistemas:

- A partir de la lectura de libros de probabilidad y artículos didácticos o de investigación que describen paradojas en probabilidad, como por ejemplo, los de Crandall y Greenfield (1986), Falk (1986), Székely (1986), Borovnick y Bentz (1991), Bohl, Liberatore y Nydick (1995), Borovcnik y Peard (1996) o Díaz y de la Fuente (2005).
- Mediante búsqueda directa en buscadores de Internet, utilizando palabras claves como "paradoja" o "problemas históricos", "intuiciones" y alguna de las siguientes

“probabilidad condicional” “probabilidad condicionada”, “independencia”, “conditional probability“, “Bayes”, “Independence” o “Bayes theorem”.

Una vez localizada una paradoja, se realiza un primer análisis de la misma, para ver si cumple los criterios especificados. En caso de interés potencial, se clasifica como paradoja asociada a la probabilidad condicional o paradoja asociada a la independencia. A continuación llevamos a cabo un análisis de dicha paradoja desde diferentes puntos de vista. En primer lugar, describimos la paradoja, proporcionando datos de su historia y referencias donde se pueda ampliar lo expuesto en este capítulo.

En segundo lugar se lleva a cabo un análisis matemático de una posible solución intuitiva correcta y de algunas soluciones incorrectas. Para el primer caso, analizamos los objetos matemáticos en juego y para el segundo los posibles conflictos semióticos que llevan a la solución incorrecta, describiendo con detalle las que se hayan descrito en las investigaciones previas, como por ejemplo en los trabajos de Díaz (2005) y Díaz y de la Fuente (2005) o hayan aparecido en los Estudios 1 o 2. Se construye una tabla de análisis semiótico, mostrando los objetos matemáticos requeridos en la solución correcta.

Para finalizar el análisis y al final del capítulo se estudian, globalmente para el conjunto de paradojas analizadas los procesos matemáticos requeridos así como la idoneidad didáctica del uso de estas paradojas en los cursos de formación. En lo que sigue, se presentan los resultados, deteniéndonos en la paradoja de Bertrand que es base para la organización del taller descrito en el capítulo 6. No pretendemos ser exhaustivos, debido a la gran cantidad de paradojas en la historia de la probabilidad, sino analizar algunas de las más conocidas, pues ello es suficiente para los objetivos del Estudio 4.

5.3.4. ANALISIS DE RESULTADOS

5.3.4.1. LA PARADOJA DE LA CAJA DE BERTRAND Y SUS VARIANTES

Comenzamos con el análisis detallado de esta paradoja, pues será la base del taller descrito en el Capítulo 6. Joseph Bertrand (1822-1900) fue un matemático francés del siglo XIX, cuyos campos de trabajo fueron la Teoría de los Números, la Geometría Diferencial, la Economía, la Termodinámica y la Teoría de las Probabilidades. En su libro "*Cálcul des probabilités*" publicado en 1888, describe numerosos ejemplos de problemas de probabilidades cuyo resultado es contrario a la intuición. Entre ellos el siguiente problema, que es ahora conocido como la "Paradoja de la caja de Bertrand":

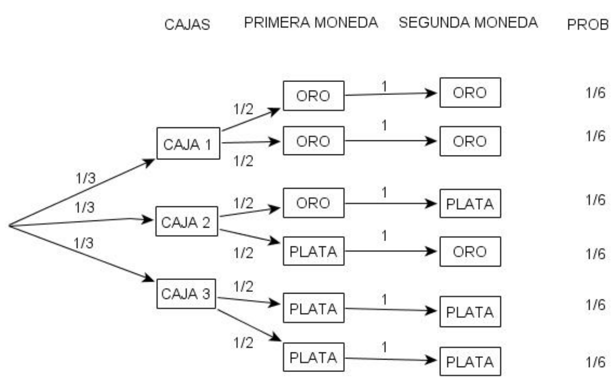
Tenemos tres cajas y cada caja tiene dos cajones con una moneda cada una: una caja contiene dos monedas de oro, otra caja dos monedas de plata, y la caja final con una de cada tipo. Después de elegir una caja al azar se toma un cajón al azar, y resulta por ejemplo que contiene una moneda de oro. ¿Cuál es la probabilidad de que la otra también sea de oro?

La solución que dan muchos estudiantes, se basa en el siguiente razonamiento. Después de elegir una caja al azar y retirar una moneda también al azar, si esta fuese una moneda de oro, sólo tenemos dos opciones: (a) que hayamos elegido la caja con dos monedas de oro; o (b) que hayamos elegido la caja con una moneda de oro y otra de plata, pues es imposible que sea la caja con dos monedas de plata. Por tanto, la probabilidad de que la otra moneda también fuese de oro es igual 1/2. Esta solución es incorrecta, pues la probabilidad de que la segunda moneda sea de oro es 2/3.

Solución intuitiva correcta

Una posible solución correcta del problema se obtiene intuitivamente con una adecuada notación al representar el espacio muestral mediante el diagrama en árbol. Como hemos supuesto, hay tres cajas. Una tiene dos monedas de oro (O, O), otra dos monedas de plata (P, P), y la última una moneda de oro y una moneda de plata (O, P). Los posibles resultados del experimento descrito se observan en la Figura 5.3.1, que es representa el espacio muestral de un experimento compuesto, junto con las probabilidades en cada rama del árbol y finales de cada suceso.

Figura 5.3.1. Solución intuitiva al problema de la caja de Bertrand



El primer experimento es elegir al azar una de las tres cajas disponibles, que tienen igual probabilidad de ser elegida (1/3). En un segundo paso, al abrir uno de los cajones se pueden encontrar los casos representados en la segunda división en ramas del árbol. El tercer experimento (última rama) es el tipo de moneda que queda en la caja, cuando se ha abierto uno de los cajones.

Sabemos que una de las monedas es de oro; por tanto quedan dos posibilidades (equiprobables); que se trate de la caja con dos monedas de oro o de la caja con una de oro y otra de plata. Para la caja (*OP*) la probabilidad de que la moneda resultante sea oro si la observada fue oro es cero, y si la observada es plata es 1. Como dicha caja sólo tiene dos posibilidades, la probabilidad de que la moneda que queda en el cajón sea de oro es 1/2 (segunda rama del árbol). Siguiendo un razonamiento similar, observamos que la probabilidad de que la moneda restante sea oro, para la caja (*OO*) es igual a 1. En consecuencia, sabiendo que una moneda es de oro, la probabilidad de que la otra sea de oro es doble (2/3) que la probabilidad de que sea de plata (1/3).

Tabla 5.3.1. Objetos matemáticos implícitos en la solución

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Predecir la segunda moneda, sabiendo la primera	- Determinar el material más probable de la segunda moneda
Lenguajes	- Verbal	- Explicación de la situación
	- Gráfico	- Diagrama en árbol - Representación icónica del juego
	- Numérico	- Expresión de números naturales y fracciones
	- Simbólico	- Expresiones de la probabilidad y de operaciones matemáticas, así como de sus resultados
Conceptos	- Experimento aleatorio	- Elegir una caja - Moneda que sale - Aceptar la predicción
	- Sucesos; espacio muestral	- Cajas 1, 2, 3 - Moneda que se elige - Moneda que queda en la caja - Aceptar/no aceptar
	- Experimento compuesto	- Composición de los experimentos anteriores
	- Sucesos en el experimento compuesto	- Producto cartesiano de los espacios muestrales anteriores (*OO*) (*OP*) (*PP*)
	- Probabilidad clásica	- Proporción de casos favorables a posibles.
	- Probabilidad condicional	- Proporción de ocurrencia suceso respecto a la ocurrencia de otro
Procedimientos	- Cálculo probabilidades intuitivo	- Aplicar reglas intuitivas
	- Representación gráfica	- Diagrama; esquema
Propiedades	- Diferencia probabilidad condicionada y simple	- Restricción del espacio muestral
	- Axioma de unión	- La probabilidad de la suma es suma de probabilidades

	- Regla de la suma	- Suma de probabilidades
	- Restricción del espacio muestral en la probabilidad condicionada	- Si una moneda es de oro, la caja no puede ser (*PP*)
Argumento	- Razonamiento deductivo	- Demostración de la solución

En la Tabla 5.3.1 se presentan los objetos matemáticos que se han utilizado en la solución correcta descrita. Observamos que el trabajo con este problema, es abordable en la educación secundaria y formación de profesores y que permite trabajar con los objetos matemáticos ligados a la probabilidad condicional analizados en el capítulo 1. Por tanto, este sencillo pero contra-intuitivo juego puede ser utilizado en el aula, ya que ilustra algunos principios básicos de la enseñanza de la teoría de probabilidades, entre ellos los axiomas de Kolmogorov. Por supuesto es posible también un trabajo con mayor nivel de formalización, lo que implicaría un conjunto diferente de objetos matemáticos usados (diferente configuración epistémica).

Dificultades posibles de los estudiantes

Al igual que en el caso de los recursos en Internet, la complejidad del problema, aparentemente simple se muestra en el análisis de los objetos matemáticos que se analizan en la Tabla 5.3.1. En la literatura relacionada con este problema se han descrito varias soluciones erróneas, relacionadas con una deficiente intuición, que comentamos a continuación. Estas soluciones pueden ser debidas a errores en el proceso de representación-interpretación o bien a la atribución de propiedades que no tienen a ciertos objetos o situaciones, como vemos en los casos que siguen:

- *No percepción de la independencia.* Un primer problema se produce porque *no se percibe la dependencia de los sucesivos experimentos* (elegir una caja) y (elegir una moneda). Es decir, o bien no se visualiza la estructura del experimento compuesto o se suponen los sucesivos experimentos independientes, atribuyendo una propiedad (independencia) que no tienen los experimentos. Pensamos que esto es un conflicto semiótico debido a fallo de interpretación de la descripción verbal, que no es más que la representación del experimento real. Este error de razonamiento es explicado por Falk (1986), mediante la “falacia del eje de tiempos” que consiste en que las personas creen erróneamente que una información actual (la moneda mostrada) no puede afectar a un suceso que ocurrió con anterioridad a la misma (caja elegida). Esta falacia se mostró en una alta proporción de participantes en el Estudio 2,

- Incorrecta *percepción del espacio muestral*. Otra posibilidad de error en este problema es una incorrecta enumeración del espacio muestral; mostrando un fallo en pasar de la idea espacio muestral (intensivo) al espacio muestral concreto (extensivo) o lo que es lo mismo, fallo en la particularización del espacio muestral en este experimento. La intuición sugiere que, una vez elegida la caja, y quitando la (*PP*), sólo quedan dos posibilidades equiprobables. Por tanto, hay un 50 % para las otras dos cajas. En este razonamiento se está realizando una incorrecta enumeración del espacio muestral al calcular la probabilidad condicionada, otro sesgo descrito por Gras y Totohasina (1995).

Paradoja de las tarjetas

Son múltiples las versiones de la paradoja de la caja de Bertrand encontradas en la literatura. Batanero, Godino y Roa (2004) describen una actividad que sirve para comparar la concepción frecuentista y Laplaciana de la probabilidad y reflexionar sobre los conceptos de experimentos dependientes y probabilidad condicional. Utilizan el siguiente enunciado de la denominada “paradoja de las tres tarjetas”:

Tomamos tres tarjetas de la misma forma y tamaño. Una es de color azul en ambos lados, la segunda es de color rojo en ambos lados y la tercera es azul de un lado y roja por el otro. Ponemos las tres tarjetas en una caja, y agitamos la caja, antes de seleccionar una tarjeta al azar. Después de seleccionar la tarjeta se muestra uno de los lados. El objetivo del juego es adivinar el color de la cara oculta. Repetimos el proceso, poniendo la tarjeta de nuevo en la caja antes de cada nueva extracción. Hacemos predicciones sobre el color del lado oculto y se gana un punto cada vez que nuestra predicción es correcta.

Esta variante de la paradoja se utiliza en el taller didáctico analizado en el capítulo 6, donde analizamos posibles soluciones con diferente nivel de formalización. Gardner (1982) denomina a esta paradoja como la estafa de las tres tarjetas, y propone un par de variantes. La primera versión consiste en extraer una tarjeta de un sombrero, y si esta tiene una marca roja, el estafador apuesta con la víctima una cierta cantidad de dinero a que la marca de la otra parte es también roja. La víctima convencida de que la apuesta es justa aceptará sin pensarlo ya que creerá que tiene igual posibilidades de ganar que de perder, pero el timador gana dinero a largo plazo, ya que ganará 2/3 de las veces que realice el juego. La segunda versión de esta paradoja, propuesta por Gardner tiene el siguiente enunciado:

Tres hombres, sentados formando un triángulo, mantienen los ojos cerrados mientras juegan. Se les pone un gorro rojo o un gorro negro en sus cabezas. Se les propone que abran los ojos a la vez y que se fijen en el gorros del de su derecha, pero no el suyo. Tienen que adivinar el color de su propio gorro. Tan pronto como esté seguro del color de su gorro, ha de decirlo.

Al igual que ocurre con las tarjetas, el jugador al observar un gorro de un determinado color tiende a pensar que hay dos sobreros de distintos colores, el suyo y el que no ve, por lo que tiende a asignar probabilidad de 1/2 a cada uno de los casos.

Dilema del prisionero

Otra variante de la paradoja de la caja de Bertrand es el "dilema del prisionero" (Hardin, 1968)**,** problema que presenta el siguiente enunciado:

Tres prisioneros esperan encarcelados su juicio. Se les informa que a uno de ellos se les condenará a muerte y que a los otros dos se les liberará. Cada preso piensa en las posibilidades que tiene de salvase, pero el juez le dice al primer preso que el tercero se salva y que si quiere puede intercambiar su suerte con el segundo. ¿Qué debe hacer el primer prisionero?

Una solución intuitiva de este dilema sería comprobar las posibilidades, es decir crear el espacio muestral de las posibilidades a priori sin tener en cuenta la información proporcionada por el juez. Antes de conocer que el tercer preso se salva, las posibilidades de vida y muerte de los prisioneros serían las indicadas en la Tabla 5.3.2. Por lo tanto, la probabilidad de que el primer preso muera es de 1/3 ya que sólo moriría en el supuesto 1. Pero como el juez le indica que el tercer preso se salva, aparentemente sólo quedan dos opciones: que muera el primero o segundo prisionero. Por lo que vemos la respuesta más usual sería decir que ambos tienen la misma probabilidad (1/2) de morir, y no merece la pena elegir la opción que proporciona el juez de intercambiar su futuro por el preso segundo. Sin embargo, este razonamiento es incorrecto.

Tabla 5.3.2. Supuestos posibles en el dilema de los tres prisioneros

	1º preso	2º preso	3º preso
Supuesto 1	Muere	Libre	Libre
Supuesto 2	Libre	Muere	Libre
Supuesto 3	Libre	Libre	Muere

Solución correcta

Una solución correcta se obtendría comparando las probabilidades de que mueran el primer y segundo preso, sabiendo que se salva el tercero.

Para calcular *P((1º muera)/(3º se salve))* habría que aplicar la fórmula de la probabilidad condicional, es decir:

$$P((1^o\, muera)/(3^o\, se\ salve)) = \frac{P((1^o\, muera) \cap (3^o\, se\ salve))}{P(3^o\, se\ salve)}.$$

Como sabemos que el tercer prisionero se salva, la probabilidad pedida es igual a la del numerador, $P((1^o\, muera) \cap (3^o\, se\ salve))$. Utilizando la fórmula de la probabilidad compuesta: $P((1^o\, muera) \cap (3^o\, se\ salve)) = P(1^o\, muera) \cdot P(3^o\, se\ salva \mid 1^o\, muera)$. Al salvarse siempre el tercer prisionero: $P(3^o\, se\ salva \mid 1^o\, muera) = P(3^o\, se\ salva) = 1$, por tanto, $P((1^o\, muera) \cap (3^o\, se\ salve)) = P(1^o\, muera)$ =1/3. Aplicando la fórmula de la probabilidad condicional, tenemos:

$$P((1^o\, muera)/(3^o\, se\ salve)) = \frac{P((1^o\, muera) \cap (3^o\, se\ salve))}{P(3^o\, se\ salve)} = \frac{1/3}{1} = 1/3.$$

La probabilidad de que muera el segundo preso, sabiendo que se salva el tercero, sería la complementaria de la anterior, pues sabemos que o bien el segundo o el tercero han de morir, por lo que la probabilidad sería:

$$P((2^o\, muera) \mid (3^o\, se\ salve)) = 1 - P((1^o\, muera) \mid (3^o\, se\ salve)) = 2/3.$$

Luego al primer preso no le conviene intercambiar su suerte con el segundo ya que así tendría el doble de posibilidades de morir, lo cual es paradójico. Como se muestra en la Tabla 5.3.3 la resolución es bastante compleja, debido a la gran cantidad de objetos matemáticos que se manejan, incluyendo varias fórmulas de cálculo de probabilidad compuesta y condicional. Como vemos, pequeñas variantes en el enunciado de la paradoja de Bertrand aumentan la complejidad de la solución, de modo que el contexto es una variable de tarea importante en este problema.

Posibles dificultades

Al igual que en el problema de la caja de Bertrand las soluciones erróneas se deben a la no percepción de la independencia de los sucesos, que apareció tanto en el

Estudio 1 como en el Estudio 2, y a la incorrecta percepción del espacio muestral. El primer problema se produce cuando no se percibe la dependencia de los sucesivos experimentos (plantear las posibilidades de morir) y (elegir intercambiarse o no). No se visualiza la estructura del experimento compuesto, otorgándole a los experimentos la propiedad de independencia que no tienen.

El otro error es debido a la incorrecta enumeración del espacio muestral, lo que es frecuente en el cálculo de la probabilidad condicional, según Gras y Totohasina (1995), mostrando un fallo en pasar de la idea espacio muestral (intensivo) al espacio muestral concreto (extensivo). La intuición nos dice que, una vez conocido que el tercer preso sobrevive, sólo quedan dos posibilidades equiprobables. Por tanto, hay un 50% de posibilidades de morir, es decir se está realizando una incorrecta enumeración del espacio muestral al calcular la probabilidad condicionada.

Tabla 5.3.3. Objetos matemáticos implícitos en la solución

Tipo	Objetos matemáticos	Significado en la situación
Problema	- Elegir una opción, intercambiarse o no por el segundo preso	- Determinar la mayor probabilidad de supervivencia
Lenguajes	- Verbal	- Explicación de la situación
	- Gráfico	- Tabla de resultados
	- Simbólico	- Expresar sucesos, probabilidades
	- Numérico: Probabilidades	- Calculo de las diferentes probabilidades
	- Icónico	- Iconos que representan los sucesos y resultados
Conceptos	- Experimento aleatorio	- Elegir el preso - Elegir intercambiarse o no - Salvar la vida
	- Sucesos; espacio muestral	- Preso 1, 2, 3 - Morir/Libre
	- Experimento compuesto	- Composición de los experimentos anteriores
	- Sucesos en el experimento compuesto	- Producto cartesiano de los espacios muestrales anteriores (*M1, L2, L3*), (*L1, M2, L3*), (*L1, L2, M3*)
	- Intersección de sucesos	- Conjunto común de sucesos
	- Probabilidad clásica	- Proporción de casos favorables a posibles
	- Probabilidad condicional	- Proporción de ocurrencia suceso respecto a la ocurrencia de otro
	- Variable aleatoria	- Número de preso que se salvará
	- Complementario	- Suceso contrario a otro
Procedimientos	- Regla de la probabilidad condicionada	- Descomposición de la probabilidad condicionada como cociente de la conjunta y la simple
	- Cálculo de probabilidad compuesta	- Regla del producto para sucesos dependientes
	- Cálculo del complementario	- La probabilidad del complementario es uno menos la probabilidad del suceso
Propiedades	- Diferencia probabilidad condicionada y simple	- Restricción del espacio muestral
	- Axioma de unión	- La probabilidad de la suma es suma de probabilidades

	- Probabilidad condicionada en función de la conjunta y la simple	- $P(A/B) = \frac{P(A \cap B)}{P(B)}$
	- Regla del complementario	- $P(\overline{A}) = 1 - P(A)$
	- Regla de la unión	- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
	- Regla del producto	- $P(A \cap R) = P(A)\ P(R/A)$
	- Regla de la independencia	- $P(A/B) = P(A)$
Argumento	- Razonamiento deductivo	- Demostración de la solución
	- Razonamiento empírico	- Comparar aciertos con distintas estrategias

Paradoja del niño o niña

Otra versión de la paradoja de Bertrand, de gran importancia por el número de investigaciones pedagógicas que la tratan, es la paradoja del niño o niña, también conocida como el problema de los dos hijos, los niños de Smith o el problema de la señora Smith (Gardner, 1959b). El enunciado de la paradoja se describió mediante dos opciones de la siguiente manera:

El Sr. Jones tiene dos hijos. El hijo mayor es una niña. ¿Cuál es la probabilidad de que ambos sean niñas?
El Sr. Smith tiene dos hijos. Al menos uno de ellos es un niño. ¿Cuál es la probabilidad de que ambos hijos sean niños?

La paradoja está planteada para comprobar si el problema tiene una configuración similar para las dos preguntas, ya que la respuesta intuitiva es 1/2 para ambas, pues la pregunta lleva al lector a creer que hay dos sucesos igualmente probables para el sexo del segundo hijo (es decir, niño y niña), y que la probabilidad de estos resultados no está condicionada a la información dada.

Esta paradoja ha estimulado una gran controversia. En primer lugar, el espacio muestral de todos los eventos posibles puede ser fácilmente enumerado: (*HH*, *HM*, *MH*, *MM*). En segundo lugar, se supone que estos resultados son igualmente probables. Los supuestos son los siguientes:

1. Que cada niño es hombre o mujer.
2. El sexo de cada hijo es independiente del sexo del otro.
3. Que cada niño tenga la misma probabilidad de ser varón como de ser mujer.

Estos supuestos no son totalmente ciertos ya que la proporción de niños y niñas no es exactamente de 50:50, y (entre otros factores) la posibilidad de gemelos idénticos significa que la determinación del sexo no es totalmente independiente. Pero se asumen como verdaderos. En la primera cuestión, el Sr. Jones tiene dos hijos y el mayor es una niña; se nos pide la probabilidad de que ambos hijos sean niñas. En el espacio muestral del problema encontramos que hay cuatro eventos igualmente probables, pero que dos de ellos no cumplen la condición de ser el hijo mayor chica (Tabla 5.3.4).

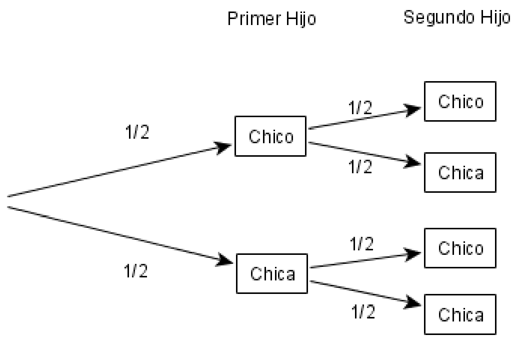
Tabla 5.3.4. Espacio muestral de la variable aleatoria

Hijo mayor	Hijo menor
Chica	Chica
Chica	Niño
~~Niño~~	~~Chica~~
~~Niño~~	~~Niño~~

Puesto que las dos posibilidades del espacio muestral nuevo (*MH*, *MM*) son igualmente probables, y sólo uno de los dos incluye dos niñas, la probabilidad de que el hijo más pequeño sea también una chica es 1/2.

En la segunda cuestión, se indica que el Sr. Smith tiene dos hijos y al menos uno de ellos es un niño y se pregunta la probabilidad de que ambos hijos sean niños. En principio esta cuestión es idéntica a la primera, excepto que en lugar de especificar que el hijo mayor es un niño, se dice que al menos uno de ellos es un niño. La primera impresión (errónea) es que como ya tiene un hijo varón, la probabilidad de que el otro también lo sea es de 1/2 ya que tiene un caso favorable de dos posibles.

Figura 5.3.2. Posibles resultados del experimento



Solución intuitiva correcta

Si suponemos que esta información se obtuvo al considerar todos los niños, como en el primer apartado, vemos que hay cuatro posibles eventos en el espacio muestral. Tres de estas familias reúnen la condición necesaria y suficiente de tener al menos un niño. El conjunto de posibilidades viene dado en la Tabla 5.3.5, donde vemos que hay un caso favorable de tres. Para responder a la pregunta, la hipótesis fundamental es cómo fue seleccionada la familia del Sr. Smith y cómo se obtuvo la información. Si no se sabe nada, son considerados todos los sucesos en los que intervienen niños, la respuesta a la cuestión 2 sería 1/3. En la Tabla 5.3.6 se presentan los objetos matemáticos implícitos en la solución.

Tabla 5.3.5. Espacio muestral de la variable aleatoria

Hijo mayor	Hijo menor
~~Chica~~	~~Chica~~
Chica	Niño
Niño	Chica
Niño	Niño

Tabla 5.3.6. Objetos matemáticos implícitos en la solución

Tipo	Objetos matemáticos	Significado en la situación
Problema	- Conocer la probabilidad de tener dos hijos o hijas	- Calcular probabilidades
Lenguajes	- Verbal	- Explicación de la situación
	- Gráfico	- Diagrama en árbol
	- Simbólico	- Expresar sucesos, probabilidades
	- Numérico: Frecuencias	- Resultados del experimento
	- Icónico	- Iconos que representan los sucesos y resultados
Conceptos	- Experimento aleatorio	- Ser niño o niña
	- Sucesos; espacio muestral	- Niño/niña, niña/niña, niña/niño, niño/niño
	- Experimento compuesto	- Composición de los experimentos anteriores
	- Sucesos en el experimento compuesto	- Producto cartesiano de los espacios muestrales anteriores
	- Intersección de sucesos	- Conjunto común de sucesos.
	- Probabilidad clásica	- Proporción de casos favorables a posibles.
	- Probabilidad condicional	- Proporción de ocurrencia suceso respecto a la ocurrencia de otro
Procedimientos	- Cálculo probabilidades intuitivo	- Aplicar reglas intuitivas
	- Cálculo probabilidades formal	- Aplicar reglas de cálculo formal
	- Representación gráfica	- Diagrama; esquema
Propiedades	- Diferencia probabilidad condicionada y simple	- Restricción del espacio muestral
	- Regla del producto	- $P(A \cap R) = P(A)\ P(R/A)$
	- Teorema probabilidad total	- Aplicar a la situación
Argumento	- Razonamiento deductivo	- Demostración de la solución
	- Razonamiento empírico	- Comparar aciertos con distintas estrategias

Posibles dificultades

Esta paradoja tiene valor pedagógico, ya que ilustra una aplicación interesante del comportamiento de la probabilidad condicional. Fox y Levav (2004) utilizaron el problema para analizar como los alumnos estiman las probabilidades condicionales. El estudio encontró que el 85% de los participantes respondieron correctamente al primer problema, mientras que sólo el 39% respondió correctamente a la segunda pregunta. Los autores alegaron que la razón se debe a la utilización ingenua de la heurística de equiprobabilidad (Lecoutre, 1992), detectado entre los futuros profesores, tanto en el Estudio 1, como en el Estudio 2.

Otro razonamiento erróneo es suponer que la familia fuese seleccionada al azar y se pidiese calcular la probabilidad de que los dos hijos fuesen niños, es decir, no se tiene en cuenta la información de que uno es chico, lo que supone confusión entre probabilidad condicional y conjunta, sesgo descrito por Pollatsek, Well, Konold y Hardiman (1987), y Einhorn y Hogarth (1986) y encontrado en los Estudios 1 y 2. Como observamos en la Figura 5.3.4.5, la probabilidad de que el primer hijo fuese chico y el segundo también es de 1/4, ya que:

$$P(1^{o}\,Chico \cap 2^{o}\,Chico) = P(2^{o}\,Chico \mid 1^{o}\,Chico)P(1^{o}\,Chico) = \frac{1}{2}\frac{1}{2} = \frac{1}{4}$$

5.3.4.2. OTRAS PARADOJAS LIGADAS A LA INDEPENDENCIA Y A LA PROBABILIDAD CONDICIONAL

En esta sección, analizamos otras paradojas que cumplen los objetivos de este estudio, sin pretender, como se ha dicho, ser exhaustivos.

Paradoja del cumpleaños

En sentido estricto, este problema no es una paradoja, ya que no tiene contradicción lógica en su solución. Pero la solución matemática contradice la intuición común. Este problema y su solución fueron propuestos por Schnabel (1938) en la teoría de estimación del total de población de peces en un lago, bajo el nombre *Captura-recaptura estadística*. Una formulación elemental de la paradoja es la siguiente:

Si hay 23 personas reunidas, hay una probabilidad del 50,7% de que al menos dos personas de ellas cumplan años el mismo día. Para 60 o más personas la probabilidad es mayor del 99%.

Una respuesta errónea es pensar que si un año tiene 365 días (366 si es bisiesto) haría falta por lo menos un grupo de personas equivalente a la mitad de los días de un año (183 personas si redondeamos) para alcanzar una probabilidad de al menos un 50%. Se supone que con 23 personas la probabilidad de ocurrencia de la afirmación debería de ser mucho menor.

Solución correcta

Una solución correcta de este problema es la siguiente: Calculemos la probabilidad aproximada de que en una habitación de n personas, ninguna cumpla años el mismo día, desechando los años bisiestos. Esta probabilidad viene dada por:

$$p=\frac{365}{365}\cdot\frac{364}{365}\cdot\frac{363}{365}\cdots\frac{365-n+1}{365}$$

En la anterior solución, calculamos probabilidades simples de que una segunda persona no pueda tener el mismo cumpleaños que el primero (364/365), que una tercera persona no puede tener el mismo cumpleaños que las dos primeras (363/365) y así sucesivamente para un n dado. La probabilidad del suceso complementario *1-p* en este caso sería la probabilidad que al menos dos personas tengan el mismo día de cumpleaños. Puesto que $1-p=1-\left(\frac{364}{365}\right)^n$ para *n*=23 se obtiene una probabilidad *p*=0.493 por lo que *1-p* es alrededor de 0,507.

Tabla 5.3.7. Objetos matemáticos implícitos en la solución

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones problemas	- Calcular la probabilidad de que dos personas tengan igual cumpleaños	- Solución a la paradoja
Lenguajes	- Verbal	- Explicación de la situación
	- Simbólico	- Expresar sucesos, probabilidades, fracciones y operaciones con fracciones
	- Numérico: Fracciones	- Probabilidades en forma de fracción
Conceptos	- Experimento aleatorio	- Cumpleaños de una persona
	- Sucesos; espacio muestral	- Personas 1, 2, 3, …23 - Fecha del cumpleaños (1 a 365) - Mismo cumpleaños/distinto cumpleaños
	- Experimento compuesto	- Composición de los experimentos anteriores
	- Sucesos en el experimento compuesto	- Producto cartesiano de los espacios muestrales anteriores
	- Probabilidad clásica	- Proporción de casos favorables a posibles
	- Probabilidad condicional	- Proporción de ocurrencia suceso respecto a la ocurrencia de otro
	- Intersección de sucesos	- Conjunto común de sucesos
	- Axiomas probabilidad	- Explicitación de los axiomas
Procedimientos	- Cálculo de probabilidad intuitivo	- Aplicar reglas intuitivas
	- Cálculo de probabilidad formal	- Aplicar reglas de cálculo formal
Propiedades	- Teorema de la probabilidad total	- Aplicar a la situación
	- Regla del producto	- $P(A\cap R)=P(A)P(R/A)$
	- Probabilidad del suceso complementario	- La suma de la probabilidad de un suceso y su complementario es la unidad

La clave para entender la paradoja del cumpleaños es que hay muchas probabilidades de encontrar parejas que cumplan años el mismo día. Específicamente, entre 23 personas, hay (23×22)/2=253 pares de persona, cada uno de ellos un candidato potencial para cumplir la paradoja. Sin embargo, si una persona entrase en una habitación con 22 personas, la probabilidad de que cualquiera cumpla años el mismo día que quien entra, es mucho más baja. Esto es debido a que ahora sólo hay 22 pares posibles. En la Tabla 5.3.7 se presenta el análisis semiótico de esta solución.

Dificultades posibles de los estudiantes

A priori, podemos identificar las siguientes posibles causas de este tipo de razonamiento.

- Razonamiento basado en el principio del palomar o principio de Dirichlet que parte de que para tener una probabilidad del 100%, hay que tener 365 personas, por ejemplo: Si se toman trece personas, al menos dos habrán nacido el mismo mes.
- Otro posible problema es la incorrecta *percepción del espacio muestral* descrito por Gras y Totohasina (1995) y mostrada en algunos participantes de los Estudios 1 y 2. La intuición nos dice que necesitaríamos al menos una cantidad de personas proporcional a la mitad de días del año para alcanzar el 50% de posibilidades sin tener en cuenta la combinación de parejas, 253 pares, que se pueden hacer con 23 elementos. El error se comete al plantear el problema como una probabilidad simple sin tener en cuenta la información disponible para crear el espacio muestral compuesto.

Otra explicación posible del error es un *deficiente razonamiento combinatorio de los estudiantes*. Según Dubois (1984), podemos clasificar las configuraciones combinatorias simples en tres modelos diferentes: *Selección*, que enfatiza la idea de muestreo, *colocación*, relacionado con el concepto de aplicación y *partición* o división de un conjunto en subconjuntos. El problema, en concreto, sería un problema combinatorio de colocación que tiene mucha dificultad para los estudiantes, pues no se enseña explícitamente en secundaria. En un problema de colocación se diferencian dos parámetros: (a) Los objetos a colocar (en este caso las 23 personas) y las celdas donde se colocan (en el ejemplo, 365 días del año). Roa (2000) y Navarro-Pelayo (1994) definen el error titulado "confusión en el tipo de celdas", consistentes en pensar que podríamos distinguir celdas idénticas o que no es posible diferenciar las celdas distinguibles. En nuestro caso, las celdas son distinguibles, pues cada persona se puede “colocar” en uno de los 365 días. Los alumnos pueden interpretar que las 22 personas siguientes a la primera han de tener alguna de ellas el cumpleaños el mismo día de la primera, por tanto, considerar celdas idénticas.

Paradoja de Simpson

Simpson (1951) publicó el problema posteriormente conocido como paradoja de Simpson, o efecto de Yule-Simpson, ya que anteriormente este fue descrito por el británico G. Udny Yule. Este fenómeno muestra que en determinados casos se produce un cambio en la asociación o relación entre un par de variables, ya sean cualitativas o cuantitativas, cuando se controla el efecto de una tercera variable. Un ejemplo de una situación concreta donde se aplica este efecto viene dado por el siguiente enunciado:

Supongamos que se quiere realizar un estudio comparativo sobre dos hospitales de una determinada ciudad, que llamaremos respectivamente A y B. Para ello la siguiente tabla muestra el número de muertes tras las operaciones realizadas en cada uno de los hospitales.

Tabla 5.3.8. Tabla de contingencia de dos variables

	Hospital A	Hospital B
Mueren	63	16
Sobreviven	2037	784
Total de pacientes operados	2100	800

Si analizamos los datos de la Tabla 5.3.8, se puede observar que en el hospital *A*, del total de los pacientes que se someten a una operación muere el 3% ($\frac{63}{2100}$) y en el hospital *B* del total de los pacientes operados muere un 2% ($\frac{16}{800}$). Por lo que inicialmente, estos resultados nos podrían llevar a pensar que el hospital más seguro para someterse a una operación sería el hospital *B*.

La paradoja aparece cuando se tiene en cuenta otra variable que influye en las anteriores, en este caso vamos a utilizar la variable: “estado de salud de los pacientes antes de la operación”. Para este ejemplo vamos a segregar los datos en dos tablas de contingencia (5.3.9 y 5.3.10) donde se tiene en cuenta el estado de salud de los pacientes antes de ser hospitalizados, y se clasifican los pacientes en dos grupos, los que tenían buena salud y los que no.

Tabla 5.3.9. Tabla de contingencia de tres variables (I)

Pacientes con buena salud	Hospital *A*	Hospital *B*
Mueren	6	8
Sobreviven	594	592
Total de pacientes operados	600	600

Tabla 5.3.10. Tabla de contingencia de tres variables (II)

Pacientes con salud delicada	Hospital *A*	Hospital *B*
Mueren	57	8
Sobreviven	1443	192
Total de pacientes operados	1500	200

Si analizamos la tasa de mortalidad de personas que han sido operadas en ambos hospitales teniendo en cuenta el estado de salud antes de la operación, se tiene que, del total de paciente con buena salud que se operaron en el hospital *A*, fallecieron un 1% ($\frac{6}{600}$) y de los que se operaron el *B*, no sobrevivieron un 1,3% ($\frac{8}{600}$). Por lo tanto bajo estas condiciones parece mejor el hospital *A*. Estudiemos ahora el caso de pacientes con salud delicada: Del total de los que se operaron en el hospital *A* un 3,8% ($\frac{57}{1500}$) murieron y en el caso del hospital *B* el 4% ($\frac{8}{200}$), por lo que también podríamos concluir que el hospital A tiene una tasa de mortalidad de pacientes recién operados de salud delicada mejor que la del hospital *B*.

Es aquí donde se presenta la paradoja, ya que sin tener en cuenta la variable estado de salud de los pacientes, el hospital *B* tenía mejor tasa de supervivencia, y sin embargo si separamos los pacientes en dos grupos, los que tenían buena salud y los de salud delicada, en cada uno de estos dos grupos, el hospital *A* obtuvo menor tasa de mortalidad tras las operaciones.

Solución correcta

La paradoja se resuelve observando la frecuencia total de pacientes que se opera en uno y otro hospital en las dos últimas tablas. La mayoría de los pacientes que se operan en el hospital *A* tienen una salud delicada y es de esperar que este tipo de pacientes tenga mayor dificultad en sobrevivir tras una operación, por lo que la tasa de mortalidad en dicho hospital es mayor que en el hospital *B*, que tiene menos pacientes y la mayoría gozan de buena salud. En la Tabla 5.3.11 se presenta el análisis semiótico de esta solución.

Dificultades posibles de los estudiantes

Un razonamiento erróneo se comete cuando se interpretan los datos sin considerarlos globalmente, es decir, sin tener en cuenta todas las variables que aparecen en la situación. Posibles explicaciones de este razonamiento son las siguientes:

- Al calcular la probabilidad no tenemos en cuenta todos los elementos del espacio muestral. Según tomemos las variables que afectan a la asociación entre "supervivencia" y "hospital", los sucesos toman una probabilidad u otra (Gras y Totohasina, 1995). Supone un fallo en pasar de la idea espacio muestral (intensivo) al espacio muestral concreto (extensivo), que apareció en algunos participantes en los Estudios 1 y 2.

Tabla 5.3.11. Objetos matemáticos implícitos en la solución

Tipo	Objetos matemáticos	Significado en la situación
Problema	- Elección del hospital	- Determinar cuál es el mejor hospital para operarse
Lenguajes	- Verbal	- Explicación de la situación
	- Simbólico	- Expresar sucesos, probabilidades
	- Numérico: Porcentajes	- Probabilidades en porcentaje
	- Numérico: Frecuencias	- Número de pacientes con cada condición
	- Icónico	- Iconos que representan los sucesos y resultados
Conceptos	- Experimento aleatorio	- Elegir hospital - Sobrevivir/No sobrevivir - Estado de salud
	- Sucesos; espacio muestral	- Hospital *A* y *B* - Sobrevivir/No sobrevivir - Salud delicada/No salud delicada
	- Experimento compuesto	- Composición de los experimentos anteriores
	- Sucesos en el experimento compuesto	- Producto cartesiano de los espacios muestrales anteriores
	- Frecuencia relativa	- Éxitos / número experimentos
	- Intersección de sucesos	- Conjunto común de sucesos.
	- Unión de sucesos	- Sucesos en uno u otro conjunto
	- Probabilidad clásica	- Proporción de casos favorables a posibles.
	- Axiomas probabilidad	- Explicitación de los axiomas
	- Probabilidad condicional	- Proporción de ocurrencia suceso respecto a la ocurrencia de otro
Procedimientos	- Cálculo de probabilidad formal	- Aplicar reglas de cálculo formal
	- Cálculo de probabilidad frecuencial	- Estimar la probabilidad mediante la frecuencia
	- Comparación de probabilidades	- Comparar los resultados para decidir cuál es mejor
Propiedades	- Diferencia probabilidad condicionada y simple	- Restricción del espacio muestral
	- Regla de la suma	- Si *A* y *B* son independientes $P(A\cup B)=P(A)+P(B)$
	- Regla del producto	- $P(A\cap R)=P(A)P(R/A)$
	- Axioma de unión	- La probabilidad de la suma es suma de probabilidades

Argumento	- Razonamiento deductivo	- Demostración de la solución
	- Razonamiento empírico	- Comparar supervivencia con distintas estrategias

- En relación al estudio de la asociación entre las dos variables, se olvida una condición necesaria en dicho estudio que es el aislamiento de otras variables espúreas. Esto quiere decir, que los pacientes que asisten a cada hospital, en el ejemplo, debieran suponerse elegidos al azar (lo cual evidentemente no se cumple, pues la tasa de pacientes con buena salud es muy diferente en cada uno de los hospitales). Si existe una variable no controlada que afecta en forma desigual a los dos grupos, el valor real del coeficiente de asociación entre dos variables puede aumentar, disminuir o incluso cambiar de signo.

La Paradoja de los dos sobres

La "paradoja de los dos sobres" también conocida como "paradoja del intercambio" o "problema de las dos billeteras" fue propuesta por Kraitchik (1953). Se recoge también en Gardner (1982). Una formulación sencilla de esta paradoja es la siguiente:

Dos personas, igualmente ricas, se reúnen para comparar el contenido de sus carteras. Cada uno ignorando el contenido de las dos carteras. El juego es el siguiente: cualquiera que tenga menos dinero recibe el contenido de la otra cartera (en el caso de que las cantidades sean iguales, no hay intercambio). ¿Es un juego equitativo? ¿Quién tiene ventaja?

Nalebuff (1989) presenta la variante del juego mediante dos sobres, que es la forma en que la paradoja se ha presentado con mayor frecuencia desde entonces.

Un benefactor pone una cierta cantidad de dinero, x en un sobre; y pone el doble de esa cantidad, 2x, en otro sobre. La cantidad x, así como la identificación del sobre que tiene la cantidad mayor, las desconoces. Para hacerlo más incierto, el benefactor selecciona al azar uno de los sobres y te lo regala. El contenido de ese sobre es tuyo. ¿Lo cambias o te quedas con el tuyo?

El razonamiento que habitualmente nos planteamos y puede conducir a error es el siguiente: Como sabemos, en uno de los sobres hay una cantidad x y por tanto el otro sobre puede tener o bien $2x$ o bien $x/2$. Elegimos un sobre (la probabilidad de elegir un sobre y de que tengamos una de las cantidades es de 1/2). Calculamos la esperanza matemática de la cantidad que contiene el otro sobre, es decir el valor esperado en función de los datos de los que dispongo. Por definición, la esperanza será la suma de las probabilidades de ocurrencia por la cantidad de cada uno, es decir: $E[x] = 0,5 \cdot 2x + 0,5 \cdot x / 2 = 1,25x$.

Ya que suponemos que *x* es un valor mayor que la unidad, *1,25x* es mayor que *x*, al concursante le vendría bien elegir el otro sobre ya que la cantidad esperada es mayor que la posee. Lo paradójico es que habríamos razonado exactamente igual si cambiamos de sobre inicial, por lo que en cualquier caso conviene cambiar de sobre.

Tabla 5.3.12. Objetos matemáticos implícitos en la solución

Tipo	Objetos matemáticos	Significado en la situación
Problema	- Elección del sobre	- Determinar cuál es la mejor elección
Lenguajes	- Verbal	- Explicación de la situación
	- Simbólico	- Expresar sucesos, probabilidades
	- Numérico	- Cantidad de dinero en los sobres; valores de las probabilidades y la esperanza
Conceptos	- Experimento aleatorio	- Elegir sobre - Dinero en cada sobre
	- Sucesos; espacio muestral	- Sobre *A* y *B* - 2000; 1000
	- Experimento compuesto	- Composición de los experimentos anteriores
	- Sucesos en el experimento compuesto	- Producto cartesiano de los espacios muestrales anteriores
	- Probabilidad clásica	- Proporción de casos favorables a posibles
	- Axiomas probabilidad	- Explicitación de los axiomas
	- Probabilidad condicional	- Proporción de ocurrencia suceso respecto a la ocurrencia de otro
	- Regla de la suma	- Probabilidad de ganar
	- Regla del producto	- Probabilidad conjunta; dependencia
	- Esperanza	- Valor esperado de una variable
	- Variable aleatoria	- Cantidad de dinero en cada sobre
Procedimientos	- Cálculo de probabilidad formal	- Aplicar reglas de cálculo formal
	- Cálculo de la esperanza matemática	- Encontrar el valor esperado a partir de los sucesos y su probabilidad
Propiedades	- Diferencia probabilidad condicionada y simple	- Restricción del espacio muestral
	- Esperanza matemática	- Suma de productos de valores de la variable multiplicado por su probabilidad
	- Axioma de unión	- La probabilidad de la suma es suma de probabilidades
Argumento	- Razonamiento deductivo	- Demostración de la solución

Solución correcta

Para resolver correctamente el problema hay que suponer una cantidad concreta en cada sobre. Supongamos que un sobre tiene 1000 euros y el otro tiene 2000. Con el razonamiento anterior, esperaríamos ganar 1500 euros cambiando de sobre. Pero, si abrimos el sobre de 1000 euros, no es cierto que el otro tenga 2000 o 500 euros con probabilidad 1/2, sino que tiene 2000 con probabilidad 1. Si elegimos el sobre de 2000 euros el otro tiene 1000 con probabilidad igual a 1. Luego el valor esperado a ganar si cambiamos de sobre sería la probabilidad de elegir el sobre con 2000 euros, multiplicado por la ganancia esperada si se elige este sobre más la probabilidad de elegir el sobre con los 1000 euros multiplicado por la ganancia esperada; es decir $\frac{(1000-1000)}{2}=0$ y por tanto da igual cambiar o no de sobre. En la Tabla 5.3.12 se presenta el análisis semiótico de esta solución.

Dificultades posibles de los estudiantes

El fallo de este razonamiento se debe a la a*signación incorrecta de probabilidades*. Ello es debido a que, para resolver el problema, partimos de probabilidades calculadas antes de conocer la ganancia esperada, en lugar de las probabilidades a posteriori. Pero, en la situación dada, el resultado del experimento modifica las probabilidades de los sucesos, lo cual es contra intuitivo para los estudiantes, quienes no comprenden que un suceso que ocurre a posteriori pueda modificar una probabilidad de otro que se dio anteriormente. Es decir, encontramos de nuevo un ejemplo de aplicación de la "falacia del eje de tiempos" (Falk, 1986) que se presentó en una proporción importante de los profesores participantes en el Estudio 2.

Paradoja de la división de las apuestas

Este problema, uno de los que originó el cálculo de probabilidades y que interesó a los matemáticos durante un largo periodo histórico, fue propuesto por primera vez por Fray Luca Pacioli en su obra "*Summa de arithmetica, geometria, proportioni et proportionalitá*". Una formulación elemental de este problema es la siguiente:

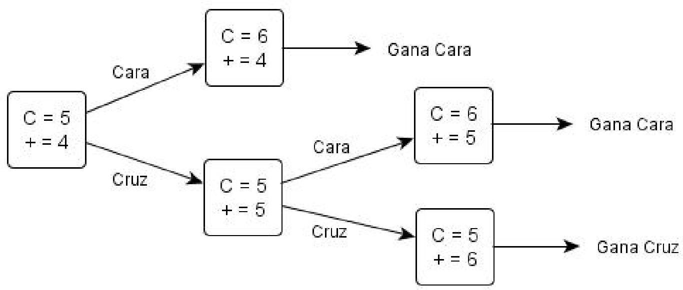
Dos jugadores compiten por un premio que se asigna después de que uno de ellos haya ganado "n" partidas en un juego. En un determinado momento, se tiene que parar el juego, y un jugador lleva mayor número de partidas ganadas que el otro. ¿Cómo debe dividirse la apuesta entre los dos jugadores?

De acuerdo a Székely (1986), Pacioli sugirió una solución de reparto proporcional a la cantidad de puntos obtenido por cada jugador al momento de interrumpir el juego. Dicha solución no fue aceptada, pues sólo tenía en cuenta lo que había ocurrido en la partida hasta el momento que se interrumpe, y no lo que pudiera suceder a continuación. Incluso cuando en un momento dado, un jugador tenga ventaja, el partido todavía podría haber sido ganado por su oponente si no se hubiera interrumpido. Székely indica que otras soluciones incorrectas fueron publicadas por Tartaglia en "*Trattato generale di numeri et misure*" y Cardano en "*Practica arithmeticae generalis*". En dichas soluciones, aunque ya se tiene en cuenta el carácter aleatorio del experimento, no se llega a una correcta enumeración del espacio muestral, ni se calcula correctamente la esperanza matemática del juego.

Solución correcta

La solución correcta fue dada por Pascal y Fermat, quienes en su correspondencia fijan las bases de la teoría de la probabilidad. Consideran por primera vez el conjunto de todas las posibilidades en este juego (que posteriormente llamaríamos espacio muestral) y resolvieron el problema para una partida con dos jugadores, pero no desarrollan una solución completa. La solución general al problema sería obtenida por Huygens en 1657, mediante la introducción de un nuevo concepto que denomina "expectatio" y que formalizado daría lugar a lo que hoy conocemos como esperanza matemática o "cantidad que se espera ganar en un juego equitativo".

Figura 5.3.3. Posibles resultados en el juego, en las condiciones dadas



Explicaremos, a continuación un caso particular de la solución dada por Pascal en su correspondencia con Fermat. En este ejemplo, suponemos que el ganador es el jugador que alcanza primero 6 éxitos (6 caras ó 6 cruces) y suponemos que el juego se interrumpe en el noveno lanzamiento obteniéndose el resultado 5 a 4 y por tanto no hay un ganador. El Caballero de Meré creía que lo más justo era repartir las apuestas en la relación 5 a 4 pero Pascal demostró que las apuestas debían repartirse en la relación 3 a 1, si se toma en consideración que, de continuar el juego, quien lleva 5 éxitos tiene 3 veces más posibilidades de ganar el juego que su contrincante. Supongamos que quien apuesta por “cara” lleva 5 éxitos (ver Figura 5.3.3). Si a partir de las condiciones dadas, calculamos las probabilidades de ganar apostando por “cara” y por “cruz” tenemos:

$$P(Ganar\ C) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$$

$$P(Ganar\ +) = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

En la Tabla 5.3.13 se presenta el análisis semiótico de esta solución.

Tabla 5.3.13. Objetos matemáticos implícitos en la solución

Tipo	Objetos matemáticos	Significado en la situación
Problema	- Repartición de la ganancia	- Determinar la mejor repartición
Lenguajes	- Verbal	- Explicación de la situación
	- Gráfico	- Diagrama en árbol
	- Simbólico	- Expresar sucesos, probabilidades
	- Numérico	- Resultados en el juego, valores de las probabilidades
Conceptos	- Experimento aleatorio	- Tirar las monedas
	- Sucesos; espacio muestral	- Cara/cruz - Ganar/no ganar
	- Experimento compuesto	- Composición de los experimentos anteriores
	- Unión de sucesos	- Sucesos en uno u otro conjunto
	- Probabilidad clásica	- Proporción de casos favorables a posibles.
	- Axiomas probabilidad	- Explicitación de los axiomas
	- Probabilidad condicional	- Proporción de ocurrencia suceso respecto a la ocurrencia de otro
	- Variable aleatoria	- Número de caras y cruces en cada momento del juego - Número de partidas necesarias para acabar el juego
Procedimientos	- Cálculo de probabilidad formal	- Aplicar reglas de cálculo formal
	- Cálculo de probabilidad intuitivo	- Aplicar reglas intuitivas de cálculo
	- Representación gráfica	- Diagrama; esquema
Propiedades	- Diferencia probabilidad condicionada y simple	- Restricción del espacio muestral
	- Regla de la suma	- Probabilidad de ganar
	- Regla del producto	- Probabilidad conjunta; dependencia
	- Teorema de la probabilidad total	- Aplicar a la situación
	- Axioma de unión	- La probabilidad de la suma es suma de probabilidades
Argumentos	- Razonamiento deductivo	- Demostración de la solución

Dificultades posibles de los estudiantes

Como en los casos anteriores, son varios los razonamientos erróneos (algunos sostenidos por matemáticos famosos) los que puede llevar a una incorrecta solución a este problema. No es extraño, en consecuencia que puedan aparecer en la clase de probabilidad o incluso sean compartidas por los profesores:

- *Razonamiento determinista sobre el problema.* Una de las dificultades a las que se puede llegar en el razonamiento de este problema es la misma que planteó el caballero de Meré a Pascal y que condujo al nacimiento de la teoría de la probabilidad tal como la conocemos. El Caballero de Meré creía que lo más justo era repartir las apuestas en la relación 5 a 4, sin tener en cuenta la ventaja del jugador que iba ganando o las posibilidades del otro de ganar. Este razonamiento, erróneo, resuelve el problema en forma determinista, dándole un sentido aritmético. El razonamiento es común en este tipo de problemas ya que la lógica dice que se debe de repartir el dinero equitativamente atendiendo al puntaje acumulado por cada bando al interrumpirse el juego. Este razonamiento coincide con el seguido por Paciolli o Tartaglia para su resolución.
- Los estudiantes podrían mostrar una percepción incorrecta de la independencia, que apareció en algunos participantes en el Estudio 2, pensando que, en este juego, la probabilidad de ganar sigue siendo 1/2, independientemente de los resultados a favor de cada jugador. Esta solución es razonable, puesto que sabemos que, independientemente de que se haya producido una racha de caras o cruces la probabilidad de cara o cruz no varía. Sin embargo, puesto que no estamos buscando la probabilidad de una cara o cruz, sino de cinco acumuladas, será más sencillo para aquél jugador que lleve ventaja.

Paradoja de Condorcet

El análisis de los métodos de votación pretende conseguir una forma de elección lo más cercana posible al sentir de la mayoría del cuerpo electoral. El problema de elegir un buen método de asignación de escaños, se puso de manifiesto por Antoine de Caritat Condorcet, sobre 1780, en uno de sus principales trabajos: el "*Ensayo sobre la aplicación del análisis a la probabilidad de las decisiones sometidas a la pluralidad de voces*". En dicho trabajo, que fue elaborado por su autor para intentar hallar el número óptimo de miembros de un jurado en los Tribunales Populares de Justicia, instaurados por la Revolución Francesa. Condorcet demostró que los sistemas de votación que aplican la regla de aceptar simplemente la opinión de la mayoría, podrían no llegar a una situación concluyente.

La paradoja advierte que, aunque las preferencias individuales en las votaciones tienen propiedad transitiva (si una persona prefiere al candidato *A* frente al *B* y prefiere al *B* frente al *C*), no tiene por qué haber transitividad en las preferencias colectivas (si la mayoría de un grupo de personas prefiere al candidato *A* frente al *B* y también la mayoría del mismo grupo prefiere a *B* frente a *C*, esto no implica que la mayoría del grupo prefiera a *A* frente a *C*). Esta es una de las situaciones, aunque no la única, en que la propiedad transitiva no se aplica en el cálculo de probabilidades (Székely. 1986).

Solución correcta

Para dar una solución intuitiva a esta paradoja usaremos el ejemplo presentado en el servidor "Matemáticas Educativas", en el epígrafe "matemáticas del voto" (www.iescarrus.com/edumat/taller/votaciones/votaciones_02.htm). Supongamos que tenemos 49 votantes que tienen que elegir entre tres candidatos *X*, *Y* y *Z*. Como es normal, cada votante tiene sus propias preferencias, por tanto es posible que un votante prefiera al candidato *X* antes que al *Y* y al *Y* antes que al candidato *Z*, por lo tanto también preferirá a *X* antes que al candidato *Z*. Es decir existe una transitividad entre las preferencias y ello nos permite asociar a cada votante con un orden de preferencias (*X*>*Y*>*Z*). En la Tabla 5.3.14 hemos presentado todos los posibles órdenes de preferencia que se pueden dar entre tres candidatos y unas frecuencias posibles, dentro de la muestra de 49 personas.

Tabla 5.3.14. Posibles frecuencias de las preferencias entre tres candidatos

X>Z>Y	Z>X>Y	Z>Y>X	Y>Z>X	Y>X>Z	X>Y>Z
21	3	4	16	5	0

Si se tienen en cuenta estas preferencias, en una primera vuelta de las elecciones, se elegiría el candidato preferido por más personas, y por tanto los resultados serían los siguientes: El candidato *X* tiene 21 (21+0) votos, el *Y* 21 votos (16+5), y el candidato *Z* tiene 7 votos (3+4) y se produce un empate.

Supongamos que para romper el empate entre *X* e *Y* se realiza una nueva votación, en la que sólo se puede votar a los candidatos que empataron en la primera vuelta, lo que es habitual en muchos sistemas de votación. En este supuesto, los votos del candidato *Z* pasan al resto de candidatos según la segunda preferencia. Los resultados entonces son: *X*, 24 votos (21+3) e *Y*, 25 votos (4+16+5) y por tanto ganaría el candidato *Y c*on 25 votos frente a los 24 de *X*.

Aparentemente esto resuelve el problema y habría que dar el puesto a *Y*. Sin embargo, si analizamos el número de personas que prefieren a *Z* antes que a *Y*, se observaría que hay 21 personas que prefieren a *Y* y 28 que prefieren a *Z*, por lo que debiera elegirse a *Z*. Por otro lado 26 personas prefieren a candidato X antes que a *Z* mientras que sólo 23 prefieren a *Z* antes que al *X*.

El resultado es paradójico porque implica que la voluntad de mayorías parciales entra en conflicto entre sí. En otras palabras es posible que un procedimiento de elección falle el criterio de que siempre hay un ganador. El hecho de que una mayoría prefiera a *X* antes que a *Z* y a *Z* antes que a *Y*, no conduce necesariamente a que prefieran a *X* antes que a *Y*. Se produce un proceso cíclico. En la Tabla 5.3.15 se presenta el análisis semiótico de esta solución.

Tabla 5.3.15. Objetos matemáticos implícitos en la solución

Tipo	Objetos matemáticos	Significado en la situación
Problema	- Elegir al candidato	- Determinar el candidato ganador
Lenguajes	- Verbal	- Explicación de la situación
	- Simbólico	- Expresar sucesos, probabilidades
	- Numérico	- Número de personas que prefieren a cada candidato
Conceptos	- Experimento aleatorio	- Candidato preferido por cada persona - Orden de preferencia de cada persona entre tres candidatos
	- Sucesos; espacio muestral	- Candidato ganador. Posibles permutaciones del orden de preferencia
	- Orden	- Orden de preferencia de candidatos
	- Comparación de probabilidades	- Comparación de preferencias de los tres candidatos
	- Unión de sucesos	- Preferir a un candidato antes que a otro
Procedimientos	- Cálculo de probabilidades	- Aplicar reglas de cálculo intuitivo
Propiedades	- Regla de Laplace	- Casos favorables/casos posibles
	- No transitividad de la probabilidad	- La probabilidad no tiene propiedad transitiva
Argumentos	- Razonamiento deductivo	- Demostración de la solución
	- Razonamiento empírico	- Comparar preferencias individuales

Dificultades posibles de los estudiantes

El principal razonamiento erróneo que puede darse en este problema es pensar que la propiedad transitiva se aplica a la probabilidad. Este error se presenta por un fallo en el proceso de generalización, ya que se aplica una propiedad (transitiva) a un caso inapropiado. El error es lógico porque los alumnos encuentran la propiedad transitiva en otras situaciones, por ejemplo, la ordenación numérica y por ello se puede crear un obstáculo cuando trabajan problemas de probabilidad, Como en otros obstáculos, el error aparece no por falta de conocimiento, sino por un conocimiento anterior que no es válido en la situación propuesta. Aunque no se ha investigado específicamente este error en los Estudios 1 y 2, si se observaron en los mismos casos de participantes que realizaban un proceso incorrecto de generalización.

5.3.5. PROCESOS MATEMÁTICOS

Al igual que en el análisis de los recursos en Internet que llevamos a cabo en la sección 5.2, además del uso diferentes tipos de objetos matemáticos, podemos observar que en el trabajo con paradojas se usan implícitamente varios de los procesos matemáticos descritos por Godino, Font y Wilhelmi (2008). Seguidamente enumeramos los diferentes procesos que podemos observar en el uso de las paradojas.

Procesos de materialización-idealización:

- Los objetos a los que hacen referencia las paradojas (cajas, sobres, monedas para el reparto, prisioneros…) e incluso las acciones que realizamos con ellos (abrir una caja, hacer una tirada,…) son objetos imaginarios que son representados por el texto del problema, es decir por objetos visibles u ostensivos.
- Las operaciones aritméticas se representan simbólicamente; por ejemplo $X/2$ representa la operación de dividir una cantidad en dos partes y también el resultado de la división, que es un número real que no percibimos directamente.
- Los diferentes símbolos representan propiedades, conceptos u operaciones. Por ejemplo, $Z>X>Y$ representa en forma visible un orden.
- El diagrama en árbol de las soluciones intuitivas representa, por un lado la estructura del experimento (por pasos) y en cada paso, los resultados del experimento y sus probabilidades.

Procesos de particularización – generalización (dualidad extensivo – intensivo):

- Sabemos que la suma total de todas las probabilidades de los sucesos en un experimento es la unidad. En cada ejemplo, particularizando llegamos a las probabilidades de los sucesos dados. Por ejemplo, en el problema del chico- chica deducimos que la probabilidad inicial de tener una chica es 1/2.
- En la paradoja de Bertrand vemos que si suponemos que la moneda es de plata estamos obteniendo el doble de aciertos que si elegimos la de oro. Aunque no

sepamos por qué, aceptamos (generalizando) que la probabilidad de acierto es mayor.

- Conocemos el axioma de la unión que se cumple en forma general. Para cada problema concreto, lo aplicamos para calcular la probabilidad total o la probabilidad de la unión de dos o más sucesos (aunque se cumple para cualquier número de sucesos, acá particularizamos para cada caso).

Procesos de representación – significación:

- Por ejemplo, el objeto "probabilidad" lo representamos por la letra *P*; La probabilidad de un suceso que denominamos *A* lo representamos mediante *P(A). A* (el suceso) a su vez representa un resultado.

Procesos de descomposición – reificación (dualidad sistémico – unitario):

- Cada suceso de un experimento aleatorio es elemental. Pero el espacio muestral del experimento es sistémico.
- Cada resultado de un experimento es elemental, pero la frecuencia relativa o el porcentaje de veces que ocurre un resultado se calcula como una operación sobre el total de los resultados al realizar n veces el experimento.
- Cada rama del diagrama en árbol es elemental, mientras que todo el diagrama en árbol es sistémico.

Procesos de personalización - institucionalización (dualidad personal – institucional):

- Como en el caso de los recursos en Internet, al comenzar un proceso de estudio basado en el trabajo con paradojas será necesario lograr que los estudiantes asuman el problema y se involucren en su solución. Pensamos que esto se consigue totalmente con estas paradojas pues aumenta el interés del alumno y olvida que se trata de una clase de matemáticas. Ha personalizado la situación.
- El proceso inverso es pasar de lo personal a lo institucional. Esto lo logra el maestro cuando discute colectivamente con los estudiantes para decidir cuáles de las soluciones son correctas y por qué.

5.3.6. IDONEIDAD DIDÁCTICA DEL TRABAJO CON LAS PARADOJAS

En lo que sigue, analizamos la idoneidad del trabajo con paradojas, siguiente lo expuesto por Godino, Wilhelmi y Bencomo (2005); Godino, Contreras y Font (2006). Dichos autores definen la idoneidad didáctica como la articulación de seis componentes cada uno de los cuales puede presentarse en mayor o menor grado en el trabajo con paradojas.

- *Idoneidad epistémica o matemática:* Representatividad de los significados institucionales implementados (o pretendidos), respecto de un significado de referencia. Pensamos que las paradojas, podrían tener una idoneidad matemática en el estudio de los conceptos de: Probabilidad condicional, experimento compuesto, dependencia e independencia y experimentos dependientes e independientes. Para conseguir esta idoneidad en los cursos dirigidos a profesores sería necesario que los participantes primero realicen algunas simulaciones de la situación y, a continuación, el formador pida que encuentren la solución correcta al problema. Debido al carácter contra-intuitivo de las paradojas, surgirá más de una solución (algunas incorrectas). El formador debe listar todas las soluciones sugeridas por los profesores en la pizarra y posteriormente sería organizado un debate para decidir cuál es la mejor solución. El objetivo de este debate, donde es posible tanto el razonamiento correcto como los conceptos erróneos, es aumentar el conocimiento probabilístico del maestro. Si no hay acuerdo final, los participantes serán animados a realizar una demostración matemática de su solución y se organizará un debate que permitirá revelar los razonamientos erróneos.
- *Idoneidad cognitiva:* Grado en que los significados pretendidos/implementados son asequibles a los alumnos, así como si los significados personales logrados por los alumnos son los significados pretendidos por el profesor. Pensamos que la situación planteada tiene suficiente idoneidad en cursos de formación de profesores, sobre todo de profesores de secundaria que son licenciados en matemáticas. Asimismo podría tener idoneidad suficiente en los últimos cursos de secundaria o Bachillerato, pues los razonamientos descritos están al alcance de los alumnos.
- *Idoneidad interaccional*: Grado en que la organización de la enseñanza permite identificar conflictos semióticos y resolverlos durante el proceso de instrucción. Este tipo de idoneidad dependerá de cómo organiza el educador el trabajo en el aula. Será

importante que los estudiantes trabajen en grupos para que surja el conflicto y se explicite. También se les pedirá organizar una solución colectiva de las soluciones para que los mismos alumnos ayuden a sus compañeros a detectar los puntos equivocados.

- *Idoneidad mediacional:* Disponibilidad y adecuación de los recursos necesarios para el desarrollo del proceso de enseñanza-aprendizaje. No se precisa de muchos recursos, pues incluso podría hacerse una simulación con objetos físicos. Pero con ayuda de un sobre o una caja o unas tarjetas los alumnos pueden trabajar estas paradojas.
- *Idoneidad emocional:* Interés y motivación del alumnado en el proceso de estudio. Pensamos que esta es la más alta de todas con estas paradojas, ya que sin duda intrigan e interesan a todo el que trata de resolverlas.

Para acabar esta parte, recordamos que González (2004) señala que el uso de la historia con fines didácticos depende del conocimiento histórico del profesor y su iniciativa para adaptar este saber a los intereses y necesidades del grupo, por lo que ha de exponer los avances de la disciplina junto con su estado actual teórico y de aplicabilidad. El estudio de la historia de la probabilidad y de las paradojas asociadas a la misma será entonces un componente importante en la preparación de formadores.

5.3.7. CONCLUSIONES DEL ESTUDIO 4

En el Estudio 4 hemos realizado un análisis de algunas paradojas ligadas a la probabilidad condicional e independencia para valorar su utilidad potencial en la formación de profesores. Los objetivos planteados fueron los siguientes:

1. *Seleccionar algunas paradojas cuya solución pueda abordarse con conocimientos elementales de probabilidad condicional e independencia*. Habría de tener potencial para organizar un debate en torno a las cuestiones matemáticas y didácticas, para contribuir a aumentar el conocimiento del profesor.
2. *Realizar un análisis de los objetos matemáticos puestos en juego en las posibles soluciones* de las paradojas seleccionadas, similar al realizado en el Estudio 3 con los recursos en Internet e *identificar algunos posibles conflictos semióticos* que pudiesen surgir durante su solución.

3. *Valorar los procesos matemáticos utilizados en el trabajo con las paradojas* seleccionadas y la *idoneidad didáctica de este recurso* en los cursos de formación de profesores.

La hipótesis que plantada fue que el *análisis de las paradojas nos permitiría mostrar una variedad de objetos matemáticos involucrados en sus posibles soluciones, así como dificultades y conflictos potenciales relacionados con las dificultades descritas en las investigaciones previas y las encontrados en los Estudios 1 y 2.* Asimismo esperábamos *mostrar la idoneidad didáctica de este recurso para la formación de profesores.*

Mediante el estudio realizado se han cumplido razonablemente los objetivos y se ha confirmado la hipótesis, pues hemos seleccionado algunas paradojas cuya solución pueda abordarse con conocimientos elementales de probabilidad condicional e independencia. Asimismo se ha realizado el análisis de los objetos matemáticos puestos en juego en las posibles soluciones, identificado algunos posibles conflictos semióticos que puedan surgir durante su solución y que fueron observados en los Estudios 1 y 2. Valoramos, asimismo, la alta idoneidad didáctica de este recurso en los cursos de formación de profesores.

Los resultados muestran la utilidad potencial e idoneidad didáctica de dichas paradojas en la formación de profesores. Estos necesitan apoyo y formación adecuada para tener éxito en el logro de un equilibrio adecuado de la intuición y el rigor en la enseñanza de la probabilidad. Lamentablemente, no todos los profesores reciben una buena preparación para enseñar probabilidad en su formación inicial, como señalan Franklin y Mewborn (2006).

Autores como León (2009) indican que la historia de la probabilidad presenta situaciones muy atractivas que pueden conducir a reflexionar sobre la presencia del azar en la cotidianidad además de servir de motivación hacia el estudio por parte de los alumnos. Como muestran Basulto y Camuñez (2007), cuando un estudiante, o en nuestro caso un futuro profesor, aborda la resolución de problemas históricos en el cálculo de probabilidades, el contexto en el que se plantea el problema en la historia y las tácticas de resolución de los primeros autores, puede dar una motivación añadida al alumno y llevarlos a un aprendizaje a partir de los errores y aciertos que estos cometían.

En el trabajo con las paradojas analizadas se pueden identificar algunas ideas fundamentales descritas por Heitele (1975): Sucesos y espacio muestral, probabilidad y convergencia, operaciones combinatorias, las reglas de adición y multiplicación de probabilidades, independencia, probabilidad condicionada, variable aleatoria, equidistribución y simetría, esperanza matemática y muestreo. Por tanto, actividades como la que se analizan en este apartado y análisis como los proporcionados en este estudio proporcionan una fuente de conocimiento matemático del contenido al profesor.

Además, el trabajo con las paradojas en los cursos para profesores puede servir para aumentar los conocimientos didácticos del contenido probabilidad de los docentes en algunas componentes de los conocimientos matemático-didácticos descritos por Godino (2009):

- Componente epistémica: Al hacer reflexionar al futuro profesor sobre los diversos significados históricos de la probabilidad (subjetiva, frecuentista y clásica) y las controversias ligadas a su definición en diferentes momentos históricos; el análisis de los objetos y procesos matemáticos de las soluciones, aumenta su conocimiento especializado del contenido.
- Componente cognitivo: Al experimentar el profesor dificultades similares a las que puede encontrar sus estudiantes aumenta su conocimiento del contenido y los estudiantes. El estudio de las dificultades de los estudiantes con los conceptos de probabilidad condicional e independencia y sus posibles razonamientos y estrategias correctas e incorrectas para resolver el problema también contribuye a este conocimiento.
- Componente afectivo: Experimentando nuevos métodos de enseñanza, basados en el juego, experimentación y debate, que permite aumentar el interés de los alumnos y su participación en la actividad.
- Componente interaccional: Aumentando su experiencia sobre la forma de organizar el discurso y el tiempo didáctico y de hacer aflorar y resolver los conflictos cognitivos de los estudiantes.
- Componente ecológico: Dado que la actividad se puede conectar con el estudio de las concepciones erróneas sobre el azar (psicología) y con los problemas sociales relacionados con la adicción a los juegos de azar (sociología).

CAPÍTULO 6

EVALUACIÓN DE UNA EXPERIENCIA DE FORMACIÓN DE PROFESORES

6.1. INTRODUCCIÓN

En los capítulos 3 y 4 se presentaron los resultados de los Estudios 1 y 2, en los que analizamos una parte de los conocimientos matemáticos de profesores en formación de educación primaria y secundaria con relación a la probabilidad condicional, mostrando algunas de sus carencias, tanto en las estrategias de resolución de problemas, como en la comprensión y aplicación de las ideas de probabilidad condicional e independencia. Asimismo mostramos la existencia de algunos sesgos en el razonamiento probabilístico condicional de estos futuros profesores.

Por otro lado, en el capítulo 5 (Estudios 3 y 4) hemos llevado a cabo un estudio de posibles recursos didácticos que permiten mejorar el conocimiento de los profesores con relación a la probabilidad condicional. Se mostró que dichos recursos tienen un gran potencial para poder plantear talleres y actividades dirigidas a los profesores que permita profundizar simultáneamente en su conocimiento matemático y su conocimiento didáctico del contenido sobre probabilidad condicional, como hemos mostrado en Batanero, Contreras, Díaz y Arteaga (2009) y Contreras, Díaz, Batanero y Ortiz (2010).

Basándonos en los resultados de los cuatro estudios anteriores, en este capítulo se evalúa una experiencia formativa dirigida a profesores de educación secundaria, tanto en su etapa de formación, como en ejercicio. Analizamos, en primer lugar los conocimientos matemáticos de los participantes y seguidamente, describimos su evolución como consecuencia de la actividad formativa, que está basada en una paradoja clásica de probabilidad analizada en el Estudio 4.

El taller plantea, en primer lugar, una actividad matemática que permite evaluar los conocimientos previos de los participantes sobre este objeto matemático, así como los posibles sesgos de razonamiento descritos en la literatura, algunos de los cuáles hemos descrito en el capítulo 2 y fueron evaluados en los Estudios 1 y 2 (capítulos 3 y 4). Dicha actividad fue descrita y analizada en Batanero, Godino y Roa (2004), quienes realizaron una primera evaluación de la misma con un grupo reducido de estudiantes de la licenciatura de estadística (n=47). Se desea comparar si las dificultades de partida y el aprendizaje a lo largo de la actividad, descrita por Batanero, Godino y Roa, se reproducen en una muestra de mayor tamaño y se conservan en el caso de los docentes en servicio. Además, se han tomado datos no sólo de profesores en formación, sino de profesores en ejercicio de España, Portugal y México, ampliando la muestra usada en Batanero, Godino y Roa. Completamos el análisis, respecto al realizado por los citados autores, identificando las estrategias iniciales y finales, realizando un análisis semiótico, identificando conflictos semióticos y relacionando todo ello con el tipo de profesor (en ejercicio o formación) y el contexto geográfico.

En la segunda parte del taller, se plantea una situación de análisis didáctico, que fue completada por una parte de la muestra, permitiendo evaluar el conocimiento del contenido didáctico de estos participantes en relación a nuestro objeto de estudio. Todos estos resultados han sido publicados en Contreras, Batanero, Fernándes y Ojeda (2010) y Batanero, Contreras, Fernándes y Ojeda (2010). Este taller se ha realizado en algunos cursos de didáctica de la probabilidad dirigidos a profesores en formación y en ejercicio, en unos casos como actividades organizadas en congresos de estadística, en otros dentro de cursos de actualización profesional o en una materia optativa de la licenciatura de estadística o matemáticas. A continuación se describen los objetivos e hipótesis del

Estudio 5, se analiza el material, se presentan las categorías de análisis y se discuten los resultados obtenidos.

6.2. OBJETIVOS E HIPÓTESIS DEL ESTUDIO 5

Como sugiere Steinbring (1990), debido a su riqueza, tanto el conocimiento estocástico, como el conocimiento didáctico del profesor necesitan ser reconstruidos por los mismos profesores y no puede transmitirse simplemente "desde fuera" o ser prescrito. Esto no significa que no podamos ayudar y proporcionar soporte en este proceso de construcción. Por su parte, Ponte y Chapman (2006) indican que debemos considerar a los profesores como profesionales, haciendo que todos los elementos de la práctica docente (preparación de las clases, tareas y materiales, la realización de clases, observación y reflexión sobre la experiencia) sean el elemento central del proceso de formación del profesorado. Siguiendo estas sugerencias, el Estudio 5 tiene los siguientes objetivos:

Objetivo 1: Hacer experimentar a los profesores una situación didáctica, basada en una paradoja clásica de la teoría de la probabilidad, para hacer aflorar algunos conocimientos matemáticos y didácticos con relación a la probabilidad condicional.

Se trata de hacer experimentar a los profesores la dificultad que encuentra un alumno al resolver las tareas o problemas que se les plantea. Puesto que esperamos que, tanto los profesores en ejercicio, como los futuros profesores (que están en sus últimos años de formación) tengan unos conocimientos suficientes sobre probabilidad condicional, buscamos algunos ejemplos de problemas, que, siendo aparentemente sencillos, puedan sin embargo, tener soluciones contraintuitivas o sorprendentes. Se pretende también que los profesores experimenten las diversas fases de la teoría de situaciones didácticas de Brousseau (1986; 1997); acción, formulación, validación e institucionalización. Puesto que la fase de acción requiere que los participantes se involucren con un "verdadero problema" (en el sentido de ser un problema no inmediatamente resoluble para ellos), se decidió elegir un problema paradójico.

Basamos el taller en la paradoja de la caja de Bertrand, descrita en el libro de Székely (1986) y que fue analizada con detalle en el Capítulo 5. Como se sugiere en dicho capítulo, una paradoja es una declaración en apariencia verdadera que conlleva a una auto-contradicción lógica o a una situación que contradice el sentido común. La identificación de paradojas basadas en problemas en apariencia razonables y simples, ha impulsado importantes avances en matemáticas y en el campo de la probabilidad. Estos problemas, así como las soluciones, tanto correctas como erróneas que algunos de participantes puedan defender durante el taller, servirán para analizar cuáles son los conceptos involucrados en la solución del problema. Deseamos también que los participantes experimenten un proceso de aprendizaje de ideas estocásticas fundamentales (Heitele, 1975), a partir de una situación didáctica que puede ser adecuada en la Educación Secundaria Obligatoria o Bachillerato y contextualizar la reflexión epistemológica sobre estas ideas, así como identificar cuáles de ellas intervienen en la situación didáctica.

Objetivo 2. Enfrentar a los participantes con sus intuiciones incorrectas y concienciarlos de la utilidad de la teoría de la probabilidad para resolver posibles conflictos en las mismas.

Algunos profesores pudieran no ser conscientes de algunas de las dificultades y sesgos probabilísticos de sus alumnos (Stohl, 2005) y pudieran ellos mismos presentar algunos de los descritos en los Estudios 1 y 2, donde se mostró la extensión de dichos sesgos entre futuros profesores. Es importante encontrar algún modo de apoyarlos, para que tomen consciencia de estos sesgos y los superen, además de pensar en actividades que les sirvan para conectar los aspectos conceptuales y didácticos (Ball, 2000).

En el taller dejaremos a los participantes que encuentren sus soluciones (correctas e incorrectas) y organizaremos un debate, para que ellos mismos lleguen a la solución correcta y perciban los puntos en que cometieron un error. Todo ello les llevará a una mayor sensibilidad hacia las dificultades de sus alumnos y el papel que tiene el error en el proceso de aprendizaje (Batanero, Godino y Roa, 2004).

Objetivo 3. Introducir algunos de los tipos de análisis didácticos propuestos por Godino, Font y Wilhelmi (2008), ejemplificándolos en la situación y evaluar los conocimientos didácticos de los participantes.

En la segunda parte del taller, se propondrá a los participantes el análisis de la tarea propuesta y del modo en que se organizó el trabajo en el taller. En concreto, se propondrán tres diferentes tipos de análisis propuestos por Godino, Font y Wilhelmi (2008): En un primer nivel, se analizan las prácticas matemáticas y los objetos involucrados en la resolución de la tarea propuesta. Seguidamente se les pide identificar los posibles errores y dificultades de los estudiantes en la tarea, que serán similares a los que ellos han experimentado a lo largo del taller. Finalmente se pide identificar las fases diferenciadas en el proceso de estudio, para introducirlos a la teoría de situaciones de Brousseau (1986; 1997). Con todo ello podremos proporcionar alguna información sobre los conocimientos didácticos de los participantes.

Hipótesis

La principal hipótesis que nos planteamos en el Estudio 5 es que *la actividad realizada en el taller, permitirá hacer aflorar algunos de los sesgos en el razonamiento condicional* que se describieron en el capítulo 2 y que se mostraron en los futuros profesores en los Estudios 1 y 2. Asimismo esperamos *una superación de algunos de estos sesgos en una parte de los profesores participantes.*

Esta hipótesis viene avalada por los resultados descritos en la experimentación de este taller con una muestra de estudiantes de estadística por Batanero, Godino y Roa (2004). En el citado estudio, a pesar de la alta formación estadística de los participantes, una gran proporción mostró dificultad inicial en percibir el suceso condicionante en el experimento y comprender su estructura, proporcionando, en consecuencia, respuestas incorrectas. Otros participantes mostraron una baja percepción de la dependencia de los experimentos aleatorios involucrados. Los autores también informan de la evolución positiva de estos sesgos como consecuencia de la actividad.

Los participantes en nuestro estudio tienen una formación estadística similar o a veces menor que los del estudio de Batanero, Godino y Roa. Por otro lado, los sesgos en el razonamiento condicional se han descrito no sólo en estudiantes, sino incluso en profesores de métodos de investigación, como se analizó en el capítulo 2. Es por ello que esperamos una replicación de los resultados de Batanero, Godino y Roa en nuestro trabajo. Los autores no analizaron los conocimientos didácticos de los participantes en su investigación. Puesto que la mayoría de los sujetos de nuestra muestra no ha tenido formación didáctica específica en relación a la enseñanza de la estadística, nuestra expectativa es que los conocimientos didácticos de los participantes en el Estudio 5 podrían tener alguna limitación. No obstante, no formulamos formalmente una hipótesis específica en este punto, al carecer de antecedentes de investigación.

6.3. MATERIAL Y MÉTODO

El material analizado son los protocolos escritos por los participantes, recogidos en los talleres sobre ideas estocásticas fundamentales, impartidos en cursos de formación de profesores o como actividades formativas en congresos (Ver anexo 3).

Figura 6.3.1. Material para el taller de ideas estocásticas fundamentales

PARTE 1. RESOLUCIÓN DE PROBLEMAS

Se toman 3 fichas de la misma forma y tamaño, de las cuales una es roja por ambas caras; otra es azul por una cara y roja por la otra, y la tercera es azul por las dos caras.
El profesor coloca las tres fichas en una caja, que agita convenientemente, antes de seleccionar una de las tres fichas, al azar. Muestra, a continuación, una de las caras de la ficha elegida, manteniendo la otra tapada, pidiendo a sus alumnos que adivinen el color de la cara oculta.
Una vez hechas las apuestas, el profesor muestra la cara oculta. Cada alumno que haya acertado en la predicción efectuada, consigue un punto.

Figura 1. Hoja de registro

Ensayo nº.	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Color de la cara mostrada										
Color predicho										
Color de la cara oculta										

1. *¿Has seguido alguna estrategia? Descríbela*
2. *¿Por qué sigues esta estrategia o por qué no sigues ninguna? Da un razonamiento*

Figura 2. Hoja de registro. Continuación del juego

Ensayo nº.	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Color de la cara mostrada										
Color predicho										
Color de la cara oculta										

3. ¿Cuál es ahora tu estrategia? ¿Por qué?
4. Si estás seguro/a de tu estrategia, da una demostración matemática de la misma
5. Para los que no están seguros, haremos una tercera ronda del juego

Figura 3. Hoja de registro

Ensayo nº.	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Color de la cara mostrada										
Color predicho										
Color de la cara oculta										

6. Si estás seguro/a de tu estrategia, da una demostración matemática de la misma

El taller puede adaptarse a un tiempo variable (entre 2 y 6 horas, dependiendo del tiempo disponible) y consta de una parte de resolución de problemas, que enfatiza el conocimiento del contenido matemático y otra parte de análisis didáctico que enfatiza el conocimiento didáctico del contenido.

A lo largo del taller, los participantes realizan varias actividades, que incluyen la presentación del taller por el formador de profesores, la realización de un juego y la resolución de un problema asociado, así como el análisis didáctico de todo el proceso de estudio. Los participantes también completan una ficha que constituye el material analizado en este estudio. En las Figuras 6.3.1 y 6.3.2 mostramos la ficha de trabajo utilizada en el taller. Seguidamente describimos el desarrollo del taller, realizamos un análisis a priori de las tareas planteadas y presentamos los datos recogidos.

Figura 6.3.2. Material para el taller de ideas estocásticas fundamentales

PARTE 2. ANÁLISIS DIDÁCTICO

7. Primer nivel de análisis: Completa la siguiente tabla de objetos que has utilizado en esta situación

Tabla 1

Tipos	Objetos matemáticos	Significado en la situación
Situaciones problemas		
Lenguaje		
Conceptos		
Procedimientos		
Propiedades		
Argumentos		

8. Segundo nivel de análisis: Lista algunos posibles conflictos y dificultades por parte de los estudiantes

9. Tercer nivel de análisis: configuraciones didácticas. ¿Qué diferentes fases didácticas puedes identificar en toda la actividad, teniendo en cuenta la forma de trabajo e interacción entre alumnos y del alumno con el profesor?

6.4. DESCRIPCIÓN DEL TALLER Y ANÁLISIS A PRIORI

El taller comienza con una introducción, realizada por el formador del profesor, quien describe los objetivos del taller, así como los supuestos que subyacen en la lista de ideas estocásticas fundamentales descritas por Heitele (1975). Este autor indica que, además de la idea básica de aleatoriedad, hay una serie de conceptos sobre los cuáles se apoya todo el cálculo de probabilidades, que son los que debemos enseñar en los niveles no universitarios. Heitele sostiene las siguientes ideas:

- Cualquier tema puede ser enseñado efectivamente en diferentes niveles, siempre que se use un lenguaje y formalización adecuados.
- La transición a un nivel cognitivo superior se facilita si el tema subyacente ha sido preparado en una representación conveniente en etapas cognitivas anteriores.

De acuerdo a Batanero, Godino y Roa (2004), las ideas fundamentales proporcionan modelos adecuados para las diferentes etapas del desarrollo cognitivo, que varían sólo en la formalización del lenguaje y en su nivel de profundización, pero no en lo esencial. En la actividad propuesta se pueden identificar algunas ideas fundamentales descritas por Heitele (1975): Sucesos y espacio muestral, probabilidad y convergencia, operaciones combinatorias, las reglas de adición y multiplicación de probabilidades, independencia, probabilidad condicionada, variable aleatoria, equidistribución y simetría, esperanza matemática y muestreo.

6.4.1. PRIMERA PARTE. RESOLUCIÓN DE PROBLEMAS

Planteamiento del juego

Una vez introducido el taller, se propone a los profesores experimentar una situación didáctica que ha sido diseñada para la enseñanza de la probabilidad en el nivel de Secundaria o Bachillerato, y también puede usarse en la universidad. Los profesores harán el papel de alumnos y el formador de profesores el papel del profesor. La actividad se inicia planteando el siguiente juego:

Se toman 3 fichas de la misma forma y tamaño, de las cuales una es roja por ambas caras; otra es azul por una cara y roja por la otra, y la tercera es azul por las dos caras. El profesor coloca las tres fichas en una caja, que agita convenientemente, antes de seleccionar una de las tres fichas, al azar. Muestra, a continuación, una de las caras de la ficha elegida, manteniendo la otra tapada, pidiendo a sus alumnos que adivinen el color de la cara oculta. Una vez hechas las apuestas, el profesor muestra la cara oculta. Cada alumno que haya acertado en la predicción efectuada, consigue un punto.

Dicho juego es una variante de la paradoja de la caja de Bertrand propuesta por Joseph Bertrand (1822-1900), que fue analizada con detalle en la sección 5.3.4 (Estudio 4). Una vez comprendido en qué consiste el juego, y tras haber hecho algunos ensayos, se pide a los estudiantes (en este caso los profesores participantes en el taller) que busquen una estrategia que les permita obtener el mayor número de puntos (aciertos) en una serie larga de repeticiones del juego. Se completan los pasos 1 a 6 de la Figura 6.3.1, alternando entre dejar un tiempo de trabajo al participante y realizar varias simulaciones, cada una de ellas de 10 jugadas del juego. Cada profesor pensará individualmente su estrategia, y finalmente deben argumentar la estrategia elegida.

Finalmente se organiza el debate en clase para decidir cuál es la mejor estrategia (que consiste en apostar al mismo color de la cara que puede verse). A algunos profesores, a pesar de su formación, les cuesta ver que esta estrategia es mejor que, por ejemplo apostar al azar. El análisis de los diferentes argumentos lleva a la idea de probabilidad condicional y sucesos dependientes, así como a concienciarse de las dificultades de los alumnos con estos conceptos.

Solución matemática del juego

Como se ha dicho, la solución que esperamos encuentren los profesores es que la mejor estrategia es dar el mismo color de la cara mostrada, es decir, si se muestra color azul predecir azul y lo mismo para el color rojo. Que esta estrategia es la óptima en el juego puede razonarse de diversas formas:

Solución 1. La forma más sencilla de encontrar la solución es razonar que, de las tres tarjetas, dos tienen las caras del mismo color. Por tanto, en el experimento consistente en obtener una tarjeta al azar tenemos tres posibilidades (las tres tarjetas). Los casos favorables son las dos tarjetas con las dos caras iguales. Una simple aplicación de la regla de Laplace sirve para obtener la probabilidad pedida: *Probabilidad (Cara oculta = cara visible) = Probabilidad (dos caras iguales) = 2/3.* En la Figura 6.4.1 presentamos el espacio muestral en este experimento y en la Tabla 6.4.1 realizamos el análisis semiótico. Observamos que en esta solución no se utiliza la idea de experimento compuesto ni de dependencia o independencia, sino sólo la idea de experimento simple, suceso y espacio muestral, regla de Laplace y equiprobabilidad de las tarjetas.

Figura 6.4.1. Espacio muestral en el experimento descrito en la solución 1

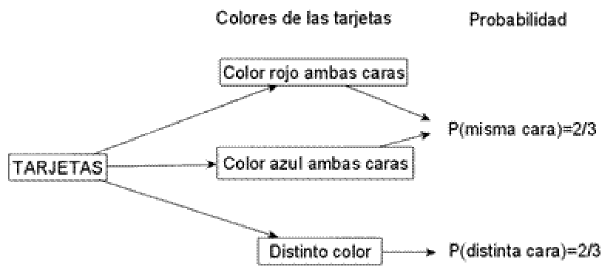


Tabla 6.4.1. Objetos matemáticos implícitos en la Solución 1

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, dependiendo del color mostrado
Lenguajes	- Tarjetas	- Espacio muestral
	- Color rojo en ambas caras; color azul en ambas caras; distinto color	- Sucesos posibles
	- *P*(misma cara), *P*(distinta cara)	- Probabilidad simple
	- Diagrama en árbol	- Representa el experimento y sus resultados
Conceptos	- Experimento aleatorio	- Sacar una tarjeta
	- Sucesos	- Tarjeta obtenida
	- Espacio muestral	- Conjunto de posibilidades
	- Probabilidad	- Cociente de casos favorables entre posibles
Procedimientos	- Cálculo de probabilidades	- Calcula las probabilidades de cada tarjeta
Propiedades	- Axioma de la unión	- La suma de las probabilidades es 1
	- Regla de Laplace	- La probabilidad de un suceso es el cociente entre los casos favorables entre los posibles

Solución 2. Podemos considerar el espacio muestral en el experimento compuesto de dos experimentos simples:

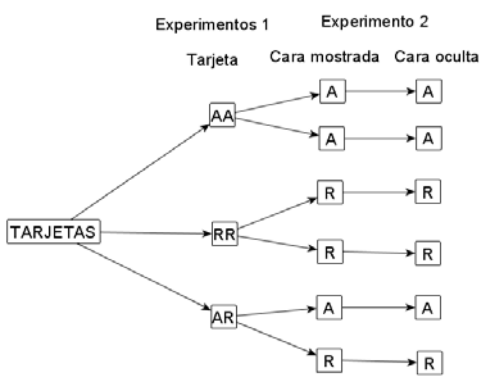
- Experimento 1: Tomar al azar una de las tres tarjetas. Cada una tiene probabilidad 1/3.
- Experimento 2: Mostrar una de las dos caras de la tarjeta, elegida al azar. En cada tarjeta, cada cara tiene una probabilidad 1/2. Respecto a los colores, en la tarjeta de dos colores, cada color tiene probabilidad 1/2; en la azul, la única posibilidad es que la cara oculta sea azul y en la roja, que sea roja.

Podemos representar este experimento mediante un diagrama en árbol (Figura 6.4.2). El espacio muestral consta de seis sucesos {*AA, AA, RR, RR, RA y AR*}, pues en las tarjetas con las dos caras iguales hay que considerar dos veces el color azul o el rojo, dependiendo de si se muestra la cara de arriba o la de abajo. El dato de que la cara mostrada es azul implica la reducción del espacio muestral {*AA, AA y AR*}. Aplicando la regla de Laplace, tenemos *P(oculta A/mostrada A)= 2/3*. En esta segunda solución se usa la idea de experimento compuesto (de tres experimentos) y de dependencia, pues la cara mostrada dependerá de la tarjeta elegida y la cara oculta de la cara mostrada. Se requiere un razonamiento combinatorio para enumerar todas las posibilidades del espacio muestral. En la Tabla 6.4.2 mostramos el análisis semiótico de la solución.

Tabla 6.4.2. Objetos matemáticos implícitos en la solución 2

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, dependiendo del color mostrado
Lenguajes	- Verbal	- Explicación de la situación
	- Gráfico	- Diagrama en árbol
	- Simbólico	- Expresar sucesos, probabilidades
Conceptos	- Experimento aleatorio	- Tarjeta que se elige; Cara que se muestra - Cara opuesta
	- Sucesos; espacio muestral	- Tarjetas (*RR*) (*AA*) (*RA*) - (*RR*) (*AA*) (*RA*) (*AR*)
	- Experimento compuesto	- Composición de los experimentos anteriores
	- Sucesos en el experimento Compuesto	- Producto cartesiano de los espacios muestrales anteriores
	- Frecuencia relativa	- Éxitos / número experimentos
	- Unión de sucesos	- Sucesos en uno u otro conjunto
	- Probabilidad clásica	- Proporción de casos favorables a posibles
	- Axiomas de probabilidad	- Explicitación de los axiomas
	- Probabilidad condicional	- Proporción de ocurrencia suceso respecto a la ocurrencia de otro
	- Regla de la suma	- Probabilidad de acertar
	- Regla del producto	- Probabilidad conjunta; dependencia
	- Variable aleatoria	- Número de aciertos
Procedimientos	- Enumeración del espacio muestral	- Conjunto de elementos en el espacio producto
	- Representación	- Construcción de un diagrama en árbol
	- Cálculo de probabilidades	- Aplicar la regla de Laplace
Propiedades	- Los experimentos son dependientes	- El número de caras azules depende de la tarjeta elegida en primer lugar
	- Reducción del espacio muestral	- Conocer una cara cambia el espacio muestral
	- Axioma de unión	- La probabilidad de la suma es suma de probabilidades en este caso
	- Regla del producto	- La probabilidad de la intersección es el producto de la probabilidad simple por la condicionada

Figura 6.4.2. Espacio muestral en el experimento descrito en la solución 2



Solución 3. El espacio muestral de este experimento será {*(AA), (AR), (RA), (RR)*}, donde los sucesos con las dos caras iguales tienen doble probabilidad que los que tienen caras diferentes (combinaciones de 2 elementos con repetición). Las probabilidades de estos sucesos son: *P(AA)=1/3; P(AR)=1/6; P(RA)=1/6; P(RR)=1/3; P(Mismo color)= P(AA)+ P(RR)=2/3.*

Figura 6.4.3. Espacio muestral del experimento descrito en la solución 3

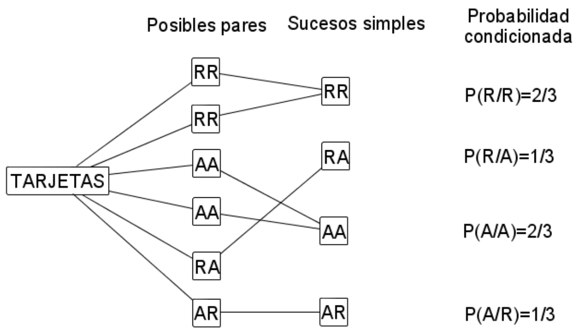


Tabla 6.4.3. Objetos matemáticos implícitos en la Solución 3

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, dependiendo del color mostrado
Lenguajes	- Verbal	- Explicación de la situación
	- Gráfico	- Diagrama en árbol
	- Tabular	- Tabla de resultados en el juego
	- Simbólico	- Expresar sucesos, probabilidades
	- Icónico (si se simula el juego)	- Iconos que representan los sucesos y resultados
Conceptos	- Experimento aleatorio	- Tarjeta que se elige - Cara que se muestra - Cara opuesta
	- Sucesos; espacio muestral	- Tarjetas: *(R/R) (A/A) (R/A)* *(R/R) (A/A) (R/A) (A/R)*
	- Experimento compuesto	- Composición de los experimentos anteriores
	- Sucesos en el experimento Compuesto	- Producto cartesiano de los espacios muéstrales anteriores
	- Unión de sucesos	- Sucesos en uno u otro conjunto
	- Probabilidad clásica	- Proporción de casos favorables a posibles
	- Axiomas probabilidad	- Explicitación de los axiomas
	- Regla de la suma	- Probabilidad de acertar
	- Regla del producto	- Probabilidad conjunta; dependencia
Procedimientos	- Enumeración del espacio muestral	- Desarrollar conjunto de elementos simples
	- Representación	- Construcción de un diagrama en árbol
	- Cálculo de probabilidades	- Aplicar la regla de Laplace
Propiedades	- Los experimentos son dependientes	- El número de caras azules o rojas depende de la tarjeta elegida en primer lugar

En esta solución el espacio muestral puede ser definido utilizando como sucesos simples los diferentes pares de resultados; es decir, el par *RR* se toma como un suceso simple sin distinguir qué cara roja salió antes, cosa que también ocurre con el suceso *AA*, aunque distinguimos como diferentes el suceso *AR* y el *RA*. Por tanto, al calcular las probabilidades de ocurrencia de los sucesos tenemos que tener en cuenta que el suceso simple *RR* aparece dos veces, siendo su probabilidad de ocurrencia de 2/6. El mismo razonamiento puede aplicarse al suceso *AA*, cuya probabilidad es 2/6. Pero al distinguir entre la tarjeta *AR* y la *RA* tenemos que la probabilidad de cada una de ellas es de 1/6 (Figura 6.4.1.3). En la Tabla 6.4.3 se presenta el análisis semiótico.

Solución 4: Aplicando la fórmula de la probabilidad compuesta para sucesos dependientes tenemos:

$$P(AA)=P(A)\cdot P(A/A)=1/2\cdot 2/3=1/3$$

$$P(RR)=P(R)\cdot P(R/R)=1/2\cdot 2/3=1/3$$

$$P(AR)= P(A)\cdot P(R/A)=1/2\cdot 1/3=1/6$$

$$P(RA)=P(R)\cdot P(A/R)=1/2\cdot 1/3=1/6$$

Por tanto, es más probable que la cara sea igual por ambos lados. En esta solución utilizamos la descomposición de la probabilidad de la intersección a partir de la definición de probabilidad condicionada.

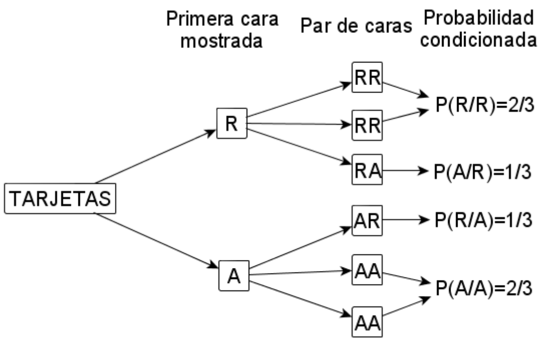
Tomamos el suceso simple *AA* como intersección del suceso “primera cara mostrada azul” (*A*) y “segunda cara mostrada azul” (*A*), por lo que a partir de la definición de probabilidad condicionada $P(A/A)=\frac{P(A\cap A)}{P(A)}$; $2/3=\frac{P(A\cap A)}{1/2}$, tenemos que $P(A\cap A)=P(A/A)\cdot P(A)=2/3\cdot 1/2=1/3$.

Calculamos el resto de probabilidades de la misma forma, teniendo en cuenta que las probabilidades condicionadas las podemos definir a partir del diagrama de la Figura 6.4.4.

Tabla 6.4.4. Objetos matemáticos implícitos en la solución 4

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, dependiendo del color mostrado
Lenguajes	- Verbal	- Explicación de la situación
	- Gráfico	- Diagrama en árbol
	- Tabular	- Tabla de resultados en el juego
	- Simbólico	- Expresar sucesos, probabilidades
	- Icónico (si se simula el juego)	- Iconos que representan los sucesos y resultados
Conceptos	- Experimento aleatorio	- Tarjeta que se elige; - Cara que se muestra - Cara opuesta
	- Sucesos; espacio muestral	- Tarjetas - *(A) (R)* - *(R/R) (R/R) (A/A) (A/A) (R/A) (A/R)*
	- Experimento compuesto	- Composición de los experimentos anteriores
	- Sucesos en el experimento Compuesto	- Producto cartesiano de los espacios muéstrales anteriores
	- Unión de sucesos	- Sucesos en uno u otro conjunto
	- Probabilidad clásica	- Proporción de casos favorables a posibles
	- Axiomas probabilidad	- Explicitación de los axiomas
	- Regla de la suma	- Probabilidad de acertar
	- Regla del producto	- Probabilidad conjunta; dependencia
Procedimientos	- Enumeración del espacio muestral	- Desarrollar conjunto de elementos simples
	- Representación	- Construcción de un diagrama en árbol
	- Cálculo de probabilidades	- Aplicar la regla de Laplace
	- Regla del producto	- Aplicación para calcular las probabilidades conjuntas
Propiedades	- Los experimentos son dependientes	- El número de caras azules o rojas depende de la tarjeta elegida en primer lugar

Figura 6.4.4. Espacio muestral del experimento descrito en la solución 4



Solución 5. Solución empírica. El alumno podría usar los datos obtenidos durante el juego y observar empíricamente cuál es la solución correcta, es decir, obtener la solución estimando las probabilidades a partir de la frecuencia relativa. Sería una comprobación empírica de la solución, aunque no llegaría a comprender por qué una estrategia es preferible a otra.

Tabla 6.4.5. Objetos matemáticos implícitos en la solución 5

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, dependiendo del color mostrado
Lenguajes	- Verbal	- Explicación de la situación
	- Tabular	- Tabla de resultados en el juego
	- Numérico: Frecuencias	- Frecuencia de resultados
	- Icónico (si se simula el juego)	- Iconos que representan los sucesos y resultados
Conceptos	- Experimento aleatorio	- Tarjeta que se elige - Cara que se muestra - Cara opuesta
	- Convergencia	- Tendencia de la frecuencia a un valor fijo
	- Unión de sucesos	- Sucesos en uno u otro conjunto
	- Estimación frecuencial de la probabilidad	- Límite de la frecuencia
	- Axiomas probabilidad	- Explicitación de los axiomas
	- Regla de la suma	- Probabilidad de acertar
Procedimientos	- Enumeración del espacio muestral	- Desarrollar conjunto de elementos simples

6.4.2. SEGUNDA PARTE. ANÁLISIS DIDÁCTICO

En esta sección realizamos un análisis didáctico de la tarea que permitirá, por un lado, prever la actividad matemática necesaria para completarla y con ello los aspectos del conocimiento común y especializado del contenido que se pueden adquirir en el desarrollo del taller. Seguidamente analizamos las posibles dificultades que pueden surgir en la solución. Al confrontar a los profesores con estas dificultades, se les ayudará a mejorar su conocimiento del contenido matemático y los estudiantes. Por último, se analizan las configuraciones didácticas. La reflexión en torno a las mismas servirá para mejorar el conocimiento didáctico del contenido y la enseñanza. Todo ello, de acuerdo a la terminología de Hill, Ball, y Schilling (2008).

6.4.2.1. OBJETOS MATEMÁTICOS IMPLÍCITOS

Una vez finalizado el taller, se procedió a estudiar con los asistentes los objetos matemáticos utilizados en cada una de las posibles soluciones, que dependerían de la solución encontrada. Según Godino (2009), este tipo de actividad permite explicitar e incrementar el conocimiento especializado del contenido por parte de los profesores.

El análisis de las soluciones correctas muestra los objetos matemáticos implícitos en este juego, resumidos en la Tabla 6.4.6. El profesor participante en el taller o bien un alumno (en caso de que el taller se hiciera con estudiantes de secundaria o universidad) usaría implícita o explícitamente algunos de estos objetos matemáticos, que varían dependiendo del tipo de solución encontrada. Hemos clasificado estos objetos de acuerdo a los considerados en nuestro marco teórico. Según Godino, Font y Wilhelmi (2008), este primer nivel de análisis se aplica tanto en la planificación como en la implementación de un proceso de estudio y pretende estudiar las prácticas matemáticas implicadas en dicho proceso. Permite describir una configuración epistémica global (previa y emergente) que determina las prácticas planificadas y realizadas.

Se observa que dependiendo del tipo de solución, hay una mayor profundidad del trabajo matemático. La solución empírica es la menos deseable, pues involucra un conjunto de objetos matemáticos notablemente menor y de menor complejidad que el resto. Asimismo, la primera solución no hace uso de las ideas de experimento compuesto, probabilidad condicional o independencia. Sería, por tanto deseable que el profesor condujera la clase hacia una de las otras soluciones.

Para cada una de las soluciones que hemos analizado, la tabla correspondiente sería una síntesis de la configuración epistémica utilizada. En nuestro caso se trataría de configuraciones previas, puesto que los profesores conocen el tema, pero si la actividad se lleva a cabo en un aula con estudiantes, los objetos probabilidad condicional, dependencia e independencia constituirían la parte emergente de la configuración.

Tabla 6.4.6. Objetos matemáticos implícitos en la situación, según solución

Tipos	Objetos matemáticos en la situación	Significado en la situación	S1	S2	S3	S4	S5
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, dependiendo del color mostrado	X	X	X	X	X
Lenguajes	- Verbal	- Explicación de la situación	X	X	X	X	X
	- Gráfico	- Diagrama en árbol - Representación icónica del juego	X X	X X	X X	X X	
	- Tabular	- Tabla de resultados en el juego					X
	- Simbólico	- Expresar sucesos, probabilidades	X	X	X	X	
	- Numérico: Frecuencias	- Frecuencia de resultados					X
	- Icónico (si se simula el juego)	- Iconos que representan los sucesos y resultados	X	X	X	X	X
Conceptos	- Experimento aleatorio	- Tarjeta que se elige - Cara que se muestra - Cara opuesta	X X X	X X X	X X X	X X X	X X X

	- Sucesos; espacio muestral	- Tarjetas - *(R/R) (A/A) (R/A)* - *(R/R) (A/A) (R/A) (A/R)* - *(R/R) (R/R) (A/A) (A/A) (R/A) (A/R)*	X X 	X X X	X X X	X X X	
	- Experimento compuesto	- Composición de los experimentos anteriores		X	X	X	
	- Sucesos en el experimento Compuesto	- Producto cartesiano de los espacios muéstrales anteriores		X	X	X	
	- Frecuencia relativa	- Éxitos / número experimentos	X	X	X	X	
	- Convergencia	- Tendencia de la frecuencia a un valor fijo					X
	- Intersección de sucesos	- Conjunto común de sucesos.				X	
	- Unión de sucesos	- Sucesos en uno u otro conjunto	X	X	X		X
	- Probabilidad clásica	- Proporción de casos favorables a posibles	X	X	X	X	
	- Estimación frecuencial	- Límite de la frecuencia					X
	- Axiomas probabilidad	- Explicitación de los axiomas	X	X	X	X	X
	- Probabilidad condicional	- Proporción de ocurrencia suceso respecto a la ocurrencia de otro		X		X	
	- Regla de la suma	- Probabilidad de acertar	X		X		X
	- Regla del producto	- Probabilidad conjunta; dependencia				X	
	- Variable aleatoria	- Número de aciertos	X	X	X	X	
	- Distribución de probabilidad	- Conjunto de valores con sus probabilidades	X	X	X	X	
	- Distribución discreta uniforme	- Conjunto finito de valores equiprobables	X	X	X	X	
	- Enumeración del espacio muestral	- Desarrollar conjunto de elementos simples	X	X	X	X	X
Procedimientos	- Representación; construcción de un diagrama en árbol	- Ver figuras	X	X	X	X	
	- Cálculo de probabilidades	- Aplicar la regla de Laplace	X	X	X	X	
	- Los experimentos son dependientes	- El número de caras azules depende de la tarjeta elegida en primer lugar		X	X	X	
	- Reducción del espacio muestral	- Conocer una cara cambia el espacio muestral		X		X	
Propiedades	- Axioma de unión	- La probabilidad de la suma es suma de probabilidades en este caso		X			
	- Regla del producto	- La probabilidad de la intersección es el producto de la probabilidad simple por la condicionada				X	
Argumentos	- Razonamiento deductivo	- Demostración de la solución	X	X	X	X	
	- Razonamiento empírico	- Comparar aciertos fallos con distintas estrategias					X

La finalidad de este primer nivel de análisis didáctico, en talleres para profesores es que los profesores reflexionen sobre el proceso de resolución del problema y los objetos que hemos descrito y lleguen a identificar algunos de ellos. No se pretende que lleguen a completar toda la tabla, pero sí que identifiquen en la actividad algunas de las ideas estocásticas fundamentales listadas por Heitele. También se trata de que el profesor observe cómo la matemática que se trabaja en la clase no sólo depende del problema planteado, sino de los recursos matemáticos proporcionados a los estudiantes.

6.4.2.2. ANÁLISIS DE POSIBLES CONFLICTOS DE LOS ESTUDIANTES

Godino, Font y Wilhelmi (2008) proponen un segundo nivel de análisis, que incluye los procesos matemáticos y conflictos semióticos. Los autores señalan que en toda práctica se identifica un sujeto agente (en nuestro caso los profesores) y un medio en el que dicha práctica se realiza (en nuestro caso el taller). Este nivel de análisis tiene como objeto describir la complejidad ontosemiótica de las prácticas matemáticas como factor explicativo de los conflictos semióticos que se producen en su realización.

Este problema, aparentemente simple, conlleva varias problemáticas a la hora de su interpretación y de su razonamiento, relacionadas mayoritariamente con una deficiente intuición sobre la probabilidad. A lo largo del taller, se pedirá a los participantes que expliciten sus soluciones y que traten de justificarlas. Puesto que el juego es paradójico, diferentes participantes defenderán con interés sus estrategias, algunas de las cuáles coincidirán con las correctas que acabamos de analizar. Pero otras serán erróneas, y debidas a algunos conflictos semióticos.

Se espera que la fase de debate y validación permita a los participantes detectar que sus razonamientos iniciales eran erróneos y en consecuencia, identificar algunos posibles conflictos semióticos en el desarrollo de la actividad, que serían las siguientes:

1. *Conflicto en la identificación del experimento compuesto y del correspondiente espacio muestral.* Es plausible que algunos profesores, aunque perciban que en el juego haya tres fichas rojas y tres azules, no construyan correctamente el espacio muestral, debido a un insuficiente razonamiento combinatorio. Como vimos en la evaluación llevada a cabo en el Estudio 2 (capítulo 4), algunos futuros profesores tienen dificultades al comprender la restricción del espacio muestral en la probabilidad condicionada. Esta dificultad se debe a que no se identifica con claridad la intersección de los sucesos en el espacio muestral producto (Sánchez, 1996; Truran y Truran, 1997). Por ello, algunos sujetos podrían considerar los dos

lados de cada tarjeta como sucesos independientes entre sí. De esta forma, podría haber encuestados que identifiquen sólo dos sucesos simples *R* y *A*, cada uno de ellos con tres posibilidades, *A, A, A* y *R, R, R* en lugar de identificar los sucesos (*RR*) (*AA*) (*RA*) y (*AR*) del espacio muestral producto. Este razonamiento llevará al error de suponer que no hay estrategia en este juego.

2. *Conflicto al identificar la probabilidades de los sucesos en el espacio muestral producto*. Incluso aunque se identifiquen correctamente los sucesos (*AR*) (*RA*) (*AA*) y (*RR*) en el experimento, es posible que se asigne la misma probabilidad a los sucesos, aunque de hecho (*AA*) y (*RR*) tienen doble probabilidad. Ello se debería a un razonamiento muy similar al de los estudiantes que al trabajar con dos dados, no ven la importancia del orden y consideran, por ejemplo, que el suceso “doble seis” y “un cinco y un seis” tienen la misma probabilidad. Este sesgo, denominado, *sesgo de equiprobabilidad* ha sido descrito en la investigación de Lecoutre y Durand (1988) y apareció en el Estudio 1.
3. *Conflicto en la identificación de dependencia de los experimentos.* Como indica Maury (1986), las principales dificultades al resolver un problema de probabilidad condicional aparecen cuando no se sabe si dos sucesos son o no independientes. En nuestro caso, no hemos proporcionado a los alumnos el listado de los sucesos del experimento, sino que son ellos quienes tienen que identificarlos y esto, según Maury, añade dificultad al problema. En la evaluación llevada a cabo en el Estudio 2 observamos la dificultad de algunos futuros profesores con la idea de independencia, dificultad que también es señalada por Sánchez (1996) o Kelly y Zwiers (1986). Por ello algunos encuestados podrían pensar que la probabilidad de la cara oculta no depende de la mostrada. Por el contrario, podrían pensar que los sucesos “color de la cara oculta” son equiprobables.
4. *Conflicto en la comprensión de la independencia de ensayos del mismo juego.* Aunque los experimentos en cada realización del juego son dependientes, cada vez que se juega, el resultado es independiente de los obtenidos con anterioridad. Esto ocurre en cualquier repetición de un experimento aleatorio, pero podría no ser percibido por los participantes. Por las mismas razones señaladas en el punto anterior, algunos profesores podrían suponer dependientes los resultados de la repetición del experimento. Éstos tendrían tendencia a predecir el resultado en cada

realización del juego, en función de los resultados anteriores. Por ejemplo si ha salido 5 rojas dobles podrían pensar que es más fácil que la siguiente extracción no se repita la tarjeta. Este razonamiento mostrar una conducta propia de la *heurística de la representatividad* (Kahneman, Slovic y Tversky, 1982), por la que se espera una convergencia rápida en una serie corta de ensayos.

6.4.2.3. ANÁLISIS DE CONFIGURACIONES DIDÁCTICAS

Otro punto es el análisis de las configuraciones didácticas puestas de manifiesto a lo largo del juego. Se pretende utilizar la situación para mostrar a los profesores una posible metodología de enseñanza de la probabilidad, tomada de la teoría de situaciones didácticas de Brousseau (1986; 1997). Con ello se quiere aumentar su conocimiento didáctico del contenido y la enseñanza (Hill, Ball, y Schilling, 2008). Para Brousseau, una situación didáctica se establece entre un grupo de alumnos y un profesor que usa un medio didáctico, incluyendo los problemas, materiales e instrumentos, con el fin específico de ayudar a sus alumnos a reconstruir un cierto conocimiento. Para lograr el aprendizaje, el alumno debe interesarse personalmente por la resolución del problema planteado. Diferencia cuatro tipos de situaciones didácticas:

- *Situación de acción:* Donde se resuelve el problema planteado, que consiste en buscar la mejor estrategia en el juego.
- *Situación de formulación-comunicación:* En la que el alumno debe poner por escrito para otra persona la solución hallada, lo que le hace usar el lenguaje matemático. En nuestro caso los participantes tratan de describir su estrategia por escrito.
- *Situación de validación:* Donde se pide a los alumnos las pruebas de que su solución es la correcta. En caso de que no sea así, el debate con los compañeros les permite descubrir los puntos erróneos. En el taller se trata de razonar por qué la estrategia da buenos resultados.
- *Institucionalización:* Tiene como fin dar un estatuto "oficial" al nuevo conocimiento aparecido, ponerse de acuerdo en la nomenclatura, formulación, propiedades, para que pueda ser usado en el trabajo posterior. En el taller, una vez realizadas varias series de jugadas se organizó un debate para llegar a un acuerdo sobre la solución correcta y reflexionar sobre lo que se ha aprendido fijando conjuntamente las

soluciones correctas.

Se pretende que los participantes en el taller identifiquen estas fases en el desarrollo de la actividad vivida por ellos mismos.

6.5. DESCRIPCIÓN DE LA MUESTRA Y EL CONTEXTO

La muestra que participó en el estudio estuvo formada mayoritariamente por profesores en ejercicio de educación secundaria y algunos profesores universitarios o profesores de secundaria en formación. Participaron sujetos de varios países, puesto que el taller se ofreció como actividad optativa en algunos congresos de estadística o en cursos dirigidos a profesores. También participaron algunos profesores en formación, en los cursos ofrecidos en la Universidad de Granada, dentro de la Licenciatura de Estadística, en la asignatura de Didáctica de la Estadística, elegida por una gran parte de los futuros licenciados de estadística y también como optativa por algunos estudiantes de la Licenciatura de Matemáticas.

En los cursos y talleres citados el número de participantes fue mayor. Todos los participantes conocían que el taller era parte de una investigación y se pidió a los que quisieran voluntariamente entregar el material y cuestionario completo, explicándoles la finalidad del estudio. La mayor parte de los participantes se prestaron voluntariamente para la colaborar con ella, proporcionando sus protocolos escritos y completándolos con interés. La segunda parte del taller (análisis didáctico) sólo se pudo llevar a cabo -en función del tiempo disponible- en una submuestra, pues el tiempo para el taller no fue el mismo en todos los contextos, variando desde dos a seis horas.

La muestra incluyó participantes de tres países: España, Portugal y México. Como podemos observar en la Tabla 6.5.1, la mayoría de las personas que realizaron el cuestionario eran españoles o habían realizado el cuestionario en España, en total un 59%. Casi una cuarta parte eran mexicanos y un 16,9% portugueses.

Tabla 6.5.1. Distribución de participantes por país de origen

País	Frecuencia	Porcentaje
España	98	59,0
Portugal	28	16,9
México	40	24,1
Total	166	100,0

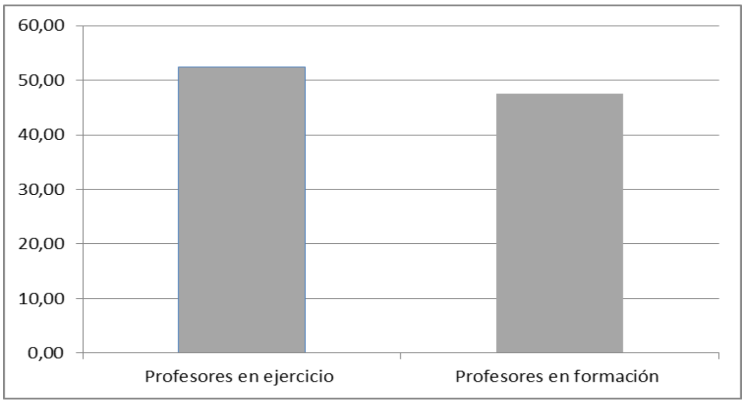
Tabla 6.5.2. Distribución de participantes según situación

Situación profesional	Frecuencia	Porcentaje
Profesor en ejercicio	87	52,4
Profesor en formación	79	47,6
Total	166	100,0

Como la Tabla 6.5.2 indica, el 47,6% de los profesores estaban en formación, la mayoría de ellos eran alumnos de la licenciatura de Ciencias y Técnicas Estadísticas o de la Licenciatura de Matemáticas. Estos profesores en formación participaron en el taller dentro de la asignatura Didáctica de la Estadística, impartida como optativa en el último año de la Licenciatura de Estadística o en el curso de Doctorado Didáctica de la matemática de la Universidad de Granada.

Los profesores portugueses en formación eran alumnos de un programa de maestría en Educación Matemática. El resto de participantes (52,4 %) estaba compuesto por profesores en ejercicio de educación superior o secundaria, la mayoría de ellos licenciados en Matemáticas, Estadística o bien en Educación Matemática (en caso de participantes de Portugal y México). En el taller impartido en México participaron también profesores en activo de estadística en diversas titulaciones universitarias, algunos de los cuales tenían una formación como ingenieros, médicos o farmacéuticos.

Figura 6.5.1. Distribución de participantes según situación



Segregando por países (Tabla 6.5.3), tenemos que en España el 61,9% de los encuestados eran futuros profesores, en Portugal el porcentaje es del 50%. En México no contamos con la participación de ningún estudiante. Podemos observar esta distribución en la Figura 6.5.2.

Figura 6.5.2. Distribución de profesores según país y situación

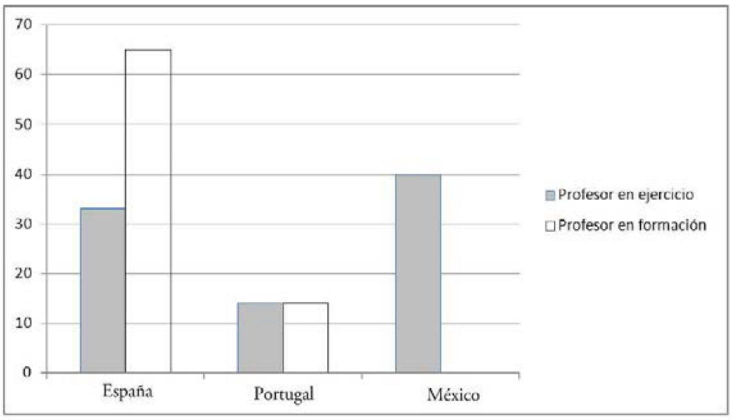


Tabla 6.5.3. Frecuencia (%) de profesores por países según situación

Situación profesional	España	Portugal	México	Total
Profesor en ejercicio	33 (38,1)	14 (50)	40 (100)	87 (52,4)
Profesor en formación	65 (61,9)	14 (50)	-	79 (47,6)
Total	98	28	40	166

Algunos profesores en activo proporcionaron datos del número de años de experiencia docente. Hubo profesores con hasta 36 años de experiencia docente (media=14,45; desviación típica=9,23). De los que dieron este dato, 63(38%) tenían experiencia docente a nivel de educación secundaria. El grupo mayoritario es el de 5 a 10 años de docencia, con un 30,2% de profesores y en el de menor porcentaje encontramos dos intervalos: con un 17,5%, los profesores de 10 a 15 años de ejercicio y los de 15 a 20 años de ejercicio.

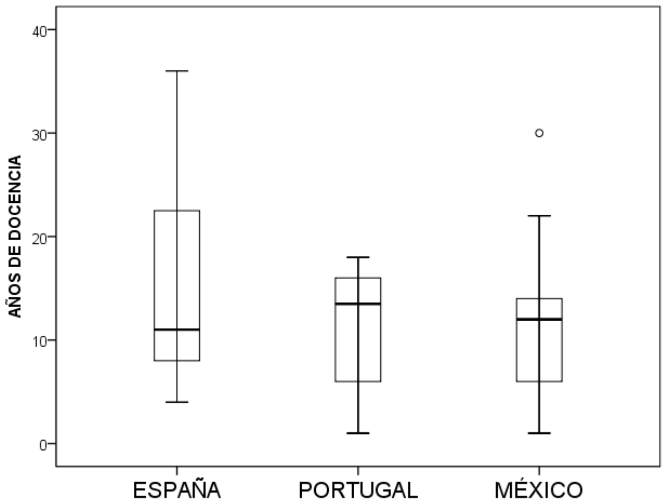
Tabla 6.5.4. Frecuencia (%) de profesores en ejercicio por años de docencia en secundaria

	España	Portugal	México	Total
Menos de 5 años	5 (12,8)	3 (21,4)	2 (20)	10 (15,8)
De 5 a 10 años	14 (35,9)	2 (14,3)	3 (30)	19 (30,2)
De 10 a 15 años	4 (10,3)	4 (28,6)	3 (30)	11 (17,5)
De 15 a 20 años	6 (15,4)	5 (35,7)	-	11 (17,5)
Más de 20 años	10 (25,6)	-	2 (20)	12 (19,0)
Total	39	14	10	63

Podemos observar las diferencias en las muestras en la Figura 6.5.3. En España el estrato con mayor porcentaje fue el de 5 a 10 años de docencia con un 35,9% y el de menor el de 10 a 15 años con un 10,3%. La media de años de docencia fue de 16,85. El valor mínimo fue de 4 años y el máximo de 36 años con una desviación típica de 10,27.

Al contrario de lo que ocurría en España, en el que todos los estratos estaban representados, en Portugal observamos que el estrato “más de 20 años de docencia” no está representado en la muestra. En este caso el estrato mayoritario es el de 15 a 20 años de docencia con un 35,7%, aunque el resto de intervalos también se encuentran bastante representados. La media de años de docencia fue de 11,36. El valor mínimo fue de 1 año y el máximo de 18 años con una desviación típica de 5,85.

Figura 6.5.3. Años de docencia de secundaria por país



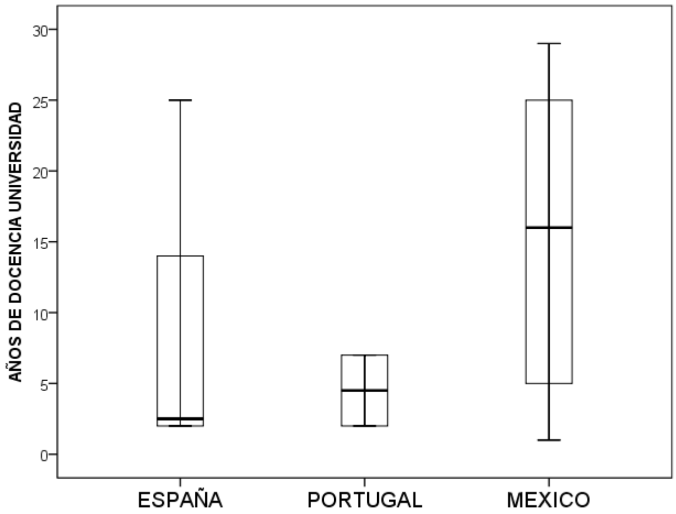
En el caso de México hay dos estratos mayoritarios: Respecto a la docencia en secundaria, el de 5 a 10 años y el de 10 a 15 años de docencia con un 30% de profesores, no estando representado el de 15 a 20 años. La media de años de docencia fue de 11,07. El valor mínimo fue de 1 año y el máximo de 30 años con una desviación típica de 7,27.

Tabla 6.5.5. Frecuencia (%) de profesores en ejercicio por años de docencia en universidad

	España	Portugal	México	Total
Menos de 5 años	3 (75)	1 (50)	4 (30,7)	8 (42,1)
De 5 a 10 años	-	1 (50)	1 (7,7)	2 (10,5)
De 10 a 15 años	-	-	1 (7,7)	1 (5,3)
De 15 a 20 años	-	-	1 (7,7)	1 (5,3)
Más de 20 años	1 (25)	-	6 (46,2)	7 (36,8)
Total	4	2	13	19

De entre los encuestados, solamente 19 (11,5%) proporcionaron datos sobre “años de docencia en universidad”, docencia que varió entre 1 y 29 años, con una media de 12,95 años, y desviación típica de 10,89. Vemos que el número de profesores universitarios que participaron en el taller fue pequeño y la mayoría de origen mexicano.

Figura 6.5.4. Distribución del número de años de docencia universitaria por países



En resumen, se trata de una muestra variada, tanto por su situación (en formación o ejercicio), formación básica (matemáticas, estadística u otra) y años de docencia (universidad o secundaria), además de diferente contexto geográfico.

En todos los talleres, los participantes mostraron un gran interés y se involucraron activamente en la actividad, con lo que se consiguió una de las finalidades principales, que es provocar la reflexión de los profesores. En lo que sigue, analizamos los resultados del análisis de los protocolos escritos proporcionados voluntariamente por los participantes.

6.6. CONOCIMIENTOS MATEMÁTICOS SOBRE PROBABILIDAD CONDICIONAL

Recogidos los protocolos, se realizó un análisis cualitativo del contenido de los mismos (Ghiglione y Matalón, 1991). Utilizando un proceso cíclico e inductivo, típico de este tipo de análisis, se definieron las principales variables y categorías analizadas en el estudio. Para asegurar la fiabilidad, se revisaron varias veces los protocolos, consultando con otros investigadores del grupo de investigación los casos en que la codificación no fuese clara. A continuación realizamos el análisis de las variables deducidas de los protocolos escritos.

6.6.1. ESTRATEGIAS INICIALES

Consideremos, en primer lugar, las estrategias que los participantes consideraron correctas al inicio del juego (estrategias iniciales). Describimos a continuación resumidamente la estrategia, mostrando un ejemplo y, mediante una tabla de análisis semiótico, identificamos los objetos matemáticos utilizados, así como posibles conflictos semióticos en caso de estrategias incorrectas. En la Tabla 6.6.1 aparecen las frecuencias de las diferentes estrategias seguidas por los participantes en la primera racha de 10 jugadas.

E1. Apostar al mismo color de la cara que se ve. Se trataría de apostar "rojo" si la cara mostrada es roja y "azul" si la cara mostrada es de dicho color. Sería la mejor estrategia en el juego, a la que el sujeto habrá llegado mediante alguno de los razonamientos correctos ya descritos.

Tabla 6.6.1. Objetos matemáticos implícitos en la estrategia E1

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, dependiendo del color mostrado
Lenguajes	- Color	- Suceso posible; que puede ser rojo o azul
	- Color que se mostraba	- Resultado del experimento "cara mostrada"
	- Color que se predecía	- Suceso esperado en el experimento "cara oculta" - Suceso más probable, dada el resultado de la cara mostrada
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara; acertar el color
	- Sucesos	- Color de las caras azul y rojo
	- Espacio muestral; espacio muestral reducido	- Conjunto de posibilidades; conjunto de posibilidades cuando se reduce el espacio muestral
	- Experimento compuesto	- Sacar una tarjeta y mostrar una de las caras
	- Probabilidad	- Proporción de casos favorables a posibles
	- Probabilidad condicional $P(B/A)$	- Proporción de casos favorables, con una condición dada
	- Variable aleatoria	- Número de aciertos siguiendo la estrategia
	- Esperanza	- Valor esperado del número de aciertos
Procedimientos	- Cálculo de probabilidades condicionadas	- Calcula las probabilidades de rojo y azul en función de lo que ha salido
	- Comparación de probabilidades	- Compara la probabilidad de azul y rojo en función de lo que ha salido
Propiedades	- Dependencia	- La cara oculta y mostrada depende de la ficha que se extrajo
	- Axioma de la unión	- La suma de las probabilidades es 1
	- Convergencia	- A la larga el suceso más probable aparece

Los alumnos que eligen esta estrategia, por tanto, han logrado calcular las probabilidades de que la cara mostrada sea igual o diferente de la oculta. Hemos de notar que dicho problema lo han debido plantear los mismos participantes, pues la consigna dada fue únicamente encontrar una estrategia que permita ganar los más puntos posibles. A continuación presentamos un ejemplo, que se analiza en la Tabla 6.6.1. Aunque la respuesta es muy escueta, involucra una gran cantidad de objetos matemáticos, no suficientemente explicitados, pero necesarios para llegar a la estrategia.

"Si, el color que se mostraba de la carta inicialmente era el color que predecía".

En el ejemplo, el sujeto debe percibir que se trata de un experimento aleatorio (o bien un experimento compuesto) y deducir el espacio muestral del experimento. Puesto que el resultado no es seguro, no se puede pedir la estrategia que produzca un acierto, sino la que a la larga proporcione mayor número de aciertos, y por tanto, involucra las ideas de esperanza matemática, variable aleatoria y convergencia. Además del cálculo de probabilidades simples y condicionales, y el axioma de la unión.

E2. Predecir el color contrario del que se muestra. Son los participantes que apuestan a rojo cuando el color mostrado es azul y a azul cuando se enseña el color rojo. En esta estrategia, no se aprecia la dependencia de los experimentos, lo cual puede deberse a varios motivos que se clarificarán cuando se pida demostrar la estrategia. Corresponderán en muchos casos a una de las dificultades descritas, la falacia del jugador, donde se espera que una serie corta de ensayos se equilibren los sucesos, y en otros casos a la falta de comprensión de la independencia de ensayos (Kelly y Zwiers, 1986; Sánchez, 1996). Otra posibilidad es una incorrecta percepción del espacio muestral considerando las tres caras rojas y tres caras azules como si fuesen sucesos simples de un experimento con tres resultados. Otra posible razón para seguir esta estrategia es que el alumno manifieste la heurística de la representatividad (Kahneman, Slovic y Tversky, 1982), esperando un cambio de color en una racha corta. Todos estos errores aparecieron en el Estudio 2. Un ejemplo se muestra a continuación y se analiza en la Tabla 6.2.2, donde se observa que, aunque se usan muchos objetos matemáticos, aparecen varios conflictos asociados a su aplicación en este caso particular.

"Aposté más a que la cara oculta fuese la opuesta a la que salió".

Tabla 6.6.2. Objetos matemáticos implícitos en la estrategia E2

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, según el color mostrado
Lenguajes	- Cara	- Suceso posible; que puede ser rojo o azul
	- La que salió	- Resultado del experimento "cara mostrada"
	- Cara oculta	- Suceso esperado en el experimento "cara oculta" suceso más probable, dada el resultado de la cara mostrada
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara
	- Sucesos	- Color de las caras azul y rojo
	- Espacio muestral	- Conjunto de posibilidades; conflicto en la enumeración de posibilidades
	- Suceso contrario	- Cara opuesta de la que salió
	- Probabilidad de cada suceso	- *Conflicto*: considera más probable la cara opuesta del color mostrado
Procedimientos	- Cálculo de probabilidades condicionadas	- Calcula las probabilidades de rojo y azul en función de lo que ha salido
	- Comparación de probabilidades	- Compara la probabilidad de azul y rojo en función de lo que ha salido. *Conflicto* al no comprender la independencia de ensayos
Propiedades	- Axioma de la unión	- La suma de las probabilidades es 1
	- Convergencia	- A la larga el suceso más probable aparece

E3. Considerar que no hay estrategia posible para mejorar la predicción, debido a la aleatoriedad. Esta estrategia es la más usual, ya que la intuición sugiere que el color de la cara oculta es independiente del de la cara mostrada, pues al mostrar por primera vez un color se descarta la opción del doble color contrario. Es decir, el profesor piensa que cuando la cara mostrada, por ejemplo, es azul se descarta la opción de que salga la carta roja/roja por lo que sólo existiría la opción de que fuese azul/roja o azul/azul y, debido al sesgo de equiprobabilidad (Lecoutre y Durand, 1988), piensa que el resultado de la cara oculta es aleatorio. El participante sólo tiene en cuenta lo que sucedió (la cara mostrada) pero no el conjunto de casos de lo que podría haber ocurrido (todo el espacio muestral). Por otro lado, olvida que la tarjeta elegida se puede mostrar de dos formas diferentes pues tiene dos caras, es decir, no percibe que la probabilidad de que la cara oculta sea igual a la mostrada es el doble de que sea diferente. Un ejemplo se muestra a continuación y se analiza en la Tabla 6.6.3.

"*Como hay la mitad roja y la mitad azul la predicción correcta depende del azar*".

Tabla 6.6.3. Objetos matemáticos implícitos en la estrategia E3

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, según el color mostrado
Lenguajes	- Mitad roja y mitad azul	- Sucesos simples del experimento, conflicto al asumir la equiprobabilidad
	- Predicción correcta	- Resultado predicho dela cara oculta
	- Azar	- El resultado de la cara oculta es aleatorio; concepción del azar como causa
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara
	- Experimento compuesto	- Cara mostrara /cara oculta; conflicto al asumir la independencia de los dos experimentos
	- Sucesos	- Color de las caras azul y rojo
	- Espacio muestral	- Conjunto de posibilidades
	- Probabilidad	- Cociente de casos favorables entre posibles
	- Probabilidad de cada suceso	- *Conflicto*: considera que no hay estrategia, al ser aleatorio
	- Aleatoriedad	- Aleatoriedad en el experimento
Procedimientos	- Cálculo de probabilidades	- Calcula las probabilidades de salir rojo o azul, equiprobabilidad
Propiedades	- Axioma de la unión	- La suma de las probabilidades es 1
	- Regla de Laplace	- La probabilidad de un suceso es el cociente entre los casos favorables entre los posibles

E4. Elección de color rojo (o color azul) en todos los ensayos. Esta estrategia está relacionada con la *E3,* ya que el alumno considera que si apuesta siempre al mismo color en la mitad de las ocasiones acertará. En consecuencia se repiten los conflictos mencionados en el caso anterior. Podría también añadirse la creencia que si un color (el color fijo) no ha salido en una serie de ensayos, aumenta su probabilidad, es decir, se muestra la falacia del jugador. Mientras el primer razonamiento está relacionado con la falta de percepción de la dependencia de los dos experimentos aleatorios (cara mostrada y oculta) en el juego, el segundo se relaciona con la falta de percepción de la independencia de ensayos sucesivos del experimento compuesto. Un ejemplo se muestra a continuación y se analiza en la Tabla 6.6.4.

"*Jugar siempre al rojo, ya que el resultado se debe al azar*".

Tabla 6.6.4. Objetos matemáticos implícitos en la estrategia E4

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, dependiendo del color mostrado
Lenguajes	- Jugar siempre al rojo	- Estrategia seguida, *conflicto* al asumir la equiprobabilidad; *conflicto* al asumir la dependencia de ensayos del mismo experimento
	- Azar	- El resultado de la cara oculta es aleatorio

Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara
	- Sucesos	- Color de las caras azul y rojo
	- Espacio muestral	- Conjunto de posibilidades
	- Probabilidad	- Cociente de casos favorables entre posibles
	- Experimento compuesto	- *Conflicto* al asumir la independencia
Procedimientos	- Cálculo de probabilidades	- Calcula las probabilidades de salir rojo
Propiedades	- Regla de Laplace	- La probabilidad de un suceso es el cociente entre los casos favorables entre los posibles

E5. Alternando colores. Al igual que ocurría con la anterior, la estrategia de ir alternando los colores responde al sesgo de equiprobabilidad (Lecoutre y Durand, 1988). De nuevo aparece la idea de que si un color no ha aparecido, aumenta su probabilidad (falta de percepción de la independencia de ensayos, que apareció en el Estudio 2), por lo que se produce otra vez la falacia del jugador, esperando convergencia en una serie corta. Del mismo modo falla la percepción de la dependencia de los experimentos simples que componen el experimento compuesto. Un ejemplo se muestra a continuación y se analiza en la Tabla 6.6.5, donde observamos un conflicto al formar el espacio muestral, asumiendo que los sucesos elementales son las caras y no las tarjetas.

> "*Atendiendo que existen tres caras rojas y tres azules la estrategia sería tomar como cara prevista azul para las cinco primeras y rojo para las siguientes*".

Tabla 6.6.5. Objetos matemáticos implícitos en la estrategia E5

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, según el color mostrado
Lenguajes	- Existen tres caras rojas y tres azules	- Espacio muestral, *conflicto* al considerar las caras como sucesos, en vez de considerar el par de caras
	- Azul para las cinco primeras y rojo para las siguientes	- El resultado de la cara oculta es aleatorio, por lo que no importa la predicción que se haga - *Conflicto*: considera que el resultado no depende de la cara mostrada
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara
	- Sucesos	- Color de las caras azul y rojo
	- Espacio muestral	- Conjunto de posibilidades; enumeración incorrecta
	- Probabilidad	- Cociente de casos favorables entre posibles
	- Azar	- Aleatoriedad; concepción de azar como un proceso equilibrado
	- Equiprobabilidad	- Igualdad de probabilidad de sucesos
Procedimientos	- Cálculo de probabilidades	- Calcula las probabilidades de salir rojo, equiprobabilidad
Propiedades	- Regla de Laplace	- La probabilidad de un suceso es el cociente entre los casos favorables entre los posibles

E6. Uso de los resultados anteriores para la predicción. En esta estrategia el estudiante otorga al juego la característica de tener memoria por lo que supone que al haberse repetido un resultado varias veces es más fácil que salga el contrario. Se percibe una concepción incorrecta sobre la independencia de las repeticiones del experimento, ya comentada. Un ejemplo se muestra a continuación y se analiza en la Tabla 6.6.6.

> *"Si hay 3 veces seguidas que sale el mismo color y las 2 primeras tiene el mismo color en la cara de atrás, la tercera vez elijo otro color porque me resulta más creíble si de 3 veces que sale el mismo color, alguna de ellas ha de ser diferente por detrás".*

Tabla 6.6.6. Objetos matemáticos implícitos en la estrategia E6

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, dependiendo del color mostrado
Lenguajes	- 3 veces seguidas que sale el mismo color	- Resultado obtenido en el juego
	- Resulta más creíble	- Tener más probabilidad de ocurrencia
	- Si de 3 veces que sale el mismo color, alguna de ellas ha de ser diferente por detrás	- *Conflicto*: los resultados dependen de los anteriores
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara
	- Sucesos	- Color de las caras azul y rojo
	- Espacio muestral	- Conjunto de posibilidades
	- Probabilidad	- Cociente de casos favorables entre posibles
	- Probabilidad condicionada	- *Conflicto*: considera que la probabilidad depende de los resultados de simulaciones anteriores no dé la cara mostrada
Procedimientos	- Cálculo de probabilidades	- Calcula las probabilidades de salir rojo, equiprobabilidad
Propiedades	- Regla de Laplace	- La probabilidad de un suceso es el cociente entre los casos favorables entre los posibles

E7. Cambiar las estrategias a lo largo de la secuencia de ensayos. En esta estrategia, algunos alumnos viendo que sus apuestas no acertaban con la estrategia elegida inicialmente, optaban por cambiar de forma de pensar y variar a una nueva estrategia a la mitad del juego. Generalmente se da en los participantes que se sienten inseguros o perdidos y tratan de adivinar al azar la estrategia pues no son capaces de razonar de una forma probabilística sobre el juego. En otros casos, se describe la estrategia correcta a mitad del juego, por lo que se opta por cambiar. Un ejemplo se muestra a continuación y se analiza en la Tabla 6.6.7.

> *"Primero traté de ver las veces que salía la tarjeta de un determinado color apostando al mismo que salía y luego aposté siempre al color rojo".*

Tabla 6.6.7. Objetos matemáticos implícitos en la estrategia E7

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de la ficha oculta, según el color mostrado
Lenguajes	- Traté de ver las veces que salía la tarjeta de un determinado color	- Simulación, mediante la observación tratar de predecir el comportamiento
	- Apostando al mismo que salía	- Tener más probabilidad de ocurrencia si es la misma cara
	- Aposté siempre al color rojo	- *Conflicto*: como entiende que los resultados son aleatorios y equiprobables toma siempre un resultado posibles
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara
	- Sucesos	- Color de las caras azul y rojo
	- Espacio muestral	- Conjunto de posibilidades
	- Probabilidad	- Cociente de casos favorables entre posibles
	- Probabilidad de cada suceso	- Conflicto: considera que la probabilidad no depende de la cara mostrada
Procedimientos	- Cálculo de probabilidades	- Calcula las probabilidades de salir rojo, equiprobabilidad
Propiedades	- Regla de Laplace	- La probabilidad es el cociente entre los casos favorables y los posibles

E8. Propiedades físicas de las tarjetas. Algunos participantes intentaron identificar el comportamiento del juego en función de las características físicas de las tarjetas utilizadas. Por ejemplo, se fijaron en que la tarjeta doble roja tenía una pequeña mancha o estaba doblada y usaron estas características para utilizar una estrategia carácter no probabilístico. Estos participantes no son capaces de abordar el problema mediante un análisis probabilístico.

Comparación por país

En la Tabla 6.6.8 presentamos las diferentes estrategias iniciales de los participantes, junto con la frecuencia y porcentajes globales y segregados por países donde se aplicó el cuestionario. Lo primero que sorprende es que al comienzo del juego, sólo la cuarta parte de los participantes la eligen, a pesar de su alta preparación y de la aparente sencillez del problema. Ello es debido al carácter contra intuitivo del juego. Vemos que se consigue nuestro primer objetivo, que es lograr un problema que no sea resoluble de forma inmediata por los profesores participantes. Esto sugiere también la existencia de sesgos en el razonamiento sobre probabilidad condicional, aun cuando la mayoría de los participantes tenían un título de licenciado en matemáticas, estadística o ciencias. Los resultados fueron peores que los descritos por Batanero, Godino y Roa (2004), donde el 38% de los participantes eligió la estrategia correcta en el comienzo del juego.

Tabla 6.6.8. Frecuencia (y porcentaje) de estrategia inicial según país

	País			
Estrategia inicial	España (n=98)	Portugal (n=27)	México (n=41)	Total (n=166)
E1. Apostar al mismo color de la cara que se ve (correcta)	33 (33,7)	5 (18,5)	3 (7,3)	41 (24,7)
E2. Predecir el color contrario del que se muestra	8 (8,2)	3 (11,1)	2 (4,9)	13 (7,8)
E3. Considerar que no hay estrategia en la predicción (aleatoriedad)	38 (38,8)	14 (51,9)	27 (65,8)	79 (47,6)
E4. Elección de color rojo (color azul) en todos los ensayos	3 (3,0)	1(3,7)	1 (2,4)	5(3,0)
E5. Alternando colores	4 (4,1)	3 (11,1)	0 (0)	7 (4,2)
E6. Uso de los resultados anteriores para la predicción	7 (7,1)	1 (3,7)	4 (9,8)	12 (7,2)
E7. Cambiar las estrategias a lo largo de la secuencia de ensayos	4 (4,1)	0 (0)	4 (9,8)	8 (4,8)
E8. Propiedades no físicas de las tarjetas	1(1,0)	0(0)	0 (0)	1 (0,7)

El porcentaje de participantes que da la respuesta correcta varía entre los diferentes países, siendo bastante mayor en el grupo español (33,7%). Pensamos que ello es debido a que este grupo contenía una proporción grande de estudiantes de estadística del último año de carrera. En el caso opuesto, está México en el que se obtuvieron los peores resultados, a pesar de que todos ellos eran profesores en activo de estadística en secundaria o a nivel universitario en diferentes especialidades. Solamente un 7,3% de los participantes mexicanos acertaron la estrategia inicialmente, lo que creemos que es debido a la heterogeneidad del grupo y en particular a su formación inicial (que incluía psicólogos, médicos, ingenieros, aunque también matemáticos y estadísticos). Este resultado nos alerta de la posibilidad de que las intuiciones incorrectas sobre la probabilidad condicional también afecten a otros profesionales que, no siendo licenciados en matemáticas o estadística, impartan docencia universitaria de estadística, situación muy común hoy día, tanto en España como en el extranjero.

Portugal también destaca por su bajo porcentaje de acierto inicial (18,5%), aunque la muestra estaba formada por estudiantes de la Licenciatura de Matemáticas y profesores de secundaria de matemáticas. La formación estadística excesivamente formal durante la Licenciatura de Matemáticas y la falta de experiencia con simulaciones y juegos de probabilidad, así como falta de formación sobre didáctica de la probabilidad, podrían explicar estos resultados.

La estrategia más elegida inicialmente fue considerar que no hay estrategia posible para mejorar la predicción, esto es, pensar que el proceso es impredecible porque se basa en la aleatoriedad. Como hemos analizado, esta creencia se relaciona con problemas en la percepción de la dependencia de los experimentos implicados en el juego, que apareció en algunos profesores en los estudios 1 y 2. El 47,6% de los encuestados dio una solución relacionada con esta estrategia. Si segregamos los resultados por países vemos que en México dos tercios de los alumnos indicaron que no había estrategia inicial o que esta era aleatoria. En Portugal más de la mitad se decidieron por la aleatoriedad y en España sólo el 38,8% se decidieron por la falta de estrategia.

También cabe destacar la estrategia de elegir siempre el mismo color en todas las elecciones, que, aunque con escasa frecuencia implica una concepción incorrecta de la independencia de ensayos sucesivos en un mismo experimento, esperando la convergencia en una serie corta de ensayos.

Otras estrategias incorrecta seguidas por los encuestados fueron el uso de los resultados anteriores para predecir los resultados, un 7,2%, que indica una percepción incorrecta de la independencia y la alternancia de colores, un 4,2% atribuyéndole al juego la característica de tener memoria, lo cual es imposible. Segregando por países vemos que casi el 10% de los encuestados en México y el 7,1% de los españoles se decantaron por la opción de tener en cuenta los resultados anteriores, lo que demuestra una mala interpretación del término probabilidad en estos casos. Contrastan estos resultados con los de Portugal en el que el 3,4% optan por esta estrategia pero un 11,1% se decantan por la alternancia de colores. La estrategia contraria a la correcta de apostar al color contrario del que se muestra también muestra una falta de percepción de la independencia se dio en un 7,8% de los casos, siendo por países de 8,2%, 11,1% y 4,9% respectivamente para España, Portugal y México.

Clasificando las estrategias en correcta o incorrectas obtenemos la Tabla 6.6.9, donde observamos una mayor proporción de estrategias iniciales correctas en España. La diferencia en proporción de estrategias correctas/incorrectas inicialmente entre países fue estadísticamente significativa (Chi-cuadrado= 10,6; 2 g.l.; *p*= 0,005).

Tabla 6.6.9. Resumen de la distribución de estrategias iniciales

Estrategia	España	Portugal	México	Total
Correcta	33 (33,7)	5 (18,5)	3 (7,3)	41 (24,7)
No percibe la independencia	15 (15,3)	4 (14,8)	6 (14,7)	25 (15,0)
No hay	50 (51,0)	18 (66,7)	32 (78,0)	100 (60,3)
Total	98	27	41	166

Comparación por situación

Tabla 6.6.10. Frecuencia (y porcentaje) de estrategia inicial según situación

Estrategia	En formación (n=79)	En ejercicio (n=87)
E1. Apostar al mismo color de la cara que se ve (correcta)	23 (28,7)	19 (21,8)
E2. Predecir el color contrario del que se muestra	9 (11,2)	4 (4,6)
E3. Considerar que no hay estrategia en la predicción (aleatoriedad)	33 (41,3)	46 (52,9)
E4. Elección de color rojo (color azul) en todos los ensayos	3 (3,7)	3 (3,4)
E5. Alternando colores	-	3 (3,4)
E6. Uso de los resultados anteriores para la predicción	7 (8,9)	4 (4,6)
E7. Cambiar las estrategias a lo largo de la secuencia de ensayos	5 (6,2)	7 (8,1)
E8. Propiedades no físicas de las tarjetas	-	1 (1,2)

Tabla. 6.6.11. Resumen de la distribución de estrategias iniciales

Estrategia	En formación	En ejercicio
Correcta	23 (28,8)	19 (21,8)
No percibe la independencia	23 (28,8)	3 (3,4)
No hay	33 (41,4)	65 (74,8)
Total	79	87

Al clasificar las estrategias por situación del profesor (en formación o ejercicio) obtenemos las Tablas 6.6.10 y 6.6.11. Comparativamente, son mejores los resultados de los profesores en formación, debido posiblemente a que en los cursos más recientes de estadística, se comienza ya a trabajar con simulaciones o recursos en Internet que permiten una mejor formación de las intuiciones. Por el contrario, los profesores en ejercicio, además de involucrar un mayor número de licenciados en especialidades no matemáticas ni estadística, estudiaron hace bastante tiempo la probabilidad con un enfoque excesivamente formal.

En la Tabla 6.6.11, donde se presentan los datos resumidos por tipo de profesor, observamos aproximadamente el mismo porcentaje de respuestas correctas en ambos grupos. Hay, sin embargo, mayor problema inicial de falta de percepción de la independencia, problema ya encontrado en los Estudios 1 y 2, en los profesores en formación, aunque, a su vez los profesores en ejercicio, consideran en mayor medida que no hay una estrategia en este juego. Puesto que la frecuencia en una celda es pequeña, hemos preferido no aplicar el Chi-cuadrado en este caso. Si se compara simplemente estrategias correctas e incorrectas, observamos poca diferencia entre los profesores en formación y ejercicio, inicialmente (Chi-cuadrado= 1,16; 1 g.l.; p= 0,28) con menos del

30 % de estrategias correctas en los dos grupos. Concluimos la similitud de intuiciones incorrectas en los dos colectivos y la utilidad para ambos de la actividad formativa propuesta.

6.6.2. ESTRATEGIAS FINALES

Los resultados analizados indican que en el transcurso del taller se produce una variedad de estrategias iníciales, por lo que creemos conseguidos dos de nuestros objetivos para esta sesión: Plantear una situación verdaderamente problemática para los alumnos que dé lugar a una variedad de soluciones y propiciar el posterior debate y reflexión. Al explicar a los participantes que el juego estaba basado en una paradoja que costó varios años en ser resuelta por matemáticos célebres, les animó a tratar de averiguar por qué el razonamiento más intuitivo era incorrecto, consiguiéndose la devolución de la situación en el sentido de Brousseau (1986, 1997).

Se continúa el juego con otra racha de 10 jugadas, permitiendo al profesor que continúe con la estrategia inicial o bien cambie a una de las ya descritas. Periódicamente se pregunta a los participantes qué estrategia siguen y se listan en la pizarra, aunque no se permite todavía dar un razonamiento de por qué se elige una estrategia. En cambio si se comparan los puntos obtenidos por unos y otros profesores, según la estrategia seguida. El debate se produce en forma espontánea, y el formador de profesores deja que sean los mismos participantes en el taller los que por ellos mismos, vayan decantándose hacia la solución correcta, sin imponer en ningún momento sus ideas al grupo. Con ello se trata de propiciar un ejemplo de metodología constructivista de aprendizaje, basada en la resolución de problemas.

Finalizadas varias repeticiones del juego, algunos participantes han cambiado su estrategia inicial. Preguntados al final por la estrategia, encontramos las recogidas en la Tabla 6.6.12. La estrategia mayoritaria al finalizar la actividad es la correcta (62,6%), en concordancia con los resultados de Batanero, Godino y Roa (2004), que obtuvieron un 68% de estrategias finales correctas. Como podemos comprobar, la realización del juego, junto con el análisis de los resultados que se van obteniendo y la comparación del éxito en las diferentes estrategias ayuda a la resolución del problema ya que de un 24,7% de participantes que contestaron adecuadamente en la primera serie de ensayos se pasa al

62,6% al finalizar el taller. Observamos de nuevo la diferencia entre países, pero esta diferencia se explica por las diferencias encontradas en el inicio del taller y las diferencias de formación inicial sugerida con anterioridad.

Tabla 6.6.12. Frecuencia (y porcentaje) de estrategia final por país

Estrategias considerado correctas por los futuros docentes	España (n=98)	Portugal (n=27)	México (n=41)	Total (n=166)
E1. Apostar al mismo color de la cara que se ve (correcta)	71 (72,5)	10 (37,0)	23 (56,1)	104 (62,6)
E2. Predecir el color contrario del que se muestra	5 (5,1)	4 (14,9)	2 (4,9)	11 (6,7)
E3. Considerar que no hay estrategia en la predicción (aleatoriedad)	12 (12,2)	7 (25,9)	7 (17,1)	26 (15,6)
E4. Elección de color rojo (color azul) en todos los ensayos	1 (1,0)	2 (7,4)	1 (2,4)	4 (2,4)
E5. Alternando colores	5 (5,1)	3 (11,1)	3 (7,3)	11 (6,7)
E6. Uso de los resultados anteriores para la predicción	3 (3,1)	-	5 (12,2)	8 (4,8)
E7. Cambiar las estrategias a lo largo de la secuencia de ensayos	-	1 (3,7)	-	1 (0,6)
E8. Propiedades no físicas de las tarjetas	1 (1,0)	-	-	1 (0,6)

Observamos que se ha producido una reflexión y cambio en las concepciones de estos profesores sobre la actividad, los conceptos implicados y su aplicación, pero no en la totalidad de la muestra. Todavía queda un 35% de participantes convencidos de que sus propias estrategias (incorrectas) son mejores que las de sus compañeros. El tiempo dedicado a la actividad fue insuficiente para lograr que todos los participantes aceptasen la solución correcta, aunque posteriormente algunos de ellos, en los días sucesivos a la celebración del taller hicieron notar al formador de profesores que por fin habían comprendido cuál era la estrategia correcta y por qué.

No hemos registrado el porcentaje de profesores que encontraron la solución correcta en otros días; algunos nos informaron posteriormente de que necesitaron una o incluso dos semanas para llegar a ella. Pero este resultado también fue educativo para los participantes, pues refleja la situación que puede presentarse en una clase de estadística en secundaria o universidad, donde no todos los estudiantes logran adquirir los conocimientos pretendidos en el tiempo disponible para la enseñanza. En consecuencia, es también importante el tiempo necesario para la reflexión y actividades de formación del profesorado.

El cambio más notable fue pasar del 47,6% que pensaba que no hay estrategia, debido a la aleatoriedad de la situación, al 15,6% con lo que muchos participantes se confirman en que se necesita una estrategia para ganar, es decir, comprenden que el problema planteado tiene una solución abordable mediante el cálculo de probabilidad. Se consigue así otro objetivo que es mostrar el interés de la estadística como instrumento en la toma de decisiones en situaciones inciertas, en este caso, tomar una decisión sobre qué estrategia utilizar en el juego. Aunque es cierto que no podemos predecir el resultado de cada jugada (cada experimento aleatorio) si se puede prever el comportamiento de una serie larga de ensayos debido a la propiedad de convergencia.

Tabla 6.6.13. Resumen de la distribución de estrategias finales

Estrategia	España	Portugal	México	Total
Correcta	71 (72,5)	10 (37,0)	23 (56,1)	104 (62,7)
No percibe la independencia	8 (8,2)	4 (14,8)	7 (17,1)	21 (11,5)
No hay	19 (19,3)	13 (48,2)	11 (26,8)	43 (25,8)
Total	98	27	41	166

Resumidos los datos en la Tabla 6.6.13, todavía un 25,8% piensan que no hay estrategia y un 11,5% muestra una percepción incorrecta de la independencia, que también aparecen en otras investigaciones con futuros profesores (por ejemplo, en la tesis de Serrano, 1999) y en los Estudios 1 y 2. Hay diferencias también en el porcentaje de creencias incorrectas por país, siendo mayor en Portugal la creencia que no hay estrategia y en México la incorrecta percepción de la independencia. Puesto que la situación inicial era diferente en los tres países, la diferencia se conserva en la estrategia final, pero no llega a ser estadísticamente significativa (Chi-cuadrado=3,71; 2 g.l., p=0,15).

Tabla 6.6.14. Frecuencia (y porcentaje) de estrategia final según situación

Estrategia	En formación (n=79)	En ejercicio (n=87)
E1. Apostar al mismo color de la cara que se ve (correcta)	53 (67,1)	55 (63,2)
E2. Predecir el color contrario del que se muestra	5 (6,3)	5 (5,8)
E3. Considerar que no hay estrategia en la predicción (aleatoriedad)	12 (15,2)	12 (13,8)
E4. Elección de color rojo (color azul) en todos los ensayos	2 (2,5)	2 (2,3)
E5. Alternando colores	2 (2,5)	6 (6,9)
E6. Uso de los resultados anteriores para la predicción	4 (5,1)	5 (5,8)
E7. Cambiar las estrategias a lo largo de la secuencia de ensayos	1 (1,3)	1 (1,1)
E8. Propiedades no físicas de las tarjetas	-	1 (1,1)

En la Tabla 6.6.14 observamos también poca diferencia entre los profesores en formación y en ejercicio, salvo que algunos más en formación no perciben la independencia. Esto es razonable, puesto que esta diferencia aparecía al comienzo de la actividad y, aunque muchos profesores en ejercicio, corrigieron este razonamiento, todavía queda un grupo que lo mantiene. Los datos se resumen en la Tabla 6.6.15. Por razones de frecuencias bajas en algunas casillas, hemos preferido comparar únicamente respuestas correctas/incorrectas en el test Chi-cuadrado. Obtuvimos un valor no significativo (Chi-cuadrado= 0,27, 1 g.l., p= 0,60) que indica similitud de resultados finales en los dos grupos con porcentajes de estrategias correctas superiores al 60%. Concluimos la similitud de intuiciones incorrectas en los dos colectivos y la utilidad para ambos de la actividad formativa propuesta.

Tabla. 6.6.15. Resumen de la distribución de estrategias finales

Estrategia	En formación	En ejercicio
Correcta	53 (66,6)	55 (63,2)
No percibe la independencia	13 (15,9)	5 (1,1)
No hay	14 (17,5)	27 (35,7)
Total	79	87

Cambio de estrategia

Como observamos en la Tabla 6.6.16, la estrategia correcta (E1) vio incrementado su porcentaje en un 38%, lo que demuestra que las nuevas simulaciones del juego y el debate organizado por el formador de profesores ayudaron a una buena interpretación de los contenidos que aparecen implícitos en él. Todas las estrategias incorrectas disminuyeron en porcentaje desde la valoración inicial a la final, excepto la estrategia E5 (alternando colores) que aumentó su porcentaje, es decir no sólo los encuestados se reafirmaron en sus ideas sino que el número de ellos aumentó.

Al realizar una comparación por países (Tabla 6.6.17) se produce un incremento de respuestas correctas en todos ellos, destacando México en donde el porcentaje aumentó un 48,8%, debido al poco número de profesores que inicialmente dio la estrategia correcta. Como ocurría en el caso general, la única estrategia que aumentó en todos los países fue la E5 (alternancia de colores) destacando de nuevo a México con un 7,3%. Vemos en este resultado uno de los peligros de la simulación y es que una persona con intuiciones incorrectas pueda verlas reafirmadas si los datos confirman estas intuiciones.

Tabla 6.6.16. Porcentaje de estrategias iniciales y finales

	Estrategia inicial	Estrategia final	Incremento
E1. Apostar al mismo color de la cara que se ve (correcta)	24,7	62,7	38,0
E2. Predecir el color contrario del que se muestra	7,8	6,7	-1,1
E3. Considerar que no hay estrategia en la predicción (aleatoriedad)	47,6	15,7	-31,9
E4. Elección de color rojo (color azul) en todos los ensayos	3,0	2,4	-0,6
E5. Alternando colores	4,2	6,7	2,3
E6. Uso de los resultados anteriores para la predicción	7,2	4,8	-2,4
E7. Cambiar las estrategias a lo largo de la secuencia de ensayos	4,8	0,6	-4,2
E8. Propiedades no físicas de las tarjetas	0,7	0,6	-0,1

Tabla 6.6.17. Porcentaje de incremento de las estrategias por países

Estrategia	España	Portugal	México
E1. Apostar al mismo color de la cara que se ve (correcta)	38,8	18,5	48,8
E2. Predecir el color contrario del que se muestra	-3,1	3,7	0,0
E3. Considerar que no hay estrategia en la predicción (aleatoriedad)	-26,5	-25,9	-48,8
E4. Elección de color rojo (color azul) en todos los ensayos	-2,0	3,7	0,0
E5. Alternando colores	1,0	0	7,3
E6. Uso de los resultados anteriores para la predicción	-4,1	-3,7	2,4
E7. Cambiar las estrategias a lo largo de la secuencia de ensayos	-4,1	3,7	-9,8
E8. Propiedades no físicas de las tarjetas	0	0	0

Al estudiar el incremento por tipo de profesor (Tabla 6.6.18) observamos de nuevo la tendencia de las tablas anteriores. La situación (en ejercicio o formación) no influye tanto en los resultados como el país.

Tabla 6.6.18. Porcentaje de incremento de las estrategias por situación

Estrategia	En formación (n=79)	En ejercicio (n=87)
E1. Apostar al mismo color de la cara que se ve (correcta)	30 (38,0)	36 (41,4)
E2. Predecir el color contrario del que se muestra	-4 (-5,1)	1 (1,1)
E3. Considerar que no hay estrategia en la predicción (aleatoriedad)	-21(-26,6)	-34 (-39,1)
E4. Elección de color rojo (color azul) en todos los ensayos	-1 (-1,3)	-1 (-1,1)
E5. Alternando colores	2 (2,5)	3 (3,5)
E6. Uso de los resultados anteriores para la predicción	-3 (-3,8)	1 (1,1)
E7. Cambiar las estrategias a lo largo de la secuencia de ensayos	-4 (-5,1)	-6 (-6,9)
E8. Propiedades no físicas de las tarjetas	0 (0)	0 (0)

Pensamos que la razón es que la probabilidad o no se enseña, o se enseña de manera formal, Por ello los profesores no tienen experiencias con simulaciones y experimentos parecidos al mostrado en el taller y no tienen ocasión de confrontar sus concepciones erróneas en estos experimentos. Por tanto, no adquieren de forma espontánea un conocimiento profesional suficiente para la enseñanza propuesta en las nuevas directrices curriculares.

6.6.3. JUSTIFICACIÓN DE LAS ESTRATEGIAS

Se proporcionó a los participantes un tiempo para que pensasen y pusiesen por escrito su demostración o lo que ellos considerasen que podría servir para demostrar la estrategia. La finalidad de esta fase es que los profesores experimenten la fase de validación, descrita por Brousseau (1986, 1997). Puesto que la matemática se construye siguiendo unas reglas establecidas, las soluciones a los problemas matemáticos, han de estar de acuerdo con los convenios y reglas matemáticos. La fase de validación permite una reflexión profunda sobre dichos convenios y reglas, así como ejercitar diversos procesos matemáticos, tales como la generalización o representación. Sin esta fase el participante puede encontrar la solución al problema, pero no comprender el porqué de dicha solución. Por tanto, no explicitarán los objetos matemáticos que han sido utilizados en el problema.

El haber obtenido un conjunto variado de estrategias finales, cada una de ellas defendida por sus partidarios, sirvió para iniciar un debate de las soluciones por parte de los alumnos. La finalidad de la misma no fue tanto el aprendizaje de la probabilidad (que también se consigue), sino concienciar a los profesores de la existencia de concepciones incorrectas sobre la probabilidad, que son visibles en algunos participantes en el taller y que seguramente encontrarán más tarde en sus alumnos. El análisis de los argumentos a favor de cada estrategia reveló estas concepciones y sirvió para analizar los contenidos matemáticos implícitos en la tarea. Se pidió a los participantes que aportaran sus justificaciones por escrito y finalmente fueron analizadas. A continuación se describen las variantes encontradas.

Justificaciones correctas y parcialmente correctas

Son todas las demostraciones correctas o parcialmente correctas (desde un punto de vista matemático) de que la estrategia de jugar al mismo color que se extrae, es la óptima. Tenemos algunas variantes, con diverso grado de formalización matemática.

J1. *Estrategia empírica.* Son los profesores que usan las frecuencias relativas obtenidas durante el desarrollo del juego, para analizar cuál de las estrategias ofrece más posibilidad de ganar. El profesor no es capaz de hallar un argumento probabilístico correcto, pero analiza la frecuencia de resultados al repetir el experimento, es decir, trata de dar una prueba empírica. Aunque no es una demostración en el sentido matemático del término, la hemos considerado parcialmente correcta para este taller, pues se aproxima a lo que sería un contraste de hipótesis estadística. Es decir, en un contraste estadístico, si tuviésemos una hipótesis, para ponerla a prueba, organizaríamos un experimento para recoger datos que confirmen o rechacen la hipótesis. Compararíamos los datos obtenidos con el patrón previsto en la hipótesis y a partir de los datos recogidos, tomaríamos una decisión sobre la validez de la hipótesis.

Tabla 6.6.19. Objetos matemáticos implícitos en la justificación J1

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, según el color mostrado
Lenguajes	- Color	- Suceso posible; que puede ser rojo o azul
	- Color que se mostraba	- Resultado del experimento
	- Color que se predecía	- Suceso más probable, dada el resultado de la cara mostrada
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara; acertar el color
	- Sucesos	- Color de las caras azul y rojo
	- Convergencia	- Resultado al que tiende la simulación
	- Unión de sucesos	- Conjunto formado por todos los elementos
	- Estimación frecuencial de la probabilidad	- Probabilidad obtenida de los datos de la simulación
Procedimientos	- Estimación frecuencial de la probabilidad	- Estima las probabilidades de rojo y azul en función de lo que ha salido
	- Comparación de probabilidades	- Compara la probabilidad de azul y rojo en función de lo que ha salido
Propiedades	- Regla de Laplace	- La probabilidad de salir rojo es el cociente entre casos favorables entre posibles
	- Regla de la suma	- Suma de las probabilidades de los dos sucesos
	- Convergencia	- A la larga el suceso más probable aparece

Algunos participantes usaron esta estrategia de justificación, aunque de manera muy informal, sin hablar de hipótesis o contraste de hipótesis, simplemente comparando los datos con lo previsto. No obstante, esta prueba podría darle una buena aproximación si hace una simulación con un número muy grande de experimentos, debido a la convergencia entre frecuencia relativa y probabilidad dada por el teorema de Bernoulli. Pero matemáticamente, la demostración no tiene un carácter válido como lo tendría una demostración deductiva y además no proporciona a los estudiantes una explicación de por qué la estrategia es válida. Un ejemplo lo presentamos a continuación y se analiza en la Tabla 6.6.19.

> *"He apostado al mismo color que sale en la cara mostrada y de forma experimental veo que tengo más aciertos que fracasos, por lo que será correcta, aunque no encuentro una justificación estadística".*

J2. *Justificación empírica, completada con el contraste formal de hipótesis.* Este razonamiento es similar al anterior, pero el estudiante ahora ha hecho un gran número de simulaciones y ha aplicado la prueba de Chi-cuadrado, por lo que la hemos considerado correcta.

Tabla 6.6.20. Objetos matemáticos implícitos en la justificación J2

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, según el color mostrado
	- Decidir si los datos se ajustan a una hipótesis	- Comprobar si es cierto que la estrategia dada por buena produce 2/3 de aciertos
Lenguajes	- Verbal	- Para designar sucesos; "misma cara", etc.
	- Numérico	- Expresión de frecuencias, operaciones y resultados
	- Simbólico	- Símbolos de operaciones; igualdad; Chi-cuadrado
	- Tabla	- Disposición tabular de cálculos
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara; acertar el color
	- Sucesos	- Color de las caras azul y rojo
	- Probabilidad simple	- Casos favorables entre posibles
	- Probabilidad condicionada	- Proporción de casos favorables, con una condición dada
	- Unión de sucesos	- Conjunto formado por todos los elementos
	- Estimación frecuencial de la probabilidad	- Probabilidad obtenida de los datos de la simulación
	- Variable aleatoria, y distribución de probabilidad	- Número de aciertos: número esperado si la estrategia da 2/3 de aciertos
	- Variable y distribución estadística	- Número de aciertos en el juego
	- Hipótesis nula y alternativa	- Los datos se ajustan o no a la distribución esperada
	- Estadístico Chi cuadrado	- Medida de la discrepancia entre hipótesis y datos
	- Estimación frecuencial de la probabilidad	- Calcula las probabilidades de rojo y azul en función de lo que ha salido
Procedimientos	- Comparación de probabilidades	- Compara la probabilidad de azul y rojo en función de lo que ha salido
	- Test Chi-cuadrado	- Compara los resultados indicando en qué medida hay diferencias entre la variable teórica y la observada
	- Cálculo de probabilidades	- La probabilidad de salir rojo es el cociente entre casos favorables entre posibles
Propiedades	- Distribución Chi-cuadrado	- Al ser un resultado pequeño indica que no hay diferencias entre la variable teórica y la observada

En el ejemplo que reproducimos, el sujeto está utilizando un modelo estadístico de distribución esperada bajo la hipótesis nula (dos tercios de aciertos) y compara los datos obtenidos en el experimento con los esperados. Calcula el estadístico, que da una medida de la discrepancia entre datos observados y esperados. Como el valor del estadístico es pequeño, no tiene motivo para rechazar la hipótesis de ajuste del modelo a los datos. Por ello, mantiene su hipótesis. Aunque no bien explicitado, este razonamiento es un ejemplo de un test de hipótesis, correctamente completado y aplicado a la situación experimental y constituiría una "prueba" estadística que debemos aceptar como buena. Mostramos un ejemplo que se analiza en la Tabla 6.6.20.

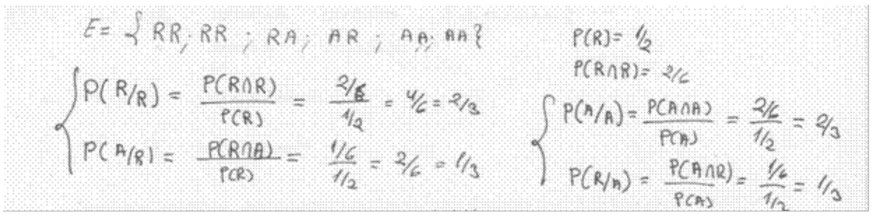
Análisis estadístico de resultados. Empíricamente lo he probado en 204 repeticiones del experimento:

	Misma cara	*Distinta cara*
Real	*138*	*66*
Esperada	*132*	*68*

$$\chi^2 = \frac{(138-132)^2}{132} + \frac{(66-68)^2}{68} = 0,3325$$

Esto nos lo confirma totalmente".

J3. *Deductivas correctas* utilizando adecuadamente los símbolos y fórmulas del cálculo de probabilidades. El participante elige una notación adecuada para representar los experimentos en el juego y para plantear el cálculo de probabilidades necesario para resolver el problema. Aplica correctamente las fórmulas y da una demostración lógicamente correcta, tanto por los conceptos, propiedades y relaciones usadas, como por la cadena de implicaciones realizadas. Un ejemplo se muestra a continuación y se analiza en la Tabla 6.6.21.


$E = \{RR, RR; RA, AR; AA, AA\}$

$P(R) = 1/2$

$P(R \cap R) = 2/6$

$P(R/R) = \frac{P(R \cap R)}{P(R)} = \frac{2/6}{1/2} = 4/6 = 2/3$

$P(A/R) = \frac{P(R \cap A)}{P(R)} = \frac{1/6}{1/2} = 2/6 = 1/3$

$P(A/A) = \frac{P(A \cap A)}{P(A)} = \frac{2/6}{1/2} = 2/3$

$P(R/A) = \frac{P(A \cap R)}{P(A)} = \frac{1/6}{1/2} = 1/3$

En este ejemplo, el profesor comienza enumerando correctamente el espacio muestral producto y haciendo referencia a la doble posibilidad de la ficha con las dos caras iguales; es por tanto consciente de la importancia del orden en este experimento. Calcula a continuación las probabilidades condicionales de obtener cara roja o azul si la mostrada es roja o azul. Ha realizado una enumeración correcta de todos los casos posibles, calculando las probabilidades de cada uno mediante la fórmula de la probabilidad condicional, correctamente aplicada al caso de dependencia. Comparando las diferentes posibilidades deduce la mayor probabilidad de que la cara oculta tenga el mismo color que la mostrada.

Tabla 6.6.21. Objetos matemáticos implícitos en la justificación J3

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, según el color mostrado
Lenguajes	- Color	- Suceso posible; que puede ser rojo o azul
	- Color que se mostraba	- Resultado del experimento “cara mostrada”
	- Color que se predecía	- Suceso esperado en el experimento “cara oculta” - Suceso más probable, dada el resultado de la cara mostrada
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara; acertar el color
	- Sucesos	- Color de las caras azul y rojo
	- Estimación frecuencial de la probabilidad	- Probabilidad obtenida de los datos de la simulación
	- Regla de la suma	- Suma de las probabilidades de los dos sucesos
	- Regla del producto	- Calculo de la probabilidad condicionada a partir del cociente de la probabilidad conjunta (intersección) y la simple del suceso condicionante
	- Axiomas de probabilidad	- Condiciones que deben verificarse para que una función definida sobre un conjunto de sucesos determine sus probabilidades
Procedimientos	- Estimación frecuencial de la probabilidad	- Estima las probabilidades de rojo y azul en función de lo que ha salido
	- Comparación de probabilidades	- Compara la probabilidad de azul y rojo en función de lo que ha salido anteriormente
Propiedades	- Regla de Laplace	- La probabilidad de salir rojo es el cociente entre casos favorables entre posibles
	- Axioma de la unión	- La suma de las probabilidades es 1
	- Regla del producto	- La probabilidad condicionada en función del cociente entre la conjunta y la simple

J4. *Deductiva informal.* El estudiante utiliza un esquema gráfico, por ejemplo, un diagrama de árbol y representa esquemáticamente su razonamiento, sin usar fórmulas ni notación simbólica para la probabilidad condicional o el espacio producto. El esquema permite ver la corrección de su razonamiento, pero no está suficientemente verbalizado. Un ejemplo se presenta a continuación y se analiza en la Tabla 6.6.22.

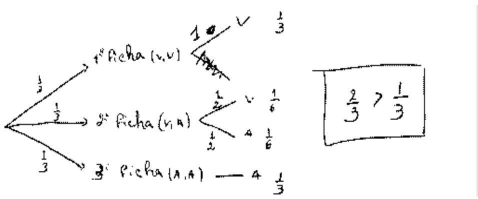


Tabla 6.6.22. Objetos matemáticos implícitos en la justificación J4

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, según el color mostrado
Lenguajes	- Color	- Suceso posible; que puede ser rojo o azul
	- Color que se mostraba	- Resultado del experimento "cara mostrada"
	- Color que se predecía	- Suceso esperado en el experimento "cara oculta" - Suceso más probable, dada el resultado de la cara mostrada
	- Gráfico	- Diagrama de árbol para diferenciar los diferentes sucesos
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara; acertar el color
	- Sucesos	- Color de las caras azul y rojo
	- Estimación frecuencial de la probabilidad	- Probabilidad obtenida de los datos de la simulación
	- Regla de la suma	- Suma de las probabilidades de los dos sucesos
Procedimientos	- Estimación frecuencia de la probabilidad	- Estima las probabilidades de rojo y azul en función de lo que ha salido
	- Comparación de probabilidades	- Compara la probabilidad de azul y rojo en función de lo que ha salido anteriormente
Propiedades	- Regla de Laplace	- La probabilidad de salir rojo es el cociente entre casos favorables entre posibles
	- Axioma de la unión	- La suma de las probabilidades es 1

Observamos en el ejemplo, la correcta partición del espacio muestral en el experimento compuesto en las tres fichas posibles (primer nivel del árbol) y la correcta asignación de la probabilidad a cada uno (equiprobabilidad, en ausencia de otra información). Para cada posible fichas, representa un segundo experimento (color de la cara oculta) dividiendo las ramas. Tiene en cuenta la importancia del orden, pues propone dos ramas para la ficha *VA* y una sola en los otros dos casos. En cada subrama del árbol asigna correctamente la probabilidad. Calcula correctamente la probabilidad producto multiplicando las probabilidades de las ramas. Parece aplicar el axioma de la unión, pues indica la mayor probabilidad de que la cara oculta sea del mismo color que la mostrada. Tanto el razonamiento como el cálculo de las probabilidades son correctos, pero no lo explicita suficientemente ni verbaliza con una adecuada notación matemática.

Justificaciones incorrectas

Se han encontrado varios tipos de justificaciones incorrectas, causadas por confusión de algunos conceptos o no reconocimiento de los mismos en el juego. Aunque la mayoría de los alumnos llega finalmente a la solución correcta, esto no quiere decir que sean capaces de argumentarla usando una argumentación matemática. Respuestas como las siguientes permiten reflexionar sobre la argumentación y sus tipos, así como sobre el tipo de evidencia que proporcionan la probabilidad teórica y empírica:

J5. *Confundir experimento simple y compuesto,* error que apareció en el Estudio 1. En el siguiente ejemplo, el razonamiento seguido consiste en considerar el conjunto de todas las caras en las tres tarjetas como si se tratase de un único espacio muestral $\{A, A, A, R, R, R\}$. Se considera el experimento como simple y no como compuesto, aplicando la regla de Laplace. Hay un conflicto en la comprensión de la naturaleza de los objetos que se manipulan en el experimento, llevando a un error en la enumeración del espacio muestral. Además, el estudiante considera el experimento como simple (extraer al azar una de las 6 caras, habiendo sacado una previamente) y no como un experimento compuesto (mostrar la cara oculta, habiendo mostrado ya una de las dos).

> *"Tomar el color contrario, ya que la probabilidad de que vuelva a ocurrir de nuevo el mismo color es menor de que salga el contrario. Es decir, si sale el rojo la probabilidad de que vuelva a salir el mismo color es 2/5 y 3/5 que salga el azul y al revés".*

Tabla 6.6.23. Objetos matemáticos implícitos en la justificación incorrecta a.

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, dependiendo del color mostrado
Lenguajes	- Probabilidad	- Casos favorables entre los posibles
	- Color contrario	- Salir distinto color del que salió en la cara inicial
	- Mismo color	- Salir el mismo color del que salió en la cara inicial
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara; acertar el color; *conflicto* al no comprender la naturaleza compuesta del experimento, confundiéndolo con uno simple
	- Sucesos	- Color de las caras azul y rojo; *conflicto* al no comprender que las caras de las tarjetas no pueden separarse
	- Espacio muestral	- *Conflicto* al enumerar el espacio muestral producto
	- Probabilidad simple	- Casos favorables entre posibles
	- Probabilidad condicionada	- Proporción de casos favorables, con una condición dada
	- Unión de sucesos	- Conjunto formado por todos los elementos
Procedimientos	- Probabilidad simple	- Cálculo de probabilidad simple
	- Comparación de probabilidades	- Compara la probabilidad de azul y rojo en función de lo que ha salido
Propiedades	- Regla de Laplace	- La probabilidad es el cociente entre casos favorables y posibles

J6. *No percibe la dependencia de los experimentos,* problema que s mostró en los estudios 1 y 2. Algunos estudiantes no perciben la dependencia de los dos experimentos. Consideran, por tanto los experimentos "tomar una tarjeta" y "enseñar una cara" como independientes, pues suponen que el color de la cara oculta no depende de la tarjeta tomada (o lo que es lo mismo, de la cara mostrada). El estudiante no tiene una buena percepción de la independencia o dependencia de los sucesos en este caso. Ello les lleva a considerar que no hay estrategia posible en la predicción. Un ejemplo se reproduce a continuación y se analizan en la Tabla 6.6.24.

"Pienso que no existe ninguna estrategia válida porque la probabilidad de la cara oculta es la misma para cada color y cada intento es independiente de los demás y supuestamente aleatorio".

Tabla 6.6.24. Objetos matemáticos implícitos en la justificación incorrecta b.

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, según el color mostrado
Lenguajes	- Probabilidad	- Casos favorables entre los posibles
	- Aleatorio	- Los resultados dependen del azar
	- Cara oculta	- Cara que se quiere predecir
	- Independiente	- El experimento no está influido por los casos anteriores
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara; acertar el color
	- Experimento compuesto	- Composición de los anteriores; conflicto al suponerlos independientes
	- Sucesos	- Color de las caras azul y rojo; sucesos en el experimento compuesto
	- Probabilidad simple	- Casos favorables entre posibles
	- Unión de sucesos	- Conjunto formado por todos los elementos
	- Independencia	- Incorrectamente percibidas
Procedimientos	- Cálculo de Probabilidad simple	- La probabilidad es el cociente entre casos favorables o posibles
	- Comparación de probabilidades	- Compara la probabilidad de azul y rojo en función de lo que ha salido
Propiedades	- Regla de Laplace	- La probabilidad es el cociente entre casos favorables y posibles

J7. *Falacia del jugador* (Tversky y Kahneman, 1982 a). En este caso la persona, no sólo no percibe la dependencia de los experimentos, sino que además, supone que al obtener un color, el contrario debe tener mayor posibilidad. Implícitamente, es posible que el estudiante haya seguido el razonamiento del ejemplo dado anteriormente para los que asumen dependencia en los datos. También apareció en el Estudio 2.

"Pienso que hay más probabilidad de alternancia de colores debido a que hay el mismo número de caras de los dos colores: 3 posibilidades de que salga rojo y 3 de que salga azul".

Tabla 6.6.25. Objetos matemáticos implícitos en la justificación incorrecta c.

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, según el color mostrado
Lenguajes	- Probabilidad	- Casos favorables entre los posibles
	- Alternancia de colores	- Los resultados de la simulación se alternan

	- 3 posibilidades de que salga rojo y 3 de que salga azul	- Conjunto de elementos que forman los sucesos simples
Conceptos	- Experimento aleatorio	- Sacar una tarjeta, mostrar una cara; acertar el color
	- Azar	- Los resultados dependen de la aleatoriedad del proceso
	- Sucesos	- Color de las caras azul y rojo
	- Probabilidad simple	- Casos favorables entre posibles
	- Unión de sucesos	- Conjunto formado por todos los elementos
Procedimientos	- Cálculo de probabilidad simple	- Aplicando la regla de Laplace
	- Comparación de probabilidades	- Compara la probabilidad de azul y rojo en función de lo que ha salido
Propiedades	- Regla de Laplace	- La probabilidad es el cociente entre casos favorables y posibles

J8. *Aplica una fórmula incorrecta*, error que apareció en el Estudio 2. Como vemos en el siguiente ejemplo el alumno conoce el comportamiento de la regla de Laplace y que tiene que dividir entre los casos favorables respecto al total pero es incapaz de delimitar los sucesos y de identificar las probabilidades. En el ejemplo que se muestra, además de confundir frecuencia absoluta con probabilidades, el alumno para definir los casos favorables de la probabilidad que salga azul en la segunda cara, suma la probabilidad de que salga azul en la primera con la segunda. En los casos posibles suma la probabilidad de que salga azul en la primera, con que salga azul en la segunda y roja en la segunda.

“ $P(A_1) = \dfrac{P(A_1) + P(A)}{P(A_1) + P(A) + P(B)}$ ”.

Tabla 6.6.26. Objetos matemáticos implícitos en la justificación J8

Tipos	Objetos matemáticos en la situación	Significado en la situación
Situaciones-problemas	- Encontrar la mejor estrategia	- Calcular la probabilidad condicional de que la ficha oculta sea roja o azul, dependiendo del color mostrado
Lenguajes	- Simbólico	- Para expresar sucesos y probabilidades
	- Fracción, suma, igualdad	- Para expresar operaciones
Conceptos	- Probabilidad simple	- Casos favorables entre posibles
	- Unión de sucesos	- Conjunto formado por todos los elementos
Procedimientos	- Cálculo de probabilidad	- Conflicto: Fórmula incorrecta
Propiedades	- Regla de Laplace	- La probabilidad es el cociente entre casos favorables y posibles

En las tablas 6.6.27 y 6.6.28 presentamos las demostraciones realizadas por los participantes de sus estrategias. Observamos que únicamente el 36% fue capaz de dar una demostración deductiva correcta y parte de ellos la llevaron a cabo únicamente en forma verbal o utilizando un diagrama en árbol, sin utilizar una notación matemática completa. Un 18% de los profesores compararon las diferentes estrategias propuestas con los datos empíricos obtenidos en el experimento y, ya que la estrategia correcta proporcionaba una clara ventaja en el juego, consideraron que esta comparación empírica era una "prueba" suficiente. Consideramos esta estrategia sólo como parcialmente correcta, pues el carácter aleatorio del experimento requeriría del uso de un contraste de hipótesis para comparar los datos observados con los resultados esperados teóricamente, para poder considerar la prueba como aceptable. Tres de los profesores realizaron este tipo de "prueba", utilizando el contraste Chi-cuadrado para comparar sus datos con el valor esperado (2/3 de aciertos) en la estrategia correcta. En este caso consideramos la prueba correcta al involucrar un contraste de hipótesis.

Tabla 6.6.27. Frecuencia de tipos de justificación

	Justificación	Frecuencia	%
Correcta	Deductiva, utilizando simbolización matemática	41	24,7
	Deductiva, mediante diagrama o verbalmente	26	15,7
	Empírica con contraste de hipótesis	3	1,8
P. correcta	Empírica	30	18,1
Incorrecta	Confunde experimento simple y compuesto	12	7,2
	Demostración deductiva incorrecta	4	2,4
	Demostración incorrecta al no percibir la dependencia	21	12,6
	No completa o no demuestra	29	17,5
	Total	166	100

Tabla 6.6.28. Resumen de tipos de justificación

Demostración	Frecuencia (%)
Correcta	70 (42,2)
Parcialmente correcta	30 (18,1)
Incorrecta	37 (22,2)
No completa	39 (17,5)

Otro 22,2% de los profesores participantes presentan una demostración matemáticamente incorrecta por diversos motivos, sobre todo, por no percibir la dependencia de los experimentos aleatorios que subyacen en este juego y seguir en la fase final apostando por una estrategia incorrecta, que es la que tratan de demostrar. Hacemos también notar que otro 17,5% de los profesores no finaliza su demostración o ni siquiera la inicia; en este grupo se incluyen aquellos que en la fase final todavía piensan que no hay una estrategia posible, por lo que no intentan la demostración.

Observamos algunas diferencias por país y por situación profesional en las justificaciones (Tablas 6.6.29 y 6.6.30). Respecto a país (Chi= 17,32; 6 g.l.; *p*= 0,0082) hubo una mayor proporción de demostraciones empíricas e incorrectas comparativamente en Portugal y las diferencias fueron estadísticamente significativas.

Tabla 6.6.29. Justificaciones por país

		España		Portugal		México	
		Fr.	%	Fr.	%	Fr.	%
Correcta	Deductiva, utilizando simbolización matemática	29	29,7	2	7,1	10	25,0
	Deductiva, mediante diagrama o verbalmente	16	16,3	1	3,6	9	22,5
	Experimental con contraste de hipótesis	2	2,0	1	3,6	0	0,0
P. correcta	Experimental empírica	15	15,3	9	32,1	6	15,0
Incorrecta	Confundir experimento simple por compuesto	3	3,1	5	17,9	4	10,0
	Demostración deductiva incorrecta	2	2,0	0	0,0	2	5,0
	Demostración incorrecta al no percibir la dependencia	12	12,2	7	25	2	5,0
	No completa o no demuestra	19	19,4	3	10,7	7	17,5
	Total	98	100	28	100,0	40	100,0

Tabla 6.6.30. Resumen de justificaciones por países

	España		Portugal		México	
Correcta	47	47,9	4	14,3	19	47,5
P. correcta	15	15,3	9	32,1	6	15
Incorrecta	17	17,3	12	42,9	8	20
No completa	19	19,4	3	10,7	7	17,5
Total	98	100	28	100	40	100

En cuanto a la situación profesional (Chi=10,51; 3 g.l., *p*=0,0147), paradójicamente, hubo mayor proporción de estrategias correctas y parcialmente correctas en los profesores en formación, quienes dejan la respuesta en blanco en menos ocasiones, pero también tienen más demostraciones incorrectas. Asumimos que el contrato didáctico para estos profesores (que estaban cursando una asignatura optativa cuando se les pasó el cuestionario) les lleva a completar la solución, incluso cuando no estén seguros de ella, mientras los profesores en ejercicio prefirieron en estos casos no responder. Al encontrarse en el último curso de la Licenciatura de Matemáticas o estadística los profesores en formación tienen más reciente el estudio de la probabilidad, lo que ha influido en la mejor respuesta.

En resumen, los resultados mostrados señalan la necesidad de mejorar las intuiciones de probabilidad en los profesores, incluso de los profesores en ejercicio, quienes, en una proporción considerable mostraron intuiciones incorrectas al comienzo de la actividad y no fueron capaces de dar una demostración completa de la estrategia, una vez identificada al final del juego. Pensamos que ello se debe a que la preparación matemática de los profesores en su formación inicial se centra en el conocimiento común del contenido, y sería necesario complementarlo con un conocimiento especializado del contenido, que incluye información sobre los sesgos en el razonamiento probabilístico.

Tabla 6.6.31. Justificaciones por situación

		En ejercicio		En formación	
		Fr.	%	Fr.	%
Correcta	Deductiva, utilizando simbolización matemática	21	24,2	20	25,3
	Deductiva, mediante diagrama o verbalmente	11	12,7	15	19,0
	Experimental con contraste de hipótesis	1	1,1	2	2,5
P. correcta	Experimental empírica	15	17,2	15	19,0
Incorrecta	Confundir experimento simple por compuesto	8	9,2	4	5,1
	Demostración deductiva incorrecta	1	1,1	3	3,8
	Demostración incorrecta al no percibir la dependencia	7	8,0	14	17,7
	No completa o no demuestra	23	26,5	6	7,6
	Total	87	100,0	79	100,0

Por otro lado, al igual que en la experiencia de Batanero. Godino y Roa (2004) el taller resultó efectivo para cambiar las creencias iniciales de una proporción importante de profesores, tanto en formación como en ejercicio y en los tres países participantes. Sin embargo el éxito no fue total pues aún una parte de los participantes persistió en sus creencias erróneas al finalizar el taller, por lo que hubiese sido necesario un mayor tiempo para la actividad para estos profesores o bien complementarla con otras.

Tabla 6.6.32. Resumen de justificaciones por situación

	Ejercicio		Formación	
	Fr.	%	Fr.	%
Correcta	33	38	37	46,8
P. correcta	15	17,2	15	19
Incorrecta	16	18,3	21	26,6
No completa	23	26,5	6	7,6
Total	87	100	79	100

6.7. CONOCIMIENTOS DIDÁCTICOS SOBRE PROBABILIDAD CONDICIONAL

El conocimiento matemático en sí mismo es insuficiente para que el profesor esté preparado para enseñar probabilidad. Por ello es también importante atender a su formación didáctica. En la segunda parte del taller se llevó a cabo el análisis didáctico del juego, siguiendo la metodología descrita en el apartado 6.4.2. Los informes escritos sobre este análisis permitieron evaluar algunos conocimientos didácticos de los profesores, siguiendo el modelo de Godino (2009), quien propone diversas actividades de análisis didáctico con fines de evaluación de los diferentes componentes del conocimiento del profesor. La metodología que propone este autor consistiría en elegir una tarea matemática (un proyecto, secuencia de actividades) cuya solución ponga en juego los principales aspectos del contenido, o de las competencias a desarrollar. A continuación se darían consignas que cubran las distintas (o principales) facetas y niveles de análisis didáctico. Nosotros hemos elegido como tarea el taller ya descrito y que fue vivido por los profesores. En lo que sigue analizamos los resultados de esta evaluación.

6.7.1. CONOCIMIENTO ESPECIALIZADO DEL CONTENIDO

En primer lugar se evalúan aspectos relevantes del conocimiento especializado del contenido. Según Hill, Ball, y Schilling (2008) éste es un conocimiento que va más allá del común e implica que los profesores sean capaces de enseñar dicho contenido. La distinción entre el conocimiento común del contenido (CCK) y el especializado (SCK) consiste en que, mientras el primero refiere al conocimiento puesto en juego para resolver problemas matemáticos, para lo cual un matemático, o incluso un sujeto adulto con suficiente conocimiento, está capacitado; el segundo incluye las competencias necesarias para desarrollar los diferentes aspectos de un contenido específico.

Para evaluarlo, hemos pedido a los profesores que identifiquen los objetos matemáticos usados implícita o explícitamente al resolver la tarea propuesta. Esta es una de las preguntas que Godino (2009) considera pertinente para evaluar este tipo de conocimiento. El autor indica que responder a la misma supone pensar de manera sistemática en los diferentes procedimientos posibles de resolución, modalidades de expresión, conceptos y propiedades que se ponen en juego con distintos grados de formalidad en su formulación, así como pensar en las maneras de argumentar o justificar los procedimientos y propiedades utilizados.

Recogidos los protocolos se realizó un análisis cualitativo de las respuestas, mediante el análisis de contenido ((Ghiglione y Matalón, 1991), anotando, para cada tipo de objeto matemático (situaciones, lenguaje, conceptos, propiedades, argumentos, procedimientos) dos variables: por un lado el número de objetos correctamente identificados y por otro lado el número de objetos incorrectamente identificados en la tarea por cada profesor. Con un proceso inductivo y cíclico, propio de este tipo de análisis, se revisaron los códigos hasta llegar a un acuerdo, asegurando así la fiabilidad del análisis. Seguidamente se procedió a realizar un análisis estadístico descriptivo de estas variables. A continuación se analizan los objetos identificados.

Situaciones problemas

La mayoría de los participantes identificó la búsqueda de la estrategia en el juego como un problema matemático, posiblemente porque al comenzar el taller se les había indicado que se trataba de una situación de enseñanza que podrían utilizar con estudiantes de secundaria o de universidad. Otro problema citado fue el cálculo de las probabilidades de obtener la cara oculta de las tarjetas o el cálculo de probabilidades de obtener una cara oculta de diferente color a la mostrada, pues, por las mismas consideraciones, este cálculo constituyó un verdadero problema para los participantes y, como hemos visto, muchos no lo consiguieron resolver. Aunque este problema es equivalente al primero, algunos los diferenciaron entre sí, pues el primero es un problema extra-matemático mientras que éste es un problema matemático.

En otros casos, simplemente se indica que hay que *"resolver un problema de probabilidad",* sin indicar cuál es la probabilidad que se busca o bien que se debe "*comparar probabilidades"*, refiriéndose al hecho de que para deducir la estrategia correcta, hay que comparar la probabilidad de acertar bajo diferentes estrategias. Por último, un problema correcto diferente a los anteriores y citado con frecuencia fue el de "*enumerar el espacio muestral"*, que constituyó quizás el principal problema para muchos participantes, pues fue difícil decidir si el orden de las tarjetas habría de tenerse en cuenta. Es decir, algunos profesores mencionaron la elaboración del espacio muestral como un problema, dado que el espacio muestral correcto no era directamente perceptible para la mayoría.

Por otro lado, se ha encontrado también identificación incorrecta de problemas, en todos los casos, porque no se trata de un problema, sino de otro tipo de objeto matemático. Se identificaron incorrectamente los siguientes problemas: "*el color de la cara mostrada*", posiblemente el profesor quería expresar que se trataba de adivinar dicho color, pero aun así, no se trata de un problema o al menos no tiene solución matemática, puesto que el color de la cara es aleatorio y el resultado de un experimento aleatorio no se puede predecir. En otros casos se identificaron incorrectamente como problemas algunos razonamientos como, por ejemplo, "*razonar sobre el experimento*" o incluso algunas de las estrategias del juego, como "*decir que no hay estrategia*", que no sería realmente un problema, sino la solución al problema planteado. También aparecieron confusiones entre problema y concepto, como por ejemplo, el profesor que citó como problema el "*espacio muestral*" e incluso con particularizaciones de conceptos, como serían "*tarjeta*" y "*cara*", que no son problemas sino casos particulares de sucesos en el experimento llevado a cabo.

Lenguajes

Dentro del lenguaje correctamente identificado, se menciona con frecuencia la tabla dada para recoger los datos como un tipo de representación que facilita la solución del problema, pues permite ver el comportamiento del proceso aleatorio y también comparar el número de aciertos en diferentes estrategias. Otro tipo de lenguaje correcto mencionado por muchos profesores es el simbólico. Algunos especifican los símbolos usados para referirse a los sucesos (R, A) o bien las palabras equivalentes (rojo, azul) para hacer referencia a los sucesos del experimento. También se citan con menor frecuencia los símbolos de unión, intersección y complementarios y diagrama en árbol.

De una forma general, se hace referencia al lenguaje matemático (para referirse a los símbolos, y términos matemáticos), al lenguaje gráfico (para hacer referencia a los diagramas en árbol), estadístico (para indicar los datos recolectados en la experiencia, y los términos estadísticos, tales como la expresión de porcentajes o frecuencias). Con menor frecuencia se cita como lenguaje los colores de las tarjetas, los dibujos de las mismas o palabras específicas usadas para referirse a elementos fenomenológicos o matemáticos.

Respecto al lenguaje incorrectamente identificado, encontramos en algunos casos referencia a conceptos como *tabla de frecuencias* (en el experimento no se elaboró ninguna), *azar, probabilidad contraria, suceso, experiencia, muestra* (que son conceptos y no lenguaje). Es posible que en este caso los profesores hicieran referencia a la palabra dada y no al concepto; nosotros lo hemos considerado como incorrecto por haber sido expresado de una manera confusa. Por este mismo motivo se ha considerado incorrecto otros ejemplos de lenguaje que hacen referencia a elementos fenomenológicos tales como *intuición, tarjeta, cara oculta, cuadritos de la experiencia, celdas, cara roja, cara azul, tarjeta monocolor, tarjeta bicolor y fracaso*.

Conceptos

Entre los conceptos correctamente identificados se nombran los de *azar, casos favorables, casos posibles, composición de experimentos, combinatoria, dependencia, espacio muestral, experimento, frecuencia, incertidumbre, independencia, muestreo, probabilidad simple, probabilidad condicionada, razón, porcentaje, suceso imposible, proporción, fracción, suceso, variable aleatoria, probabilidad compuesta, estimación frecuencial de la probabilidad.*

Observamos que la mayoría son conceptos ligados al experimento y a la probabilidad, aunque se citan unos pocos (como proporción) que no son tan específicos del juego. También observamos que se citan la mayor parte de las ideas estocásticas sugeridas por Heitele (1975), quien indica que estas aparecen en cualquier situación aleatoria. Puesto que la identificación de las ideas estocásticas fundamentales fue uno de los objetivos del taller, consideramos conseguido este objetivo.

Con menor frecuencia se citaron incorrectamente supuestos conceptos. En algunos casos se considera como concepto algunos ejemplos donde se puede aplicar el mismo, mostrando una confusión entre los planos intensivo /extensivo de un objeto matemático (Godino, 2002). Por ejemplo se nombra *acierto* y *fallar la predicción*, que serían sucesos en el experimento, es decir un ejemplo del concepto “suceso”, pero no el concepto en sí mismo. Hay también confusión entre concepto y procedimiento, por ejemplo, en los casos de citar como conceptos *estrategia, descartar, predicción, sumar probabilidades, escoger fichas.* Tampoco distinguen propiedad y concepto en unos pocos casos, como al citar como concepto *muy probable, bastante probable*, que son propiedades.

Procedimientos

Entre los procedimientos correctamente identificados se hace referencia a procedimientos efectivamente realizados por el profesor, como la *construcción del diagrama en árbol, la enumeración del espacio muestral, o el cálculo de la probabilidad condicionada.* También se hace referencia a los procedimientos generales utilizados en la sesión como *interpretar los datos del problema, utilizar estrategias, completar algoritmos, realizar operaciones, realizar una simulación, y llevar a cabo una experimentación.*

En unos pocos casos se confunden procedimientos y lenguaje: por ejemplo al indicar "*tablas", en* lugar de citar "*realización de la tabla.* Unos pocos profesores confunden el procedimiento con el problema, como cuando se indica "*encontrar la mejor estrategia*", que fue el problema planteado. Finalmente en otros casos la confusión se presenta entre procedimiento y argumento, como cuando se indica "*deductivo*", y "*sacar conclusiones*".

Propiedades

Se identificaron correctamente, entre otras, las siguientes propiedades: *la aleatoriedad presente en el experimento, las leyes de los grandes números y la convergencia de las frecuencias relativas a la probabilidad teórica*, que fue muy patente cuando aumentó el número de ensayos. También se identificaron la *regla de la suma, regla de Laplace y los axiomas de probabilidad: la probabilidad es siempre mayor o igual que cero, la probabilidad es menor o igual a 1 y la suma de probabilidades es la unidad.* Un profesor hace referencia a estos mismos axiomas y su cumplimiento en el caso de la probabilidad condicional. También se cita la equiprobabilidad de los sucesos en el experimento inicial (elegir una tarjeta).

Respecto a las propiedades incorrectamente identificadas, en algunos casos se confunden con conceptos, como *muestreo, sesgo, probabilidad experiencia aleatoria*. En otros con argumento (*argumento deductivo*) o procedimiento (*utilizar el diagrama en árbol*).

Argumentos

Respecto a los argumentos correctamente identificados el citado con mayor frecuencia fue los argumentos aportados por los alumnos o por el profesor. En relación a tipos generales de argumentos se hizo referencia a argumentos combinatorios, y deductivos. En forma más imprecisa se indicaron como argumentos *razonar matemáticamente, demostración, demostración informal.*

En relación a los argumentos incorrectos se mencionan algunas propiedades, como *la regla de la suma* y algunos procedimientos tales como *asociar cada suceso a un valor numérico, táctica elegida*.

En la Tabla 6.7.1 presentamos el número medio de objetos correcta e incorrectamente identificados en cada categoría, junto con algunas medidas de dispersión, tales como el número mínimo de objetos, el máximo y la desviación típica del número de objetos encontrados. En esta tabla y las siguientes tenemos en cuanta sólo los profesores que contestaron a alguna de las preguntas, por lo que el tamaño de la muestra se reduce un poco. Observamos que el profesor en promedio reconoce correctamente entre 1 y 2 elementos de lenguaje y procedimientos y entre 2 y 3 conceptos; estas fueron las categorías más sencillas de reconocer, posiblemente porque son términos más usados por los profesores, ya que las directrices curriculares de los pasados años hacían mucho énfasis en la diferencia entre conocimiento conceptual y procedimental. El resto de los objetos son más difíciles de identificar.

Tabla 6.7.1. Objetos matemáticos correctamente identificados en la tarea (n=156)

Tipos	Mínimo	Máximo	Media	Desviación típica
Situaciones / problemas	0	1	0,59	0,70
Lenguaje	0	7	1,40	1,92
Conceptos	0	11	2,53	1,95
Procedimientos	0	3	1,49	1,72
Propiedades	0	3	0,35	0,80
Argumentos	0	2	0,41	0,64

El pequeño número de objetos citados puede ser debido a que la mayoría de profesores trabajó la demostración del problema, bien en forma empírica, bien utilizando una demostración intuitiva informal, con poco aparato matemático. Esto hace que muchos no reconozcan en el problema los conceptos de experimento compuesto, espacio muestral compuesto o dependencia e independencia, por ejemplo, limitándose a los conceptos, que son más evidentes como probabilidad, suceso y espacio muestral. No obstante, al comparar el número de objetos matemáticos correcta e incorrectamente identificados (Tabla 6.7.2), observamos que los primeros son más numerosos, lo que sugiere un cierto conocimiento especializado del contenido en los profesores, así como la utilidad del taller para hacer aflorar este conocimiento.

Tabla 6.7.2. Objetos matemáticos incorrectamente identificados en la tarea (n=156)

Tipos	Mínimo	Máximo	Media	Desviación típica
Situaciones / problemas	0	2	0,19	0,49
Lenguaje	0	5	0,61	1,26
Conceptos	0	4	0,99	0,92
Procedimientos	0	4	0,42	0,89
Propiedades	0	3	0,64	1,05
Argumentos	0	1	0,29	0,65

Tras realizar un contraste de hipótesis para comparar si las diferencias entre el número medio de objetos matemáticos correcta e incorrectamente identificados (Tabla 6.7.3), obtuvimos diferencias estadísticamente significativas, excepto para el caso de los argumentos, pues los participantes fueron capaces de mostrar pocos ejemplos. Concluimos con ello que los profesores mostraron un conocimiento especializado del contenido aceptable en esta tarea, dado las restricciones de tiempo y el hecho que una parte importante la trabajaron sólo a un nivel intuitivo. Pensamos que con mayor tiempo disponible, se podría haber analizado con los profesores las diferentes soluciones para hacerles profundizar en su análisis del contenido matemático. Actividades como ésta pueden ser importantes para desarrollar el conocimiento del profesor.

Tabla 6.7.3, Resultado de contrastes de diferencias de muestras relacionadas (n=156)

Diferencia correctos e incorrectos	Media	Desv. típ.	Error típ.	Inter. confianza 95%		t	Sig. (bilateral)
				Inferior	Superior		
			0,07	0,24	0,54	5,23	0,000
Lenguajes	0,78	2,53	0,20	0,38	1,18	3,88	0,000
Conceptos	1,53	2,82	0,22	1,09	1,98	6,81	0,000
Procedimientos	1,07	2,04	0,16	0,75	1,40	6,59	0,000
Propiedades	0,29	1,39	0,11	-0,51	-0,07	-2,63	0,009
Argumentos	0,12	1,02	0,08	-0,04	0,28	1,48	0,140

6.7.2. CONOCIMIENTO DEL CONTENIDO Y EL ESTUDIANTE

En la parte final del taller tomó parte únicamente una submuestra de 70 de profesores, pues el tiempo disponible no fue el mismo en todos los talleres, como se ha indicado. Se pidió identificar algunas posibles dificultades y errores que podrían tener los estudiantes de los profesores participantes, si llevasen a cabo la actividad. Esta pregunta va orientada a evaluar el conocimiento que tienen los profesores del contenido y los estudiantes para este problema particular. Godino (2009) indica que pedir al profesor que describa los principales tipos de conflictos de aprendizaje en la resolución de este tipo de tareas por los alumnos es una tarea útil para desarrollar y evaluar este tipo de conocimiento y que esta tarea les permitirá en su práctica profesional conocer si se han adquirido los objetivos de aprendizaje.

Recogidos los protocolos, se prepararon dos nuevas variables para codificar el número de conflictos correcta e incorrectamente identificados por los profesores. Entre los conflictos correctamente identificados encontramos los siguientes:

- No percibir la dependencia en los experimentos que intervienen en el juego. Ya describimos en el análisis a priori que esta es una de las posibles dificultades más frecuentes y algunos profesores la identificaron, quizás porque ellos mismos consideraron al principio de la actividad que los experimentos eran independientes. Por ejemplo: "*Se puede suponer que el hecho de saber el color de una cara de la tarjeta no afecta al color de la que hay detrás"; "no se entiende como el color de la tarjeta (que lo vemos después de haber elegido una tarjeta el profesor) puede dar información de la tarjeta que se ha elegido; cuesta mucho porque la tarjeta se tomó antes de que se viese la cara"*. En este argumento vemos una referencia a la falacia del eje de tiempos descrita por Falk (1986) donde se piensa que un suceso no puede condicionar a otro que ocurrió con anterioridad, por una confusión entre condicionamiento y condicionalidad.
- Construir el espacio muestral, por ejemplo una participante sugiere: "*Una equivocación seria considerar las caras como sueltas en las tarjetas, es decir contar tres caras rojas y tres azules, sin darse cuenta que están pegadas en las tarjetas". "Cuesta trabajo saber cuáles son los casos favorables y posibles en este juego".*
- Asignación incorrecta probabilidades a los sucesos, otro conflicto que se previó en el análisis a priori. Efectivamente, hay una tendencia en este juego a considerar los colores como equiprobables. Algunos profesores detectan esta dificultad, por ejemplo: "*Que piensen que hay la misma probabilidad puesto que se toman los colores como tarjetas independientes*"; "*Es difícil entender que en la tarjeta con las*

dos caras del mismo color hay que considerar dos posibilidades; no se ve la diferencia de esta tarjeta con las otras dos".

- No percibir la estructura del experimento. En el juego se propone un experimento compuesto: En una primera fase se elige al azar una tarjeta entre tres disponibles, en una segunda se elige una de las dos caras para mostrarla y finalmente se predice el color de la cara oculta. Pero a veces el segundo experimento (voltear la tarjeta y elegir al azar una de las caras) no se percibe como relevante para el problema, a pesar de que el formador de profesores insistió en este punto. Algunos participantes aludieron a esta dificultad: *"Yo no entendí porque era importante voltear la tarjeta,… y supongo que mis alumnos no lo comprenderían"*.
- Resistencia a considerar que los fenómenos aleatorios se puedan estudiar de forma matemática. Este es un nuevo conflicto no previsto en nuestro análisis a priori y que fue descrito por Ruiz (2006) en relación a la variable aleatoria. En efecto, algunos estudiantes pueden pensar que no hay una matemática que pueda analizar los fenómenos aleatorios. Aunque el resultado de cada experimento no se puede prever, el cálculo de probabilidades nos permite predecir el comportamiento del colectivo, lo que puede ser difícil para los estudiantes al iniciar el estudio del tema. Esta dificultad fue sugerida por algunos participantes: *"Que les cueste identificar la estrategia y la vean azarosa y les cueste ya que no la atribuyan a una explicación matemática"*. Se hace una referencia a los estudiantes que no comprenden la utilidad de la estadística para analizar fenómenos aleatorios.

También algunos profesores indicaron algunos conflictos que no aparecen en la situación o que no son conflictos importantes, como los siguientes ejemplos: *"Los estudiantes podrían no estar atentos o no interesarse por el problema"; "Se pueden hacer errores al anotar el número de aciertos con cada estrategia y entonces no se dan cuenta de la estrategia óptima"* o bien fueron imprecisos *"Pueden dar una justificación errónea de la estrategia"; "Pueden elegir una estrategia equivocada"*.

En la Tabla 6.7.3 presentamos los resúmenes estadísticos del número de conflictos correcta e incorrectamente identificados por los profesores. Observamos que el número de conflictos identificados fue pequeño, conformándose la mayoría con dar uno o dos ejemplos, aunque en algunos casos llegan hasta 5. Una posible razón es el cansancio natural, pues esta parte ya se hizo al final del taller. Por otro lado, eran pocos los participantes que habían hecho un curso de didáctica de la estadística y, como hemos dicho, el tipo de actividad era nueva para ellos, por lo que no tenían experiencia con sus alumnos sobre situaciones parecidas.

Tabla 6.7.3. Número medio de dificultades identificadas por los profesores (n=70)

Tipos de dificultades	Mínimo	Máximo	Media	D. típica
Identificadas correctamente	0	5	1,63	1,24
identificadas incorrectamente	0	3	0,24	0,67

No obstante, el número de conflictos correctamente identificados fue mayor que el de los incorrectos (siendo las diferencias estadísticamente significativas con un valor p= 0,001), lo que de nuevo indica que el taller permitió la reflexión de los profesores sobre este punto. Pero, también los resultados apuntan a la importancia de realizar actividades de este tipo para los profesores con el fin de aumentar su conocimiento del contenido y los estudiantes.

6.7.3. CONOCIMIENTO DEL CONTENIDO Y LA ENSEÑANZA

Según Brousseau (1986), el alumno aprende adaptándose a un medio lleno de contradicciones, de dificultades y desequilibrios. Este saber, fruto de la adaptación del alumno, se manifiesta por respuestas nuevas que son la prueba del aprendizaje. El autor llama "situación" a un modelo de interacción de un sujeto con cierto medio que determina a un conocimiento dado y es el recurso del que dispone el sujeto para alcanzar o conservar este conocimiento. Algunas de estas "situaciones" requieren de la adquisición anterior de todos los conocimientos y esquemas necesarios, pero hay otras que ofrecen una posibilidad al sujeto para construir por sí mismo un conocimiento nuevo. La teoría distingue cuatro tipos de situaciones didácticas:

- *Situaciones de acción:* El alumno debe actuar sobre un medio (material, o simbólico); la situación requiere solamente la puesta en acto de conocimientos implícitos. En nuestro caso sería la fase de juego y búsqueda de la estrategia ganadora.
- *Situaciones de formulación:* Un alumno (o grupo de alumnos) emisor debe formular explícitamente un mensaje destinado a otro alumno (o grupo de alumnos) receptor que debe comprender el mensaje y actuar (sobre un medio, material o simbólico) en

base al conocimiento contenido en el mensaje. La finalidad es ejercitar el lenguaje matemático, oral y escrito. En nuestro caso sería tanto la fase de escritura de la estrategia, como de la lectura de la estrategia de otro compañero.

- *Situaciones de validación:* Dos alumnos (o grupos de alumnos) deben enunciar aserciones y ponerse de acuerdo sobre la verdad o falsedad de las mismas. Las afirmaciones propuestas por cada grupo son sometidas a la consideración del otro grupo, que debe tener la capacidad de "sancionarlas", es decir ser capaz de aceptarlas, rechazarlas, pedir pruebas, oponer otras aserciones. Es la fase de demostración de la solución y de debate en el taller de las diferentes soluciones hasta llegar a un consenso.
- *El último tipo es el de institucionalización.* "La consideración "oficial" del objeto de enseñanza por parte del alumno, y del aprendizaje del alumno por parte del maestro, es un fenómeno social muy importante y una fase esencial del proceso didáctico: Este doble reconocimiento constituye el objeto de la institucionalización. La institucionalización es de alguna manera complementaria a la devolución. Brousseau (1986) reconoce en estos dos procesos los roles principales del maestro, y afirma que pone a la luz uno de los aspectos teóricos y prácticos más delicados de la articulación entre ambos procesos: los comportamientos o las producciones "libres" del alumno durante las fases a-didácticas de aprendizaje indican el sentido de los conocimientos que los alumnos construyen. Por ello la institucionalización supone preservar el sentido de los conocimientos construidos por los alumnos. Esta es la fase final en que el formador de profesores pone en común lo que se aprendió a lo largo del taller.

Fueron 70 los profesores que llegaron a esta última fase del taller. En la Tabla 6.7.4 presentamos los porcentajes de profesores que identificaron correctamente las diferentes fases en el proceso de estudio implementado, observando que el porcentaje de identificación fue bastante razonable, pues algunos de ellos no conocían la teoría de situaciones con anterioridad a la realización del taller. Las identificaciones incorrectas se deben, bien a confusión de un tipo de situación con otra, bien a no llegar a completar esta parte, posiblemente de nuevo porque el tiempo necesario tendría que haber sido mayor.

Tabla 6.7.4. Frecuencia de profesores que identifican correctamente la teoría de Brousseau (n=70)

Tipos	Correctamente	Incorrectamente
Acción	51 (78,5)	14 (21,5)
Formulación	48 (73,8)	17 (26,2)
Validación	45 (69,2)	20 (30,8)
Institucionalización	41 (63,1)	24 (36,9)

6.8. CONCLUSIONES DEL ESTUDIO 5

En el Estudio 5, descrito en este capítulo nos planteamos los siguientes objetivos:

Objetivo 1: Hacer experimentar a los profesores una situación didáctica, basada en una paradoja clásica de la teoría de la probabilidad, para hacer aflorar algunos conocimientos matemáticos y didácticos con relación a la probabilidad condicional.

Objetivo 2. Enfrentar a los participantes con sus intuiciones incorrectas y concienciarlos de la utilidad de la teoría de la probabilidad para resolver posibles conflictos en las mismas.

Objetivo 3. Introducir algunos de los tipos de análisis didácticos propuestos por Godino, Font y Wilhelmi (2008), ejemplificándolos en la situación y evaluar los conocimientos didácticos de los participantes.

La principal hipótesis que nos planteamos en el Estudio 5 fue que *la actividad realizada, permitiría hacer aflorar algunos de los sesgos en el razonamiento condicional encontrados en los Estudios 1 y 2 y superarlos en una parte de los profesores participantes.*

Para lograr estos objetivos se organizó un taller, basado en un trabajo previo de Batanero, Godino y Roa (2004), y complementado con una fase de análisis de la actividad del taller y de la forma en que se había trabajado en el aula. Los datos se recogieron mediante un protocolo abierto. Dicho taller se ofertó en congresos y jornadas para profesores y dentro de algunas asignaturas optativas en tres contextos diferentes. A lo largo del capítulo se ha realizado un análisis cuidadoso de las estrategias iniciales y finales en el juego propuesto en el taller, la demostración de las soluciones y el análisis didáctico llevado a cabo en una parte de los talleres.

Nuestra hipótesis inicial se confirma y los resultados mostrados señalan la necesidad de mejorar el conocimiento común y especializado de la probabilidad condicional en los profesores, incluso de los profesores en ejercicio. En una proporción considerable de los participantes se mostraron intuiciones incorrectas al comienzo de la actividad, la mayoría de las cuáles ya se habían mostrado en los Estudios 1 y 2. Por ejemplo, se observa en muchos participantes la falta de percepción de la independencia de los diferentes experimentos aleatorios que configuran el juego, considerando que no hay estrategia posible. Dicha creencia implica, además, que no se comprende la utilidad del análisis de los datos estadísticos o de la teoría de la probabilidad para mejorar la predicción global de la distribución de aparición de los diferentes sucesos y para mejorar la toma de decisiones.

En otros casos, los profesores manifiestan la heurística de la representatividad (Kahneman, Slovic y Tversky, 1982), apostando al color contrario del que se muestra en la tarjeta, esperando una convergencia de las frecuencias en pocos ensayos. Pensamos también que estas creencias implican a la falacia del eje de tiempos descrita por Falk (1986) donde se piensa que un suceso no puede condicionar a otro que ocurrió con anterioridad, que también se mostró con alta frecuencia en el Estudio 2.

Señalamos que muchos profesores no fueron capaces de dar una demostración completa de la estrategia, una vez identificada al final del juego, pues no identificaron fácilmente cuál eran los sucesos condición y condicionante, y fallaron en las fórmulas de la probabilidad condicional o conjunta. En otros casos se apoyaron simplemente en justificaciones empíricas, que no tienen la misma validez matemática que una demostración formal. Pensamos que todos estos problemas se deben a que la preparación matemática de los profesores en su formación inicial se centra en el conocimiento común del contenido, y sería necesario complementarlo con un conocimiento especializado del contenido, que incluye información sobre los sesgos en el razonamiento probabilístico. Por otro lado, como consecuencia del taller, alrededor de un 50% de los participantes pasa de una estrategia incorrecta a la correcta, mostrando un reconocimiento de sus sesgos de razonamiento y la superación de los mismos.

Para proporcionar a los participantes en el taller estos conocimientos, la actividad se complementó con una segunda sesión de análisis didáctico sobre el conocimiento matemático que necesitan los estudiantes para resolver el problema, los errores potenciales de los estudiantes y las fases didácticas del taller. A pesar del poco tiempo, los profesores lograron realizar un primer análisis didáctico del taller, y mostraron algunos conocimientos didácticos sobre el tema. Sus respuestas fueron razonables, dado el tiempo disponible y muestran la utilidad del taller para promover la reflexión didáctica de los profesores. La actividad también ayudó a aumentar algunas componentes de los conocimientos didácticos (Godino, 2009):

- *Componente epistémica*: Al hacer reflexionar al futuro profesor sobre los diversos significados de la probabilidad (subjetiva, frecuentista y clásica) y la forma en que aparecen en el taller. También al hacerles reflexionar sobre las ideas estocásticas fundamentales (Heitele, 1975) presentes en la actividad y de la importancia que estas ideas tienen en toda situación aleatoria. Asimismo al hacerles reflexionar sobre los diferentes tipos de objetos matemáticos y como se entrelazan en esta situación.
- *Componente cognitiva*: Adquiriendo conocimientos sobre las dificultades de los estudiantes con los conceptos de probabilidad condicional e independencia y sobre posibles razonamientos y estrategias correctas e incorrectas para resolver el problema.
- *Componente afectiva*: Experimentando nuevos métodos de enseñanza, basados en el juego, experimentación y debate, que permite aumentar el interés de los alumnos y su participación en la actividad.
- *Componente interaccional*: Aumentando su experiencia sobre la forma de organizar el discurso y el tiempo didáctico y de hacer aflorar y resolver los conflictos cognitivos de los estudiantes.
- *Componente ecológica*: Conectado la actividad con el estudio de las concepciones erróneas sobre el azar (psicología) y con los problemas sociales relacionados con la adicción a los juegos de azar (sociología).

No obstante, los resultados sugieren que profesores requieren mejorar su conocimiento didáctico de la probabilidad condicional, incluyendo aspectos como conocimientos del contenido y los estudiantes y el contenido y la enseñanza. Como indican Ponte y Chapman (2006), debemos considerar a los profesores como profesionales, y formarlos en la práctica profesional, haciendo que los elementos de la práctica (tareas y materiales, realización de clases, observación y reflexión sobre la experiencia) sean el elemento central del proceso de formación del profesorado. Actividades como la analizada pueden servir al mismo tiempo para aumentar los conocimientos de probabilidad en los docentes y sus conocimientos didácticos, pues pueden posteriormente ser utilizadas por los profesores con sus estudiantes.

CAPÍTULO 7

CONCLUSIONES

7.1. INTRODUCCIÓN

En esta memoria hemos presentado una serie de estudios, todos ellos relacionados con la formación de profesores para enseñar probabilidad condicional y que comparten el mismo marco teórico.

En los Estudios 1 y 2 se abordó la evaluación de los conocimientos de los futuros profesores de educación primaria y de educación secundaria sobre probabilidad condicional, para recabar información sobre las necesidades formativas de los profesores en los niveles no universitarios, en algunos puntos de su conocimiento matemático y didáctico.

En los Estudios 3 y 4 se han analizado algunos recursos didácticos que podrían ser útiles para atender a estas necesidades de formación del profesorado. Los abundantes recursos en Internet, así como las paradojas clásicas de la historia de la probabilidad, permiten organizar situaciones de reflexión matemática y didáctica para los profesores. Esta utilidad se pone de manifiesto en el Estudio 5, en que se muestran los resultados de un taller formativo para profesores, basado en uno de estos recursos.

Aunque cada uno de estos estudios es independiente de los demás, todos ellos están relacionados con la probabilidad condicional y la formación de profesores. Además, todos ellos se han apoyado en las mismas nociones teóricas del Enfoque Ontosemiótico de la Didáctica de la Matemática propuesto por Godino y sus colaboradores: Teoría de los significados institucionales y personales de los objetos matemáticos (Godino y Batanero, 1994, 1998a y b); Teoría de las funciones semióticas (Godino, 2002; Godino, Batanero y Roa, 2005) y Teoría de las configuraciones didácticas (Godino, Wilhelmi y Bencomo, 2005; Godino, Contreras y Font, 2006), junto con los tipos de análisis asociados a dichas teorías. Por otro lado, en los Estudios 1, 2 y 5 se considera el marco teórico de Hill, Ball, y

Schilling (2008), para describir las categorías de conocimientos del profesor que serán objeto de estudio, y la metodología propuesta para la evaluación de estos conocimientos por Godino (2009). En resumen, esta problemática, marco teórico e instrumentos analíticos compartidos (en mayor o menor grado) en los diferentes estudios otorgan una unidad a la investigación en su conjunto.

Puesto que cada uno de los estudios citados se han las conclusiones obtenidas respecto a los objetivos e hipótesis específicos con suficiente detalle, en este capítulo final, para no ser reiterativos, nos limitaremos a exponer las conclusiones generales que se deducen del conjunto de estudios realizados. Dichas conclusiones se discuten en relación con los objetivos generales del trabajo, las aportaciones y limitaciones del estudio y las líneas de investigación futuras.

7.2. CONCLUSIONES RESPECTO A LOS OBJETIVOS GENERALES DEL TRABAJO

En el primer capítulo se describieron los objetivos generales planteados en el trabajo, que fueron los siguientes:

Objetivo 1. La investigación pretendía proporcionar alguna información sobre el conocimiento matemático de los futuros profesores de educación primaria y secundaria, respecto a la probabilidad condicional.

Objetivo 2. Asimismo se deseaba evaluar algunos aspectos del conocimiento didáctico sobre probabilidad condicional en futuros profesores de educación primaria.

Para responder a estos dos objetivos, se organizaron dos estudios diferenciados. En el Estudio 1, *Evaluación de conocimientos de futuros profesores de educación primaria*, presentado en el capítulo 3, se propuso una tarea abierta, con datos representados en una tabla de contingencia, a una muestra de 183 futuros profesores. La primera parte de la tarea constaba de tres preguntas en las que se había de calcular la probabilidad simple, condicional y compuesta. En esta parte se compararon los resultados con los obtenidos por Estrada y Díaz (2007) en una tarea semejante y una muestra reducida de futuros profesores de educación primaria de la Universidad de Lleida que habían seguido previamente un curso optativo de didáctica de la estadística.

El análisis semiótico de las respuestas a estos problemas describir el conocimiento matemático de los futuros profesores, definiendo categorías de respuestas correctas e incorrectas y conflictos semióticos. Además de cumplir el primer objetivos planteados para la Tesis, en este estudio se confirman las hipótesis previas planteadas en el Estudio 1 pues, el análisis y resultados de la primera parte de este estudio nos indicó que la lectura de tablas estadísticas de doble entrada y el cálculo de probabilidades a partir de ellas no fue una tarea trivial, para los futuros profesores de Educación Primaria de la muestra.

La mayoría de los estudiantes de la muestra fue incapaz de calcular correctamente la probabilidad compuesta y condicional, habiendo coincidencia en los errores más frecuentes con los resultados proporcionados por Einhorn y Hogarth (1986), Ojeda (1995) y Estrada y Díaz (2007), quienes sugieren que la principal dificultad reside en la confusión de la probabilidad condicional y la conjunta.

En el caso del cálculo de la probabilidad simple, un tercio de los futuros profesores cometen algún tipo de error, destacando la confusión de probabilidad con frecuencia absoluta, un error que nos parece importante dado que implica la no apreciación de los axiomas de probabilidad. El número de estudiantes que dejaron en blanco las tareas, indica también una pobreza de conocimientos básicos de probabilidad.

Por otro lado, aunque se confirman los conflictos semióticos identificados en trabajos previos, también se aportan otros no descritos en los antecedentes tales como la confusión de un suceso con su complementario, errores en la fórmula de cálculo, errores de cálculo, o la errónea identificación de los datos del enunciado. Y, en menor medida, la confusión de frecuencia absoluta con la probabilidad o invertir la fórmula del cálculo de probabilidades, obteniendo valores mayores que uno. Nuestros resultados fueron peores que los del estudio Estrada y Díaz (2007), un 10%, 11,4% y 11,7% respectivamente menos de respuestas correctas en las tres preguntas que en el estudio de Estrada y Díaz.

La segunda y tercera preguntas planteadas en el Estudio 1, estaban dirigidas a evaluar algunos conocimientos didácticos del contenido de los futuros profesores, respecto a dos componentes en el modelo de Hill, Ball, y Schilling (2008): (a) conocimiento especializado del contenido, y (b) conocimiento del contenido y la enseñanza.

Para esta segunda parte del Estudio 1, la hipótesis era que nuestros resultados mostrarían un conocimiento insuficiente de los futuros profesores de la muestra en relación a las componentes evaluadas del conocimiento didáctico del contenido. Esta hipótesis se vio también confirmada por las escasas respuestas de los estudiantes y los errores de identificación, tanto de los objetos matemáticos en la tarea, como de las variables de tarea. Los futuros profesores logran identificar algunos objetos matemáticos, principalmente conceptos y procedimientos, aunque insuficientes para la correcta resolución del problema. Respecto a la valoración del conocimiento del contenido y la enseñanza, los futuros profesores tuvieron problemas para identificar variables didácticas que pudiesen variar la dificultad del problema o variar su contenido matemático, principalmente a la hora de identificar variables del sujeto y de la situación de resolución. En este punto, nuestro trabajo aporta resultados originales, pues no se conocen antecedentes sobre evaluación de estos conocimientos en la tarea dada en futuros profesores.

En el capítulo 4 se ha presentado otro estudio de evaluación (Estudio 2), sobre algunos conocimientos matemáticos y sesgos en relación a la probabilidad condicional de futuros profesores de educación secundaria. Se propuso a una muestra de 196 futuros profesores de educación secundaria de varias universidades españolas (95 alumnos de la licenciatura de Matemáticas y 101 alumnos del Máster de formación del profesorado de secundaria una parte del cuestionario RPC (Díaz, 2007). Las respuestas de opción múltiple y las abiertas del cuestionario permitieron evaluar por un lado los principales sesgos sobre razonamiento en probabilidad condicional que se describen en los antecedentes. Por otro, el conocimiento del concepto y propiedades, así como la capacidad de resolver los principales tipos de problemas identificados en el apartado 1.4, en nuestro análisis del significado de referencia de la probabilidad condicional en este trabajo.

En este estudio también se confirmó parcialmente nuestra hipótesis previa. Por un lado, encontramos un alto porcentaje de futuros profesores que incurre en las diferentes falacias o sesgos relacionados con la probabilidad condicional, especialmente la *falacia del eje de tiempos,* la *falacia de las tasas base*, y *confusión entre probabilidad conjunta y condicional*, con uno de cada tres encuestados. El resto de sesgos, *falacia de la conjunción* y de la *condicional transpuesta*, se dieron en un porcentaje pequeño de futuros profesores. Por otro lado, los resultados en algunos de los ítems de opción múltiple han sido peores que los obtenidos por Díaz (2007) en estudiantes de Psicología, a pesar de la mayor preparación matemática de los alumnos de nuestra muestra.

Al contrario de los resultados obtenidos en los ítems que miden los sesgos, en los ítems abiertos el porcentaje de alumnos que cometieron errores fue mucho menor que los de la muestra del estudio de Díaz. Por tanto, también se confirmó nuestra hipótesis al respecto. Solamente en el caso del ítem relacionado con el teorema de Bayes se dio un porcentaje de errores mayor en la muestra de futuros profesores que en la de psicología. Por otro lado, los resultados coinciden el ítem 9, relacionado con el cálculo de la probabilidad condicional, que fue el ítem que presenta más dificultades por parte de ambas muestras.

El análisis semiótico de las respuestas abiertas en el Estudio 2 nos ha permitido detectar diferentes conflictos semióticos, confirmando la presencia de sesgos detectada en los ítems de opciones múltiples. Destacamos la confusión entre condicionamiento y causalidad, presentada en un 32% de las respuestas del ítem 8, la falacia del eje temporal, presentada en un 17,3% de las respuestas la ítem 8, la falacia del jugador se presenta en un 6,1% de estudiantes en el ítem 11, confusión a la hora de determinar el espacio muestral (16%) en el ítem 9, o dificultades anteriormente descritas por Sánchez (1996) y Kelly y Zwiers (1986) como confundir sucesos independientes con sucesos excluyentes.

En resumen, los Estudios 1 y 2 de evaluación contribuyen a proporcionar información detallada de las necesidades formativas de los futuros profesores de educación primaria y secundaria respecto a la probabilidad condicional, cumpliéndose, de este modo, nuestros dos primeros objetivos.

Objetivo 3. Se pretendía analizar algunos recursos educativos que tuviesen potencialidad para ayudar a los futuros profesores o profesores en ejercicio respecto a las limitaciones observadas en su conocimiento común y especializado del contenido sobre probabilidad condicional.

Para cumplir con este objetivo, en el capítulo 5 se describen dos estudios relacionados. En el Estudio 3, *Análisis de recursos para la enseñanza de la probabilidad condicional,* se analizan algunos recursos disponibles en Internet útiles para la enseñanza de la probabilidad condicional en diferentes niveles educativos y la formación de los profesores. En el Estudio 4, *Paradojas como recurso didáctico en el estudio de la probabilidad condicional,* se analizan algunas paradojas clásicas ligadas a los conceptos de probabilidad condicional e independencia. En ambos casos se estudia la actividad matemática ligada a los recursos, las posibles dificultades en el trabajo con ellos y su idoneidad didáctica.

Nuestra primera hipótesis sobre estos dos estudios fue que *podríamos identificar una variedad de recursos que mostrasen la riqueza de objetos matemáticos relacionados con la probabilidad condicional*. Asimismo *pensábamos mostrar la* *idoneidad didáctica de dichos recursos* para la formación de profesores en el campo de la probabilidad condicional.

Las conclusiones de los Estudios 3 y 4 indican el cumplimiento de los objetivos e hipótesis previas. El resultado del Estudio 3 fue un listado bastante amplio de recursos que proporcionan ayudas de aprendizaje, ampliando el libro de texto clásico; plantean una visión diferente del término probabilidad condicionada, al basarse en muchas situaciones de la vida diaria; plantean problemas paradójicos, proporcionando explicaciones de sus soluciones; permiten la visualización de objetos abstractos, simulación y experimentación con experimentos aleatorios. En resumen proporcionan al estudiante un apoyo visual y experimental del que carecen los libros de texto. El análisis de la idoneidad didáctica en cada uno de sus componentes mostró que un uso adecuado de estas ayudas, debidamente insertado en el proceso de aprendizaje ayudará al estudiante o al futuro profesor a comprender las propiedades y aplicaciones de la probabilidad condicional.

También en el Estudio 4 se mostró la utilidad potencial e idoneidad didáctica de dichas paradojas en la formación de profesores. En el trabajo con dichas paradojas se identificaron la mayor parte de ideas fundamentales descritas por Heitele (1975): Sucesos y espacio muestral, probabilidad y convergencia, operaciones combinatorias, las reglas de adición y multiplicación de probabilidades, independencia, probabilidad condicionada, variable aleatoria, equidistribución y simetría, esperanza matemática y muestreo. Además, el trabajo con las paradojas en los cursos para profesores tienen potencial para aumentar los conocimientos didácticos del contenido probabilidad de los docentes en algunas componentes de los conocimientos matemático-didácticos descritos por Godino (2009).

Finalmente indicamos que se han cumplido asimismo los objetivos de valoración de los objetos y procesos matemáticos implícitos en el uso de los dos tipos de recursos, junto con la valoración de las dificultades potenciales del trabajo con los mismos. Ya que, como indicamos, estos recursos por si mismos no resuelven las dificultades presentadas a la hora de plantear una enseñanza de la probabilidad, por lo que se ha de incorporar estos recursos dentro de unas unidades didácticas acordes a los estudiantes o futuros profesores. Con todo ello se cumple el tercer objetivo general del trabajo.

Objetivo 4. Se deseaba evaluar el cambio en el conocimiento matemático de los profesores como consecuencia de una actividad formativa basada en los recursos analizados en el objetivo 4.

En el capítulo 6 se presenta el Estudio 5, *Evaluación de una experiencia de formación de profesores de educación secundaria*, donde se analiza una experiencia formativa basada en una paradoja clásica, dirigida a profesores de educación secundaria o superior en formación y ejercicio. En dicho estudio se analizan algunos conocimientos matemáticos de los participantes y su evolución, junto que aspectos relevantes de sus conocimientos didácticos.

Un taller formativo basado en una de las paradojas descritas en el capítulo 4 y propuesto por Batanero, Godino y Roa (2004) es experimentado en varios cursos, dirigidos a profesores, con un total de 166 profesores en formación o ejercicio en tres contextos geográficos diferentes. El análisis semiótico de las estrategias iniciales y finales en un juego y contraste Chi-cuadrado permitió evaluar el cambio de concepciones y test t se utilizó para analizar los resultados sobre conocimientos didácticos de los participantes.

Los resultados del taller indicaron la necesidad de mejorar el conocimiento especializado del contenido probabilístico en los profesores, incluso de los profesores en ejercicio, quienes en una proporción considerable mostraron intuiciones incorrectas al comienzo de la actividad y no fueron capaces de dar una demostración completa de la estrategia una vez identificada al final del juego.

Sin embargo, la actividad sirvió para cambiar las concepciones de una parte importante de la muestra, confirmándose nuestra hipótesis al respecto. Hubiese sido necesario un mayor tiempo para la actividad, pues 1/3 de los participantes siguieron manteniendo sus concepciones erróneas al finalizar la misma. Destacan los mejores resultados presentados por los profesores en formación que por los profesores en ejercicio, que siguen manteniendo en un alto porcentaje la creencia que no es posible una estrategia óptima, debido a aleatoriedad de la experiencia.

Asimismo, observamos mejores resultados de la muestra española en comparación con la de los otros dos países, lo que creemos que es debido a la mayor formación matemática de esta muestra. Estos resultados replican los de Batanero, Godino y Roa (2004), con una muestra más reducida de futuros profesores, y se repiten en diferentes contextos educativos, no habiendo gran diferencia entre profesores en formación y en ejercicio. Por otro lado, ampliamos las categorías de respuestas y realizamos un análisis más profundo de las mismas, que la llevada a cabo por Batanero, Godino y Roa.

Los resultados en la evaluación del conocimiento didáctico de los profesores fueron pobres, pero posiblemente el resultado fuese debido al tiempo limitado dedicado al taller. Por otro lado, la actividad se mostró útil para provocar la reflexión didáctica del profesor. Entre otros aspectos, esta actividad hizo reflexionar al profesor sobre los diversos significados de la probabilidad y los tipos de objetos matemáticos implícitos (componente epistémica), les hizo adquirir conocimientos sobre las dificultades de los estudiantes, los posibles razonamientos y estrategias correctas e incorrectas (componente cognitiva), experimentar nuevos métodos de enseñanza que aumenten el interés y la participación de los alumnos (componente afectiva), aumentar su experiencia sobre la forma de resolver los conflictos cognitivos de los estudiantes (componente interaccional) y conectar la actividad con otras investigaciones de otras áreas de estudio relacionadas (componente ecológica).

En consecuencia, creemos que el objetivo 4 se ha completado de forma razonable mediante la información proporcionada.

7.3. APORTACIONES Y LIMITACIONES DEL ESTUDIO

El trabajo realizado presenta una serie de aportaciones de interés para la didáctica de la probabilidad y la formación de profesores, las principales de las cuales se resumen a continuación.

En primer lugar proporcionamos una información detallada sobre los sesgos y dificultades en el razonamiento relacionado con la probabilidad condicional, tanto en futuros profesores de educación primaria (Estudio 3), como futuros profesores (Estudio 4) y profesores en ejercicio (Estudio 6) de educación secundaria.

Aunque las tareas y el cuestionario utilizados no son originales, sino tomados de investigaciones previas, la mejora en los protocolos de recogida de datos y el análisis semiótico de las respuestas a los mismos es mucho más profundo que el realizado en las investigaciones de donde se tomaron los diferentes instrumentos, lo que confiere a nuestros estudios mayor validez, respecto a los originales. Los mayores tamaños de muestra y diferentes localizaciones geográficas, para los Estudios 2 y 5 permiten dotar de mayor fiabilidad a nuestros resultados.

Todo ello permite no sólo confirmar los hallazgos de Batanero, Godino y Roa (2004) y Díaz (2007) sino identificar nuevos conflictos semióticos o explicar conflictos ya descritos en la comprensión de la probabilidad condicional en profesores en formación y ejercicio. En el Estudio 1 fueron los siguientes:

- Confundir probabilidad simple y condicional o bien la simple con la compuesta, debido a una falta de capacidad para la lectura de la tabla de contingencia, no llegándose al nivel de "leer los datos" en la terminología de Curcio (1989).
- Confundir frecuencia absoluta o casos posibles con probabilidad simple o compuesta; que, además, implica el no reconocimiento de uno de los axiomas de probabilidad (la probabilidad de un suceso es un número siempre comprendido entre 0 y 1).
- Confusión de la probabilidad simple con la probabilidad de un suceso elemental. Este conflicto no aparece descrito en las investigaciones previas.
- Suponer independencia en los datos, a pesar de que la dependencia es claramente visible, conflicto tampoco descrito en las investigaciones previas. Por un lado, ello implica un olvido de la definición de la independencia y por otro, de nuevo se muestra poca competencia en la lectura de datos de la tabla.

Respecto a los futuros profesores de educación secundaria, en el Estudio 2 hemos aportado información sobre la frecuencia con que aparecen en la muestra diferentes sesgos sobre el razonamiento condicional, ya descritos en las investigaciones reseñadas en el Capítulo 2 o en la de Díaz (2007). Proporcionamos también información numérica sobre su conocimiento para resolver problemas de probabilidad condicional, analizando el número de problemas correctamente resueltos. El análisis semiótico de las respuestas en problemas abiertos permite ampliar notablemente las categorías cuantitativas descritas en la investigación de Díaz (2007) y proporcionar explicación de los errores en la resolución, en términos de conflictos semióticos.

Por otro lado, se realiza una comparación detallada de los resultados en cada ítem y en el número de ítems correctos (relacionados con sesgos o problemas abiertos) con la muestra de estudiantes de Psicología de Díaz (2007). Se obtiene la sorprendente conclusión de que, aunque los conocimientos para resolver problemas son mejores en los futuros profesores, en general, no ocurre así en los problemas relacionados con el teorema de Bayes. Las definiciones del concepto dadas por los estudiantes de psicología son también mejores que las dadas por los futuros profesores. Finalmente, en algunos ítems relacionados con sesgos, los futuros profesores tienen peores resultados.

Al estudiar la posible relación entre sesgos de razonamiento y conocimiento formal, se confirmó la hipótesis de ausencia de correlación significativamente fuerte entre uno y otro tipo de ítems. Asimismo se replicaron de una forma bastante consistente los resultados del análisis factorial realizado por Díaz (2007) del cuestionario completo en una muestra de 600 estudiantes. En consecuencia, también en el análisis factorial de los resultados en nuestra muestra los ítems que evalúan los sesgos aparecen en diferentes factores que los relacionados con la resolución de problemas. Este resultado nos sirvió para confirmar la validez de constructo del cuestionario de Díaz (2007).

Respecto a los resultados del taller formativo, el mayor tamaño de nuestra y la inclusión de profesores en ejercicio en el Estudio 5, respecto al realizado con la misma tarea por Batanero, Godino y Roa (2004) nos permite confirmar la presencia de sesgos importantes en profesores en formación y extender los resultados a profesores en ejercicio (no estudiados por los citados autores). También se ha mostrado que los resultados no dependen del contexto educativo y que los resultados en profesores en ejercicio no difieren de los presentados en profesores en formación.

Por otro lado, nuestra evaluación de algunos aspectos del conocimiento didáctico en los Estudios 3 y 6 proporciona resultados originales, en cuanto este conocimiento no fue evaluado en los trabajos de donde se tomaron los instrumentos de evaluación. Aunque las tareas presentadas fueron muy limitadas, los resultados son importantes, en cuanto muestran las necesidades formativas en didáctica de la probabilidad no sólo de los profesores en formación, sino también de los profesores en ejercicio.

Otra aportación original es el análisis proporcionado de los recursos en Internet y las paradojas, que puede ser útil tanto a los formadores de profesores como a los mismos profesores, para organizar actividades formativas sobre probabilidad condicional.

Finalmente, hemos realizado un estudio de las investigaciones previas relacionadas con nuestro objeto de investigación, que clasificamos en dos apartados:

1. Investigaciones sobre formación de profesores para enseñar probabilidad. Se incluyen los aspectos afectivos que influyen en la labor del profesor (actitudes y creencias), estudios de evaluación de los conocimientos probabilísticos de los profesores, y los relacionados con su conocimiento didáctico para enseñar probabilidad. Se hace también un pequeño resumen de algunos modelos específicos que describen los componentes del conocimiento profesional del profesor en el campo de la estadística.
2. Investigaciones sobre la comprensión de la probabilidad condicional y temas relacionados con ella. Analizamos la comprensión conceptual y procedimental,

experimentos de enseñanza del tema e investigaciones centradas en el análisis de libros de texto.

Destacamos asimismo el hecho de que, a lo largo del periodo de elaboración de la Tesis, estos resultados se han ido publicando en revistas y congresos de didáctica de la matemática y estadística, destacando nuestras publicaciones en la revista Quadrante, Educaçao Matemática e Pesquisa, Suma, Epsilon y BEIO, y en los congresos ICOTS-8, CERME 7, PME 34, JAEM XIV, SEIEM 14, RELME 24, CIAEM-13, IX Congreso Galego de Estadística, I CIFOP, XIII CEAM y XXXII Congreso Nacional de Estadística.

Limitaciones del trabajo

Como todo trabajo empírico, esta Tesis tiene sus limitaciones. En cada uno de los estudios de evaluación realizados el tamaño de la muestra ha sido razonable, pero las muestras son intencionales. En este sentido, los resultados de los contrastes estadísticos deben ser interpretados con precaución, y siempre referidos a los contextos utilizados en el estudio. Por otro lado, los contextos geográficos han sido limitados (salvo en el Estudio 4, cuya muestra es mucho más representativa de la situación actual de formación de profesores en España). Sería necesario en todos los casos completar nuestros estudios con nuevas muestras en diferentes contextos que permitiesen ver si se confirman nuestros resultados.

Por otro lado, los datos en los estudios de evaluación se toman a partir de cuestionarios o protocolos escritos, sin tener una interacción directa con los participantes; por lo tanto, no pueden describirse de una forma completa la complejidad y riqueza de los conocimientos de los futuros profesores. Los resultados podrían complementarse con entrevistas en profundidad a los futuros profesores o profesores que presenten los sesgos más importantes, con la finalidad de explicar con mayor detalle las causas de los mismos. En el caso del Estudio 2, sería interesante ampliar el cuestionario proponiendo a los futuros profesores de educación primaria otras tareas que confirmen los conflictos detectados en la propuesta.

Finalmente, el trabajo realizado se centra principalmente en la evaluación de los conocimientos matemáticos, con un peso pequeño del componente didáctico de los conocimientos del profesor, sobre el cual se podría complementar nuestros resultados con nuevos estudios. Asimismo, apenas se ha tocado el tema de la formación de profesores, centrándonos más bien en la evaluación de sus conocimientos en la situación actual.

7.4. LINEAS DE INVESTIGACIÓN FUTURAS

Como consecuencia de las limitaciones señaladas en nuestro trabajo, surgen también nuevas líneas de investigación para continuar el análisis de la formación de profesores para enseñar probabilidad condicional.

Los estudios de evaluación realizados sobre el conocimiento matemático de los profesores en este tema podrían complementarse con nuevas tareas, pues, como se ha indicado, la investigación previa es muy escasa. Tanto en la evaluación de los profesores de educación primaria como en los de secundaria, sería posible ampliar notablemente el rango de problemas planteados. Por ejemplo, se podrían ensayar versiones abiertas de los ítems de opciones múltiples utilizados en el Estudio 4, tanto con profesores de primaria como de secundaria para confirmar si los resultados replican los obtenidos con ítems de opción múltiple.

Otra línea muy amplia de investigación es la de evaluación y formación de los conocimientos didácticos de los profesores, que se estudia de forma limitada en los Estudios 3 y 5 y que se revela claramente insuficiente en nuestros resultados. El éxito de las nuevas propuestas curriculares para la Educación Primaria y Secundaria dependerá de la formación de los profesores que han de llevarla a cabo. Desafortunadamente, nuestro trabajo ha mostrado que dicha formación es insuficiente para el tema de la probabilidad condicional.

En este sentido, y aprovechando el trabajo de análisis realizado de recursos en Internet y paradojas en el capítulo 5, sería necesario continuar el trabajo de diseño de situaciones didácticas para la formación de profesores y realizar nuevas evaluaciones de su idoneidad didáctica, para complementar el trabajo presentado en el Estudio 5.

REFERENCIAS

Afifi, A. A. y Clark, V. (1990). *Computer-aided multivariate analysis.* London: Chapman and Hall.

Albert, J. H. y Rossman, A. (2001). Workshop statistics. *Discovery with data. A bayesian approach.* Bowling Green. OH: Key College Publishing.

American Psychological Association, American Educational Research Association y National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washinton, DC: American Psychological Association.

Arnold, J. C. (2004). *Probabilidad y estadística con aplicaciones para ingeniería y ciencias computacionales*. McGraw-Hill Interamericana.

Azcárate, P. (1995). *El conocimiento profesional de los profesores sobre las nociones de aleatoriedad y probabilidad*. Tesis Doctoral. Universidad de Cádiz.

Azcárate, P. y Cardeñoso, J. M. (2001). Probabilidad. En: E. Castro (Ed.). *Didáctica de la matemática en la educación primaria* (pp. 591-618). Madrid: Síntesis.

Ball, D. L. (2000). Bridging practices: Intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education, 51,* 241-247.

Ball, D. L., Lubienski, S. T., y Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. En V. Richardson (Ed.), *Handbook of research on teaching* (pp. 433-456). Washington, DC: American Educational Research Association.

Ball, D. L., Thames, M. H., y Phelps, G. (2005). *Articulating domains of mathematical knowledge for teaching.* Online: www-personal.umich.edu/~dball/.

Basulto, J. y Camuñez, J. A. (2007). *La geometría del azar. La correspondencia entre Pierre de Fermat y Blaise Pascal*, Nivola, Madrid.

Batanero, C. (2003). La simulación como instrumento de modelización en probabilidad. *Educación y Pedagogía, 35,* 37-64.

Batanero, C., Arteaga, P. y Contreras, J. M. (2009). Estadística en la educación primaria. Retos para la formación de profesores. En M Guzmán (Coord.), *Arte, humanidades y educación. Libro homenaje a Carmen Medialdea* (pp. 49-62). Universidad de Granada.

Batanero, C., Burrill, G. y Reading, C. (Eds.) (En prensa). *Teaching statistics in school mathematics. Challenges for teaching and teacher education. A joint ICMI and IASE* *study*. New York: Springer.

Batanero, C., Burrill, G., Reading, C. y Rossman, A. (Eds.) (2008). *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.* Monterrey: ICMI and IASE. Online: www.stat.auckland.ac.nz/~iase/*publications.*

Batanero, C., Cañizares, M. J. y Godino, J. (2005). Simulation as a tool to train pre-service school teachers. *Proceedings of the First African Regional Conference of ICMI.* Ciudad del Cabo: ICMI. CD ROM.

Batanero, C., Contreras, J. M. y Arteaga (2009). Taller: Ideas estocásticas fundamentales. *Actas de las XIII Jornadas Nacionales de Enseñanza y Aprendizaje de las Matemáticas* [CD-ROM]. Gerona: Federación Española de Profesores de Matemáticas.

Batanero, C., Contreras, J. M., Díaz, C. y Arteaga, P. (2009). Paradojas en la historia de la probabilidad como recurso didáctico. En C. Cañadas y J. M. Contreras (Eds.), *Actas de las XV Jornadas de Investigación en el Aula de Matemáticas.* [CD-ROM]. Granada: Sociedad Thales de Educación Matemática.

Batanero, C., Contreras, J. M., Fernándes, J. A. y Ojeda, M. M. (2010). Paradoxical games as a didactic tool to train teachers in probability, En C. Reading (Ed.), *Proceedings of the Eight International Conference on Teaching Statistics*. Lubjana: International Association for Statistical Education.

Batanero, C. y Díaz, C. (2005). La probabilidad condicional en los textos de estadística para psicología. Trabajo presentado en el *V Congreso Iberoamericano de Educación Matemática.* Guimaraes, Portugal Julio 2005.

Batanero, C., Estepa, A., Godino, J. y Green, D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education*, *27* (2), 151 – 169.

Batanero, C., Fernándes, J. A. y Contreras, J. M. (2009). Un análisis semiótico del problema de Monty Hall e implicaciones didácticas. *SUMA*, *62*, 11-18.

Batanero, C. y Godino, J. D. (2002). *Estocástica y su didáctica para maestros*. Granada: Departamento de Didáctica de la Matemática. Universidad de Granada.

Batanero, C., Godino. J. D. y Cañizares, M. J. (2005). Simulation as a tool to train Pre-service School Teachers, *Proceedings of the First ICMI African Regional Conference.* [CD-ROM]. Johannesburg: ICMI.

Batanero, C., Godino, J. D. y Roa, R. (2004). Training teachers to teach probability. *Journal of Statistics Education, 12.* On line: www.amstat.org/publications/jse/.

Batanero, C., Henry, M. y Parzysz, B. (2005). The nature of chance and probability. En G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 15-37). New York: Springer.

Begg, A. y Edwards, R. (1999). Teachers' ideas about teaching statistics. *Annual Meeting of the Australian Association for Research in Education and the New Zealand Association for Research in Education*. Melbourne, Australia: AARE.

Ben-Zvi, C. (2000). Towards understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning, 2*(1y 2), 127-155.

Berry, D. A. (1995). *Basic statistics: A Bayesian perspective.* Belmont, CA: Wadsworth.

Biehler, R. (1997). Software for learning and for doing statistics. *International Statistical Review*, *65*(2), 167-190.

Bohl, A. H., Liberatore, M. J., y Nydick, R. L. (1995). A tale of two goats... and a car, or the importance of assumptions in problem solutions. *Journal of Recreational Mathematics, 1*, 1–9.

Borim, C. y Coutinho, C. (2008). Reasoning about variation of a univariate distribution: a study with secondary mathematics teachers. En C. Batanero, G. Burrill, C. Reading, y A. Rossman (2008).

Borovcnik, M., Bentz, H. J. y Kapadia, R. (1991). Empirical research in understanding probability. En R. Kapadia y M. Borovcnik (Eds.), *Chance encounters: probability in education* (pp. 73-105). Dordrecht: Kluwer.

Borovcnik, M. y Peard, R. (1996). Probability. En A. Bishop, et al. (Eds.), *International handbook of mathematics education* (pp. 239-288). Dordrecht: Kluwer.

Botella, J., León, O. G. y San Martín, R. (1993). *Análisis de datos en psicología I.* Madrid: Pirámide.

Brousseau, G. (1986). Fondements et Méthodes de la Didactique des Mathématiques. *Recherches en Didactique des Mathématique, 7*(2), 33-115.

Brousseau, G. (1997). *Theory of didactical situations in mathematics. Didactique des* *mathématiques, 1970-1990*. Dordrecht: Kluwer.

Brown, C. y Borko, H. (1992). Becoming a mathematics teacher. En D. A. Grouws (Ed.), *Handbook of research in mathematics teaching and learning* (pp. 209-239). New York, NY: MacMillan.

Burgess, T. (2008). Teacher knowledge for teaching statistics through investigations. En C. Batanero, G. Burrill, C. Reading y A. Rossman (2008).

Canada, D. L. (2008). Conceptions of distribution held by middle school students and preservice teachers. En C. Batanero, G. Burrill, C. Reading y A. Rossman (2008).

Cañizares, M. J. (1997). *Influencia del razonamiento proporcional y de las creencias subjetivas en las intuiciones probabilísticas primarias*. Tesis Doctoral. Universidad de Granada.

Carmines, E. G. y Zéller, R. A. (1979). *Reliability and validity assessment*. London: Sage.

Cardeñoso, J. M., Azcárate, P. y Serradó, A. (2005). Los obstáculos en el aprendizaje del conocimiento probabilístico: Su incidencia desde los libros de texto. *Statistics Education Research Journal, 4*(2), 59-81.

Carnell, L. J. (1997). *Characteristics of reasoning about conditional probability (preservice teachers).* Tesis doctoral. University of North Carolina-Greensboro.

Castro, E. (1991). *Resolución de problemas aritméticos de comparación multiplicativa*. Memoria de Tercer Ciclo. Universidad de Granada.

Chaput, B., Girard, J. C. y Henry, M. (2008). Modeling and simulations in statistics education. En C. Batanero, G. Burrill, C. Reading, y A. Rossman (2008).

Chick, H. L. y Pierce, R. U. (2008). Teaching statistics at the primary school level: beliefs, affordances, and pedagogical content knowledge. En C. Batanero, G. Burrill, C. Reading, y A. Rossman (2008).

Consejería de Educación. Junta de Andalucía (2007a). *ORDEN de 10 de agosto de 2007, por la que se desarrolla el currículo correspondiente a la Educación Primaria en Andalucía.*

Consejería de Educación. Junta de Andalucía (2007b). *ORDEN de 10 de agosto de 2007, por la que se desarrolla el currículo correspondiente a la Educación Secundaria Obligatoria en Andalucía.*

Consejería de Educación. Junta de Andalucía (2008). *ORDEN de 5 de agosto de 2008, por la que se desarrolla el currículo correspondiente al Bachillerato en Andalucía*.

Contreras, J. M. (2009). *Recursos en Internet para la enseñanza de la probabilidad condicionada.* Tesis de Master. Universidad de Granada

Contreras, J. M., Batanero, C. Díaz, C. y Fernandes, J. A. (2011). Prospective teachers' common and specialized knowledge in a probability task. Trabajo presentado en *CERME 7*, Rzeszlow, Polonia.

Contreras, J. M., Batanero, C. y Fernández, J. A. (2010). Problema de Monty Hall. Un análisis semiótico. *XIII Simposio de la SEIEM*. Comunicaciones en los grupos de trabajo.

Contreras, J. M., Batanero, C., Fernándes, J. A. y Ojeda, M. M. (2010) Análisis de una experiencia de formación de profesores en diferentes contextos. En J. Costa, R. Fernández, M. Pesaro y J. Vilar (Eds.), *XXXII Congreso Nacional de Estadística e Investigación Operativa* [CD-ROM]. Universidad de la Coruña y SEIO.

Contreras, J. M., Díaz, C. y Batanero, C. (2009). Recursos para exploración de la probabilidad en Internet. *IX Congreso Galego de Estatística e Investigación de Operacións.* (pp. 201-206). Ourense: Sociedad Galega de Estadistica. Orense: SGAPEIO.

Contreras, J. M., Díaz, C., Batanero, C. y Ortiz, J. J. (2010). Razonamiento probabilístico de profesores y su evolución en un taller formativo. *Educaçao Matemática e Pesquisa*. *12*(2), 181-198.

Contreras, J. M., Estrada, A., Díaz, C. y Batanero, C. (2010). Dificultades de futuros profesores en la lectura y cálculo de probabilidades en tablas de doble entrada. En M. Moreno, A. Estrada, J. Carrillo y T. Sierra (Eds.), *Investigación en Educación Matemática XIV* (pp. 271-280). Lleida: Sociedad Española de Investigación en Educación Matemática.

Crandall, C. S. y Greenfield, B. (1986). Understanding the conjunction fallacy: A conjunction of effects? *Social Cognition, 4*, 408-419.

Cuadras, C. M. (1981). *Análisis multivariante*. Barcelona. Eudeba.

Cuadras, C. M., Echevarría B., Mateo, J. y Sánchez, P. (1988). *Fundamentos de estadística. Aplicación a las ciencias humanas.* Madrid: Promociones Publicaciones Universitarias.

Curcio, F. R. (1989). *Developing graph comprehension*. Reston, VA: N.C.T.M.

Delaney, S., Ball, D. L., Hill, H. C., Schilling, S. G. y Zopf, D. (2008). Mathematical knowledge for teaching: Adapting U.S. measures for use in Ireland. *Journal of*

Mathematics Teacher Education, *11*, 171-197.

Devore, J. L. (1998). *Probabilidad y estadística para ingeniería y ciencias*. Internacional Thomson Editores.

Díaz, C. (2004). *Elaboración de un instrumento de evaluación del razonamiento condicional. Un estudio preliminar*. Trabajo de Investigación Tutelada. Universidad de Granada.

Díaz, C. (2005). Evaluación de la falacia de la conjunción en alumnos universitarios. *Suma*, *48*, 45-50.

Díaz, C. (2007). *Viabilidad de la enseñanza de la inferencia bayesiana en el análisis de datos en psicología*. Tesis doctoral. Universidad de Granada.

Díaz, C., Batanero, C. y Contreras, J. M. (2010). Teaching independence and conditional probability. *Boletín de Estadística e Investigación Operativa. 26* (2), 149-162.

Díaz, C. y de la Fuente, I. (2005a). Razonamiento sobre probabilidad condicional e implicaciones para la enseñanza de la estadística. *Epsilon*, *59*, 245-260.

Díaz, C. y de la Fuente, I. (2005b). Recursos para la enseñanza del razonamiento bayesiano en Internet. Trabajo presentado en el Congreso Internacional: *El Profesorado ante el reto de las Nuevas Tecnologías en la Sociedad del Conocimiento.* Granada: Departamento de Didáctica y Organización Escolar. Universidad de Granada.

Díaz, C. y de la Fuente, I. (2006). Enseñanza del Teorema de Bayes con apoyo tecnológico. En P. Flores y J. Lupiáñez (Eds.), *Investigación en el aula de matemáticas. Estadística y Azar*. [CD ROM]. Granada: Sociedad de Educación Matemática Thales.

Díaz, C. y de la Fuente, I. (2007). Validación de un cuestionario de razonamiento probabilístico condicional. *REMA*, *12*(1).

Dubois, J. G. (1984). Une systematique des configurations combinatoires simples. *Educational Studies in Mathematics*, *15*(1), 37-57.

Dugdale, S. (2001). Pre-service teachers use of computer simulation to explore probability. *Computers in the Schools*, *17*(1/2), 173-182.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. En D. Kahneman, P. Slovic y Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Eichler, A. (2008). Germany**,** teachers' classroom practice and students' learning. En C. Batanero, G. Burrill, C. Reading y A. Rossman (2008).

Einhorn, H. J. y Hogart, R. M. (1986). Judging probable cause. *Psychological Bulletin. 99*, 3–19.

Estepa, A. (1993). *Concepciones iniciales sobre la asociación estadística y su evolución como consecuencia de una enseñanza basada en el uso de ordenadores*. Tesis Doctoral. Universidad de Granada.

Estrada, A. (2007). Actitudes hacia la estadística: un estudio con profesores de educación primaria en formación y en ejercicio. En M. Camaño, P. Flores y P. Bolea (Eds.), *Investigación en Educación Matemática XI* (pp. 121-140). Tenerife: Sociedad Española de Investigación en Educación Matemática.

Estrada, A. y Batanero, C. (2008). *Explaining teachers' attitudes towards statistics.* En C. Batanero, G. Burrill, C. Reading y A. Rossman (2008).

Estrada, A., Batanero, C. y Fortuny, J. M. (2004). Un estudio comparado de las actitudes hacia la estadística en profesores en formación y en ejercicio. *Enseñanza de las Ciencias*, *22*(2), 263-274.

Estrada, A. y Díaz, A. (2007). Errores en el cálculo de probabilidades en tablas de doble entrada en profesores en formación. *UNO*, *44*, 48-58.

Estrada, A., Díaz, C. y de la Fuente, I. (2006). Un estudio inicial de sesgos en el razonamiento sobre probabilidad condicional en alumnos universitarios. En M. Moreno, A. Estrada, J. Carrillo y T. Sierra (Eds.). *Investigación en Educación Matemática XIV* (p. 271-280). Lleida: Sociedad Española de Investigación en Educación Matemática.

Even, R. y Ball, D. (2008) (Eds.), *The professional education and development of teachers of mathematics. The 15th ICMI Study*. New York: Springer.

Falk, R. (1986). Conditional probabilities: insights and difficulties. En R. Davidson y J. Swift (Eds.), *Proceedings of the Second International Conference on Teaching Statistics.* (pp. 292 – 297). Victoria, Canada: International Statistical Institute.

Falk, R. y Konold, C. (1992). The psychology of learning probability. En F. Gordon y S.Gordon (Eds.), *Statistics for the Twenty-First Century* (pp. 151–164). Washington, DC: Mathematical Association of America.

Feller, W. (1973). *Introducción a la teoría de las probabilidades y sus aplicaciones.* México: Limusa.

Fernándes, J. A., Batanero, C., Contreras, J. M. y Díaz, C. (2009). A simulação em Probabilidades e Estatística: potencialidades e limitações. *Quadrante. XVIII* (1 y 2), 161-183.

Fiedler, K. (1988). The dependence of the conjunction fallacy on subtle lingüistic factors. *Psychological Research, 50,* 123-129.

Fischbein, E. (1987). *Intuition in science and mathematics. An educational approach*. Dordrecht: Reide.

Fox, C. R. y Levav, J. (2004). Partition–edit–count: naive extensional reasoning in judgment of conditional probability. *Journal of Experimental Psychology, 133,* 626–642.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M. y Scheaffer, R. (2005). *Guidelines for assessment and instruction in statistics education (GAISE) report: A Pre-K-12 curriculum framework*. Alexandria, VA: American Statistical Association. On line: www.amstat.org/Education/gaise.

Franklin, C. y Mewborn, D. (2006). The statistical education of PreK-12 teachers: A shared responsibility. En G. Burrill (Ed.), *NCTM 2006 Yearbook: Thinking and reasoning with data and chance* (pp. 335-344). Reston, VA: NCTM.

Galmacci, G. (2001). The impact of Internet on the researchers' training. En C. Batanero (Ed.), *Training researchers in the use of* statistics (pp. 159-169). Granada: International Statistical Institute.

Gardner, M. (1959a). Mathematical games. *Scientific American*, 180–182.

Gardner, M. (1959b). The two children problema. *Scientific American,*163-170.

Gardner, M. (1982). *¡Ajá!* Barcelona: Editorial Labor.

Garfield, J. B. y Ben-Zvi, D. (2008). Preparing school teachers to develop students' statistical reasoning. En C. Batanero, G. Burrill, C. Reading y A. Rossman (2008).

Gigerenzer, G. (1994). Why the distinction between single-event probabilities and frequencies is important for psychology (and vice-versa). En G. Wright y P. Ayton (Eds.), *Subjective probability* (pp. 129-161). Chichester: Wiley.

Gigerenzer, G. y Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats (pp. 129-161). *Psychological Review, 102*, 684-704.

Godino, J. D. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathematiques, 22* (2 y 3), 237-284.

Godino, J. D. (2003). *Teoría de las funciones semióticas. Un enfoque ontológico-semiótico de la cognición e instrucción matemática.* Departamento de Didáctica de la Matemática. Universidad de Granada. Online: www.ugr.es/local/jgodino.

Godino, J. D. (2009). Categorías de análisis de los conocimientos del profesor de matemáticas. *UNIÓN, Revista Iberoamericana de Educación Matemática, 20*, 13-31.

Godino, J. D. y Batanero, C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathematiques, 14*(3), 325-355.

Godino, J. D. y Batanero, C. (1998a). Clarifying the meaning of mathematical objects as a priority area of research in Mathematics Education. En A. Sierpinska y J. Kilpatrick (Eds.). *Mathematics education a research domain: A search for identity* (pp. 177-195). Dordrecht: Kluwer.

Godino, J. D. y Batanero, C. (1998b). Funciones semióticas en la enseñanza y aprendizaje de las matemáticas. En I. Vale y J. Portela (Eds.). *IX Seminário de Investigaçao em Educaçao Matemática* (pp. 25-45). Associaçao de Profesores de Matemática. Portugal.

Godino, J. D., Batanero, C. y Cañizares, M. J. (1987). *Azar y probabilidad. Fundamentos didácticos y propuestas curriculares*. Madrid: Síntesis.

Godino, J. D., Batanero**,** C. y Flores, P. (1999). El análisis didáctico del contenido matemático como recurso en la formación de profesores. En *Homenaje al profesor Oscar Sáenz Barrio* (pp. 165-185). Granada: Departamento de Didáctica y Organización Escolar.

Godino, J. D., Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education *ZDM, The International Journal on Mathematics Education, 39*(1-2) 127-135.

Godino, J. D., Batanero, C. y Roa, R. (2005). An onto-semiotic analysis of combinatorial problems and the solving processes by university students. *Educational Studies in Mathematics, 60* (1), 3-36.

Godino, J. D., Batanero, C., Roa, R., y Wilhelmi, M. R. (2008). Assessing and developing pedagogical content and statistical knowledge of primary school teachers through project work. En C. Batanero, G. Burrill, C. Reading y A. Rossman (2008).

Godino, J. D., Contreras, A. y Font, V. (2006). Análisis de procesos de instrucción basado en el enfoque ontológico-semiótico de la cognición matemática. *Recherches en Didactiques des Mathematiques*, *26* (1), 39-88.

Godino, J. D., Font, V. y Wilhelmi, M. R. (2008). Análisis didáctico de procesos de estudio matemático basado en el enfoque ontosemiótico. *Publicaciones, 38,* 25-48.

Godino, J., Wilhelmi, M. y Bencomo, D. (2005). Suitability criteria of a mathematical instruction process. A teaching experience of the function notion. *Mediterranean Journal for Research in Mathematics Education*, *4*(2), 1-26.

Godino, J. D., Wilhelmi, M. y Font, V. (2008). Análisis didáctico de procesos de estudio matemático basado en el enfoque ontosemiótico. *Publicaciones, 38,* 25-49.

Goldin, G. A. y McClintock, C. E. (Eds.) (1984). *Task variables in mathematical problem solving.* Philadelphia, PA: The Franklin Institute Press.

González, P. (2004). La historia de la matemática como recurso didáctico e instrumento para enriquecer culturalmente su enseñanza. *Suma*, *45*, 17-28.

Gotzsche, P. y Olsen, O. (2000). Is screening for breast cancer with mammography justifiable? *Lancet, 355,* 129-34.

Gras, R. y Totohasina, A. (1995). Chronologie et causalité, conceptions sources d’obstacles épistémologiques à la notion de probabilité conditionnelle *Recherches en Didactique des Mathématiques*, *15*(1), 49-95.

Green, D. R. (1983). A survey of probabilistic concepts in 3000 pupils aged 11-16 years. En D. R. Grey et al. (Eds.), *Proceedings of the First International Conference on Teaching Statistics* (v.2, pp. 766-783). Universidad de Sheffield.

Groth, R. E. (2008). Navigating layers of uncertainty in teaching statistics through case discussion. En C. Batanero, G. Burrill, C. Reading, A. Rossman (2008).

Gutiérrez R., Martínez A. y Rodríguez C. (1993). *Curso básico de probabilidad*. Madrid Ediciones Pirámide.

Haller, S. K. (1997). *Adopting probability curricula: The content and pedagogical content knowledge of middle grades teachers.* Tesis Doctoral. University of Minnesota.

Hardin, G. (1968). The tragedy of commons. *Science, 162*, 1243-1248.

Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics, 6,* 187–205.

Hill, H., Ball, D. L., y Schilling, S. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers topic-specific knowledge of students, *Journal for Research in Mathematics Education*, *39*, 372-400.

Hill, H. C., Sleep, L., Lewis, J. M. y Ball, D. (2007). Assessing teachers’ mathematical knowledge. En F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 111-155). Greenwich, CT: Information Age Publishing y NCTM.

Huerta, M. P. y Lonjedo, M. A. (2003). La resolución de problemas de probabilidad condicional. Trabajo presentado en el *Séptimo Simposio de la Sociedad Española de Investigación en Educación Matemática*. Granada.

Hwei, P. y Hsu, P. D. (2003). *Theory and problems of probability, random variables, and random processes.* New York: Schaum.

Jaworski, B. (2001). Developing mathematics teaching: teachers, teacher educators and researchers as co-learners. En L. Lin y T. J. Cooney (Eds.), *Making sense of mathematics teacher education* (pp. 295-320). Dordrecht: Kluwer.

Jaworski, B. y Gellert, U. (2003). Educating new mathematics teachers: Integrating theory and practice, and the roles of practising teachers. En A. J. Bishop, M. A. Clements, C. Keitel, J. Kilpatrick y F. K. S. Leung (Eds.), *Second international handbook of mathematics education* (pp. 795-828). Dordrech: Kluwer.

Jones, G., Langrall, C. y Mooney, E. (2007). Research in probability: responding to classroom realities. En F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning.* Greenwich, CT: Information Age Publishing y NCTM.

Kahneman, D., Slovic, P., y Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* New York: Cambridge University Press.

Kelly, I. W. y Zwiers, F. W. (1986). Mutually exclusive and independence: Unravelling basic misconceptions in probability theory. *Teaching Statistics 8*, 96-100.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction, 6*, 59-98.

Konold, C. (1994). Teaching probability through modeling real problems. *The Mathematics Teacher, 87*(4), 232-235.

Kraitchik, M. (1953). *Mathematical recreations.* New York: Dover.

Kvatinsky, T. y Even, R. (2002). Framework for teacher knowledge and understanding of probability. En B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics* [CD-ROM]. Voorburg, Netherlands: International Statistical Institute.

Lecoutre, M. P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics, 23*, 557-568.

Lecoutre, M. P. y Cordier, J. (1990). Effet du mode de présentation d'un problème aleatoire sur les modèles développés par les élèves. *Bulletin de l'APMEP*, *372*, 9-22.

Lecoutre, M. P. y Durand, J. L. (1988). Jugéments probabilistes et modéles cognitifs: étude d'une situation aléatoire. *Educational Studies in Mathematics, 19*, 357-368.

Lee, H. S. y Hollebrands, K. (2008). Preparing to teach data analysis and probability with technology. En C. Batanero, G. Burrill, C. Reading y A. Rossman (2008).

León, N. (2009) La historia como elemento motivador hacia el estudio de la probabilidad: el problema de la apuesta interrumpida, *Sapiens: Revista Universitaria de Investigación, 1*, 69-88.

Lesser, L. (1998). Countering indifference – Using counterintuitive examples. *Teaching Statistics*, *20*(1), 10-12.

Lester, F. K. (1983). Trends and issues in mathematical problem solving research. En R. Lesh y M. Landau (Eds.), *Acquisitions of mathematics concepts and processes.* London: Academy Press.

Lipschutz, S. y Lars, M. (2001). *Schaum, Teoría y problemas de probabilidad*. Madrid: Macgraw-Hill.

Llinares, S. y Krainer, K. (2006). Mathematics (student) teachers and teacher educators as learners. En A. Gutierrez y P. Boero (Eds) *Handbook of research on the psychology of mathematics education* (pp. 429 – 459). Rotherdam / Taipei: Sense Publishers.

Loeve, M. (1976). Teoría de la probabilidad. Madrid: Editorial Tecnos.

Lonjedo, M. A. (2003). *La resolución de problemas de probabilidad condicional. Un estudio exploratorio con estudiantes de bachiller*. Universidad de València. Memoria de tercer ciclo.

Lonjedo, M. A. y Huerta M. P. (2005). La naturaleza de las cantidades presentes en el problema de probabilidad condicional. Su influencia en el proceso de resolución del problema. En A. Maz, B. Gómez y M. Torralbo (Eds.), *Investigación en Educación Matemática IX. Noveno Simposio de la Sociedad Española de Investigación en Educación Matemática* (pp. 203-212). Córdoba: SEIEM.

Lonjedo, M. A. y Huerta M. P. (2007). Análisis del comportamiento de los estudiantes en la resolución de problemas isomorfos de probabilidad condicional. En M. Camacho, P. Flores y P. Bolea (Eds). *Investigación en educación matemática XI. Once Simposio de la Sociedad Española de Investigación en Educación Matemática* (pp 273-282). Badajoz: SEIEM.

López, C. (2006). Stochastics and the professional knowledge of teachers. En A. Rossman y B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. [CD-ROM] Salvador (Bahía) Brasil: International Statistical Institute.

Martin-Pliego, F. J. y Ruiz-Maya, L. (2006), *Fundamentos de probabilidad*. Internacional Thomson Editores.

Martínez-Arias, R. (1995). *Psicometría: Teoría de los tests psicológicos y educativos*. Madrid. Síntesis.

Martignon, L. y Wassner, C. (2002). Teaching decision making and statistical thinking with natural frequencies. En B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics* [CD-ROM].Ciudad del Cabo: IASE.

Maury, S. (1985). Influence de la question dans une épreuve relative á la notion d'independance. *Educational Studies in Mathematics*, *16*, 283-301.

Maury, S. (1986). *Contribution à l'étude didactique de quelques notions de probabilité et de combinatoire à travers la résolution de problémes.* Tesis doctoral. Universidad de Montpéllier II.

Mayén, S. (2009). *Significados de las medidas de posición central para estudiantes mexicanos de Bachillerato.* Tesis doctoral. Universidad de Granada.

MEC (2006). *Real Decreto 1513/2006, de 7 de diciembre, por el que se establecen las enseñanzas mínimas de la Educación Primaria.*

MEC (2007a). *Real Decreto 1631/2006, de 29 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación Secundaria Obligatoria.*

MEC (2007b). *Real Decreto 1467/2007, de 2 de noviembre, por el que se establece la estructura del bachillerato y se fijan sus enseñanzas mínimas.*

Mills, J. D. (2002). Using computer simulation methods to teach statistics: a review of the literature. *Journal of Statistics Education 10*(1).

Montgomery, D. C. y Runger, G. C. (2002). *Applied statistics and probability for engineers*. New York: John Wiley.

Morales, P. (1988). *Medición de actitudes en psicología y educación*. San Sebastián: Universidad de Comillas.

Muñiz, J. (1994). *Teoría clásica de los tests.* Madrid: Pirámide.

Navarro-Pelayo, V. (1994). *Estructura de los problemas combinatorios simples y del razonamiento combinatorio en alumnos de secundaria*. Tesis Doctoral. Universidad de Granada.

Nalebuff, B. (1989). Puzzles: the other person's envelope is always greener, *Journal of Economic Perspectives*, *3*, 171-181.

NCTM (2000). *Principles and standards for school mathematics.* Reston, VA; NCTM. Online: standards.nctm.org/.

Nisbett, R. y Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgments*. Englewood Cliffs, NJ: Prentice Hall.

Nortes, A. (1993). *Estadística teórica y aplicada*. Barcelona: PPU.

Ojeda, A. M. (1995). Dificultades del alumnado respecto a la probabilidad condicional. *UNO,* 5, 37-55.

Olivo, E. (2008). *Significados de los intervalos de confianza para los estudiantes de ingeniería en México.* Tesis doctoral. Universidad de Granada.

Ortiz, J. J. (1999). *Significados de los conceptos probabilísticos en los libros de texto de Bachillerato.* Tesis doctoral. Universidad de Granada.

Ortiz, J. J., Nordin, M., Batanero, C., Serrano, L. y Rodríguez, J. (2006). Comparación de probabilidades en maestros en formación. En Bolea, P., González, M. J. y Moreno, N. (Eds.), *Investigación en Educación Matemática. Actas del X Simposio de la Sociedad Española de Investigación en Educación Matemática* (pp. 268-276). Huesca: SEIEM.

Peña, D. (1986). *Estadística. Modelos y métodos 1. Fundamentos.* Madrid: Alianza Editorial.

Penalva, C. y Posadas, J. A. (2009). El planteamiento de problemas y la construcción del teorema de Bayes. *Enseñanza de las Ciencias, 27* (3), 331-342.

Pérez-Echeverría, M. P. (1990). *Psicología del razonamiento probabilístico.* Madrid: Ediciones de la Universidad Autónoma de Madrid.

Pereira-Mendoza, L. (2002). Would you allow your accountant to perform surgery? Implications for the education of primary teachers. En B. Phillips (Ed.), *Proceedings of the Sixth International Conference on the Teaching of Statistics*.

Pesci, A. (1994). Tree graphs: use as an aid in casual compounds events. En J. P. Ponte y J. F. Matos (Eds.), *Proceedings of the PME XVIII*, (v.4, pp. 25-32). Lisboa: Departamento de Educaçao. Universidade de Lisboa.

Phillipp, R. A. (2007). Mathematics teachers' beliefs and affects. En F. Lester (Ed.). *Second handbook of research on mathematics teaching and learning* (pp. 257-315). Charlotte, NC: Information Age Publishing y NCTM.

Ponte, J. P. (2008). Investigar a nossa própria prática: uma estratégia de formacão e de construcão do conhecimento profissional. *PNA, 2(4)*, 153-180 .

Ponte, J. P. y Chapman, O. (2006). Mathematics teachers' knowledge and practices. En A. Gutierrez y P. Boero (Eds.), *Handbook of reaserch on the psychology of mathematics education: Past, present and future* (pp. 461-494). Roterdham: Sense.

Posadas, J. A. (2009). *Estudio de la comprensión de contenidos de probabilidad de estudiantes universitarios*. Tesis doctoral. Universidad de Alicante.

Pollatsek, A., Well, A. D., Konold, C. y Hardiman, P. (1987). Understanding conditional probabilities. *Organitation, Behavior and Human Decision Processes, 40*, 255–269.

Poularikas, A. D. (1999), *Probability and stochastic processes*, *The handbook of formulas and tables for signal processing.* New York: Wiley.

Puig, L. (2008). Presencia y ausencia de la resolución de problemas en la investigación y en el currículo. En R. Luengo, B. Gómez, M. Camacho y L. Blanco (Eds.), *Investigación en Educación Matemática XII*. Badajoz: Sociedad Española de Investigación en Educación Matemática.

Roa, R. (2000). *Razonamiento combinatorio en alumnos con preparación matemática avanzada.* Tesis Doctoral. Universidad de Granada.

Russell, S. (1990). Issues in training teachers to teach statistics in the elementary school: A world of uncertainty En A. Hawkins (Ed.), *Training teachers to teach statistics Proceedings of the International Statistical Institute Round Table Conference* (pp. 59- 71). Voorburg, Netherlands: International Statistical Institute.

Sáenz, C. (1999). Teaching probability for conceptual change. *Educational Studies in Mathematics, 35*, 233-254.

Sáenz, C. (2007). La competencia matemática (en el sentido de PISA) de los futuros maestros. *Enseñanza de las Ciencias, 25*(3), 355–366.

Sánchez, E. (1996). Dificultades en la comprensión del concepto de eventos independientes. En F. Hitt (Ed.), *Investigaciones en Educación Matemática* (pp. 389-404). México.

Ruiz, B. (2006). *Un acercamiento cognitivo y epistemológico a la didáctica del* *concepto de variable aleatoria*. Tesis de Maestría. CICATA. México.

Schnabel, Z. M. (1938). The estimation of the total fish population of a lake, *American Mathematical Monthly, 45,* 348–352.

Schoenfeld, A. (1985). *Mathematical problem solving.* New York: Academic Press.

Sedlmeier, P. (1999). *Improving statistical reasoning. Theoretical models and practical implications.* Mahwah, NJ: Erlbaum.

Serradó, A., Azcárate, P. y Cardeñoso, J. M. (2006). Analyzing teacher resistance to teaching probability in compulsory education. En A. Rossman y B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. Salvador, Bahia, Brazil: International Statistical Institute e International Association for Statistical Education. Online: www.stat.auckland.ac.nz/~iase/publications.

Selvin, S. (1975). A problem in probability. *American Statistician, 29*(1), 67-80.

Serrano, L. (1996). *Significados institucionales y personales de objetos matemáticos ligados a la aproximación frecuencial de la enseñanza de la probabilidad.* Tesis doctoral. Universidad de Granada.

Serrano, L., Batanero, C., Ortiz, J. J. y Cañizares, M. J. (2001). Concepciones de los alumnos de secundaria sobre modelos probabilísticos en las secuencias de resultados aleatorios. *Suma, 36,* 23-32.

Shulman (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15* (2), 4-14.

Shulman, L. S. (1987). Knowledge and teaching: foundations of new reform. *Harvard Educational Review, 57* (1), 1-22.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables, *Journal of the Royal Statistical Society* (Series B), *13,* 238–241.

Sowder, J. T. (2007). The mathematical education and development of teachers. En F. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 157-224). Charlotte, N.C: Information Age Publishing y NCTM.

Spiegel, M. R. (2000). *Estadística.* Mc Graw Hill.

Steinbring, H. (1990). The nature of stochastical knowledge and the traditional mathematics curriculum. Some experiences with in-service training and developing materials. En A. Hawkins (Ed.). *Training teachers to teach statistics* (pp. 2-19). Voorburg: International Statistical Institute.

Stohl, H. (2005). Probability in teacher education and development. In G. Jones (Ed.).

Exploring probability in schools: Challenges for teaching and learning (345-366). New York: Springer.

Székely, G. J. (1986). *Paradoxes in probability theory and mathematical statistics.* Dordrecht: Reidel.

Tarr, J. E. y Jones, G. A. (1997). A framework for assessing middle school students' thinking in conditional probability and independence. *Mathematics Education Research Journal*, 9, 39-59.

Tarr, J. E. y Lannin, J. K. (2005). How can teachers build notions of conditional probability and independence? En G. Jones (Ed.), *Exploring probability in school. Challenges for teaching and learning* (pp. 215-238). New York: Springer.

Teigen, K. H., Brun, W. y Frydenlund, R. (1999). Judgments of risk and probability: the role of frequentist information. *Journal of Behavioral Decision Making*, *12(2)*, 123.

Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. En D. A. Grouws (Ed.), *Handbook on mathematics teaching and learning* (pp. 127-146). New York: Macmillan.

Totohasina, A. (1992). *Méthode implicative en analyse de données et application á l'analyse de conceptions d'étudiants sur la notion de probabilité conditionnelle.* Tesis Doctoral. Universidad Rennes I.

Truran, J. M. y Truran, K. M. (1997). Statistical independence: One concept or two? En B. Phillips (Ed.), *Papers from Statistical Education Presented at ICME 8* (pp. 87–100). Swinburne University of Technology.

Tversky, A. y Kahneman, D. (1982a). Causal schemas in judgment under uncertainty. En D. Kahneman, P. Slovic y A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases* (pp. 117-128). Cambridge, MA: Cambridge University Press.

Tversky, A. y Kahneman, D. (1982b). On the psychology of prediction. En D. Kahneman, P. Slovic y A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases* (pp. 69-83). Cambridge, MA: Cambridge University Press.

Vallecillos, A (1994). *Estudio teórico experimental de errores y concepciones sobre el contraste de hipótesis en estudiantes universitarios.* Tesis doctoral. Universidad de Granada.

Watson, J. M. (2001). Profiling teachers' competence and confi dence to teach particular mathematics topics: The case of data and chance. *Journal of Mathematics Teacher Education*, *4*, 305-337.

Wild, C. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 221-248.

Wood, T. (Ed.) (2008). *The international handbook of mathematics teacher education*. Rotterdam: Sense Publishers.

ANEXOS

ANEXO 1

TAREA DE EVALUACIÓN PROPUESTA EN EL ESTUDIO 1

La lectura de tablas de doble entrada es uno de los contenidos incluidos en Educación Primaria.

a. Resuelve el siguiente ejercicio.
b. ¿Qué contenidos matemáticos se trabaja? (clasifícalos según los objetos matemáticos estudiados en clase: problemas, lenguaje, conceptos, propiedades, procedimientos, argumentos).
c. Indica algunas variables que se pueda cambiar en esta tarea para variar la dificultad o para variar el contenido matemático.

Tarea. En un colegio se pregunta a los alumnos, obteniendo los siguientes resultados:

	Chicos	Chicas	Total
Le gusta el tenis	400	200	600
No le gusta	50	50	100
Total	450	250	700

Si elegimos al azar uno de estos alumnos:

a. ¿Cuál es la probabilidad de que le guste el tenis?
b. ¿Cuál es la probabilidad de que sea una chica y además le guste el tenis?
c. Sabiendo que el alumno elegido es chica, ¿Cuál es la probabilidad de que le guste el tenis?

ANEXO 2

CUESTIONARIO DE RAZONAMIENTO SOBRE PROBABILIDAD CONDICIONAL UTILIZADO EN EL ESTUDIO 2 (RPC)

CUESTIONARIO SOBRE PROBABILIDAD CONDICIONAL

Nombre: ____

Universidad: _________________ ___________

Curso que sigues:

1. Máster de secundaria_______ 2. Licenciatura:______

Indica cuál es tu formación básica:

1. Licenciado en Matemáticas_______2. Otro ¿Cuál?________________

El cuestionario que te presentamos es parte de una investigación orientada a mejorar la formación del profesor para enseñar probabilidad. Agradecemos sinceramente tu colaboración

---

Parte 1.

Lee atentamente los enunciados y marca la respuesta que consideres correcta.

Ítem 1. Un taxi se vio implicado en un accidente nocturno. Hay dos compañías de taxis en la ciudad, la Verde y la Azul. El 85% de los taxis de la ciudad son Verdes y el 15% Azules. Un testigo identificó el taxi como Azul. El tribunal comprobó la fiabilidad del testigo bajo las mismas circunstancias que había la noche del accidente y llegó a la conclusión de que el testigo identificaba correctamente cada uno de los colores en el 80% de las ocasiones y fallaba en el 20%. ¿Cuál es la probabilidad de que el taxi implicado en el accidente fuera en efecto Azul?

- ☐ a. 80 %
- ☐ b. 15%
- ☐ c. (15/100) x (80/100)
- ☐ d. 41 %

Ítem 2. Se extrae una carta al azar de una baraja española (40 cartas con números del 1 al 7, sota caballo y rey). Sea A el suceso "se extrae una carta de oros " y B el suceso "se extrae un rey" ¿Los sucesos A y B son independientes?

- ☐ a. No son independientes porque en la baraja hay un rey de oros.
- ☐ b. Sólo si sacamos primero una carta para ver si es rey y se vuelve a colocar en la baraja y luego sacamos una segunda carta para mirar si es oros.
- ☐ c. No, porque P(rey de oros)= P(rey) x P(oros)
- ☐ d. Sí, en todos los casos

Ítem 3. La probabilidad de que una mujer de 40 años tenga una mamografía positiva es el 10,3%. La probabilidad de que una mujer de 40 años tenga cáncer de pecho y una mamografía positiva es de 0,8%. Una mujer de 40 años se realiza una mamografía y resulta positiva. ¿Cuál es la probabilidad de que realmente tenga cáncer?

☐ a. $\frac{0,8}{10,3} = 0,0776$, probabilidad de 7'76%
☐ b. $10,3 \times 0,8 = 8,24$, probabilidad del 8'24%
☐ c. 0,8 %

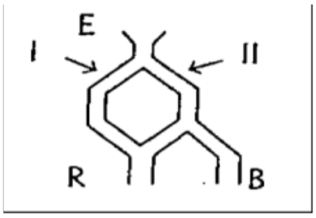
Ítem 4. Supón que Rafa Nadal alcanza la final de Roland Garros en 2010. Para ganar el partido hay que ganar tres sets. ¿Cuál de los siguientes sucesos consideras más probable?

☐ a. Rafa Nadal pierde el primer set.
☐ b. Rafa Nadal pierde el primer set pero gana el partido.
☐ c. Los dos sucesos son iguales de probables.

Ítem 5. Un test diagnóstico de cáncer fue administrado a todos los residentes de una gran ciudad. Un resultado positivo en el test es indicativo de cáncer y un resultado negativo es indicativo de ausencia de cáncer. ¿En cuál de las siguientes predicciones tienes más confianza?

☐ a. Predecir que una persona tiene cáncer si ha dado positivo en el test de diagnóstico.
☐ b. Predecir un resultado positivo en el test de diagnóstico si la persona tiene cáncer.
☐ c. Tengo la misma confianza en ambas predicciones.

Ítem 6. Una bola se suelta por la entrada E. Si sale por R, ¿Cuál es la probabilidad de que haya pasado por el canal 1?



☐ a. 0,50
☐ b. 0,33
☐ c. 0,66
☐ d. No se puede calcular

Ítem 7. Una urna contiene dos bolas blancas y dos bolas negras. Extraemos a ciegas dos bolas de la urna, una detrás de otra, sin reemplazamiento. ¿Cuál es la probabilidad de extraer una bola negra en segundo lugar, habiendo extraído una bola negra en primer lugar? P (N_2/ N_1).

☐ a. 1/ 2
☐ b. 1/ 6
☐ c. 1/ 3
☐ d. 1/ 4

¿Cuál es la probabilidad de extraer una bola negra en primer lugar, sabiendo que hemos extraído una bola negra en segundo lugar? P (N_1/ N_2).

☐ a. 1/ 3
☐ b. 1/4
☐ c. 1/ 6
☐ d. 1/ 2

Parte 2.

Lee atentamente los enunciados y responde las preguntas o resuelve los problemas. Escribe los pasos que sigas en la solución

Ítem 8. Explica con tus propias palabras la diferencia entre una probabilidad simple y una probabilidad condicional.

Ítem 9. Hemos lanzado dos dados y sabemos que el producto de los dos números obtenidos ha sido 12 ¿Cuál es la probabilidad de que ninguno de los dos números sea un 6?

Ítem 10. En una ciudad hay 60 hombres y 40 mujeres por cada 100 habitantes. 50 de cada 100 hombres y 25 de cada 100 mujeres fuman. Si cogemos 200 personas al azar, ¿cuántas de estas 200 personas fumarán?

Ítem 11. Una persona lanza un dado y anota si saca un número par o impar. Estos son los resultados al lanzarlo 15 veces:

"par, impar, impar, par, par, impar, par, par, par, par, impar, impar, par, par, par"

Lanza el dado de nuevo ¿Cuál es la probabilidad de sacar un número par en esta nueva tirada? ¿Por qué?

Ítem 12. En una encuesta publicada en un periódico se informa que el 91% de los habitantes de una ciudad mienten usualmente y de ellos el 36% mienten sobre cosas importantes. Si elegimos al azar a una persona de esta ciudad, ¿cuál es la probabilidad de que mienta sobre cosas importantes?

Ítem 13. Una fábrica dispone de dos máquinas M1 y M2 que fabrican bolas. La máquina M1 fabrica el 40 % de las bolas y la M2 el 60%. El 5% de las bolas fabricadas por M1 y el 1% de las fabricadas por M2 son defectuosas. Tomamos una bola al azar que resulta ser defectuosa. ¿Cuál es la probabilidad de que haya sido fabricada por M1?

Ítem 14. Un grupo de alumnos de un colegio se examina de Matemáticas e inglés. La proporción de alumnos que aprueban Matemáticas es del 80% y la proporción de alumnos que aprueba inglés es del 70%. Suponiendo que las notas en cada una de las asignaturas es independiente, ¿Cuál es la probabilidad de que un alumno escogido al azar haya aprobado ambas asignaturas?

ANEXO 3.
TALLER SOBRE IDEAS ESTOCÁSTICAS FUNDAMENTALES
Estudio 5

Nombre: País

Formación previa:

Años de docencia en Secundaria o Primaria En Universidad

¿Has explicado alguna vez estadística?

¿Ha estudiado algún curso de estadística? ¿Cuántos?

Decripción del juego

Se toman 3 fichas de la misma forma y tamaño, de las cuales una es roja por ambas caras; otra es azul por una cara y roja por la otra, y la tercera es azul por las dos caras.

El profesor coloca las tres fichas en una caja, que agita convenientemente, antes de seleccionar una de las tres fichas, al azar. Muestra, a continuación, una de las caras de la ficha elegida, manteniendo la otra tapada, pidiéndo a sus alumnos que adivinen el color de la cara oculta.

Una vez hechas las apuestas, el profesor muestra la cara oculta. Cada alumno que haya acertado en la predicción efectuada, consigue un punto.

Figura 1. Hoja de registro

Ensayo n.	1	2	3	4	5	6	7	8	9	10
Color de la cara mostrada										
Color predicho										
Color de la cara oculta										

1. ¿Has seguido alguna estrategia? Descríbela

2. ¿Por qué sigues esta estrategia o por qué no sigues ninguna? Da un razonamiento

Figura 2. Hoja de registro. Continuación del juego

Ensayo n.	1	2	3	4	5	6	7	8	9	10
Color de la cara mostrada										
Color predicho										
Color de la cara oculta										

¿Cuál es ahora tu estrategia?

¿Por qué?

Si estás seguro/a de tu estrategia, da una demostración matemática de la misma

Para los que no están seguros, haremos una tercera ronda del juego

Figura 3. Hoja de registro

Ensayo n.	1	2	3	4	5	6	7	8	9	10
Color de la cara mostrada										
Color predicho										
Color de la cara oculta										

Si estás seguro/a de tu estrategia, da una demostración matemática de la misma

ANÁLISIS DIDÁCTICO

Primer nivel de análisis: Completa la siguiente tabla de objetos que has utilizado en esta situación

Tabla 1

Tipos	Objetos matemáticos	Significado en la situación
SITUACIONES PROBLEMAS		
LENGUAJE		
CONCEPTOS		
PROCEDIMIENTOS		
PROPIEDADES		
ARGUMENTOS		

Segundo nivel de análisis: Lista algunos posibles conflictos y dificultades por parte de los estudiantes

Tercer nivel de análisis: configuraciones didácticas. ¿Qué diferentes fases didácticas puedes identificar en toda la actividad, teniendo en cuenta la forma de trabajo e interacción entre alumnos y del alumno con el profesor?

ANEXO 4.

CONTENIDOS SOBRE ESTADÍSTICA Y PROBABILIDAD EN LOS DECRETOS DE ENSEÑANZAS MÍNIMAS

1. EDUCACIÓN PRIMARIA

La probabilidad se contempla en la Educación Primaria, tanto por parte del Ministerio de Educación, como de la Consejería de Educación de la Junta de Andalucía.

El Ministerio (MEC, 2006a) se incluye este contenido dentro del bloque 4, *Tratamiento de la información, azar y probabilidad*. Dicho bloque "adquiere su pleno significado cuando se presentan conectados con problemas que implican a otras áreas de conocimiento. El trabajo en este Bloque ha de incidir de forma significativa en la comprensión de las informaciones presentes en los medios de comunicación. Debe también aumentar el interés por la materia y ayudar a valorar el beneficio que los conocimientos estadísticos proporcionan ante la toma de decisiones en muchas y variadas situaciones" (p. 43096). Asimismo, en la citada página se sugiere la importancia dentro del bloque los contenidos actitudinales, por ejemplo al promover la presentación de los datos de forma ordenada y gráfica, y descubrir que las matemáticas facilitan la resolución de problemas de la vida diaria, así como iniciar en el uso crítico de la información recibida por diferentes medios.

El Ministerio (MEC, 2006a) incluye los siguientes contenidos específicos de estadística y probabilidad en este bloque:

- *Primer Ciclo* (p. 43098):

- "Gráficos estadísticos: Descripción verbal, obtención de información cualitativa e interpretación de elementos significativos de gráficos sencillos relativos a fenómenos cercanos. Utilización de técnicas elementales para la recogida y ordenación de datos en contextos familiares y cercanos.
- Carácter aleatorio de algunas experiencias. Distinción entre lo imposible, lo seguro y aquello que es posible pero no seguro, y utilización en el lenguaje habitual, de expresiones relacionadas con la probabilidad.
- Participación y colaboración activa en el trabajo en equipo y el aprendizaje organizado a partir de la investigación sobre situaciones reales. Respeto por el trabajo de los demás".

- *Segundo Ciclo* (p. 43099)*:*

- "Gráficos y tablas: Tablas de datos. Iniciación al uso de estrategias eficaces de recuento de datos. Recogida y registro de datos sobre objetos, fenómenos y situaciones familiares utilizando técnicas elementales de encuesta, observación y medición. Lectura e interpretación de tablas de doble entrada de uso habitual en la vida cotidiana. Interpretación y descripción verbal de elementos significativos de gráficos sencillos relativos a fenómenos familiares. Disposición a la elaboración y presentación de gráficos y tablas de forma ordenada y clara.
- Carácter aleatorio de algunas experiencias, Valoración de los resultados de experiencias en las que interviene el azar, para apreciar que hay sucesos más o menos probables y la imposibilidad de predecir un resultado concreto. Introducción

al lenguaje del azar. Confianza en las propias posibilidades, y curiosidad, interés y constancia en la interpretación de datos presentados de forma gráfica".

- *Tercer Ciclo* (p. 43101)*:*

- "Gráficos y parámetros estadísticos: Recogida y registro de datos utilizando técnicas elementales de encuesta, observación y medición. Distintas formas de representar la información. Tipos de gráficos estadísticos. Valoración de la importancia de analizar críticamente las informaciones que se presentan a través de gráficos estadísticos.
- La media aritmética, la moda y el rango, aplicación a situaciones familiares. Disposición a la elaboración y presentación de gráficos y tablas de forma ordenada y clara. Obtención y utilización de información para la realización de gráficos.
- Carácter aleatorio de algunas experiencias.Presencia del azar en la vida cotidiana. Estimación del grado de probabilidad de un suceso.
- Valoración de la necesidad de reflexión, razonamiento y perseverancia para superar las dificultades implícitas en la resolución de problemas. Confianza en las propias posibilidades e interés por utilizar las herramientas tecnológicas en la comprensión de los contenidos funcionales".

Entre otras orientaciones curriculares, se señala que los juegos de azar proporcionan ejemplos que permitirán introducir las nociones de probabilidad e incertidumbre. La evaluación considerará además de los aspectos propios de la clasificación y representación de datos, la capacidad para deducir relaciones entre ellos y, sobre todo, la deducción de conclusiones y estimaciones a partir de los datos representados. Se pretende que el alumnado sea capaz de razonar sobre los posibles resultados de un experimento aleatorio sencillo y saber asignar probabilidades a sucesos equiprobables y no equiprobables, con estrategias tales como recuento. También se sugiere usar la asignación frecuencial, a partir de experimentos organizados en la clase, que permiten enlazar estadística y probabilidad.

Encontramos también en este documento los siguientes criterios de evaluación, relacionados con el tema:

- Primer Ciclo: "Realizar interpretaciones elementales de los datos presentados en gráficas de barras. Formular y resolver sencillos problemas en los que intervenga la lectura de gráficos. Se trata de valorar la capacidad de interpretar gráficos sencillos de situaciones familiares y verificar la habilidad para reconocer gráficamente informaciones cuantificables. También se pretende evaluar si los niños y las niñas están familiarizados con conceptos y términos básicos sobre el azar: seguro, posible, imposible" (p. 43098).

- *Segundo Ciclo:* "Recoger datos sobre hechos y objetos de la vida cotidiana utilizando técnicas sencillas de recuento, ordenar estos datos atendiendo a un criterio de clasificación y expresar el resultado de forma de tabla o gráfica. Este criterio trata de valorar la capacidad para realizar un efectivo recuento de datos y representar el resultado utilizando los gráficos estadísticos más adecuados a la situación. Es asimismo motivo de evaluación la capacidad para describir e interpretar gráficos sencillos relativos a situaciones familiares" (p. 43100).

- *Tercer Ciclo:* "Realizar, leer e interpretar representaciones gráficas de un conjunto de datos relativos al entorno inmediato. Hacer estimaciones basadas en la experiencia sobre el resultado (posible, imposible, seguro, más o menos probable) de situaciones sencillas en las que intervenga el azar y comprobar dicho resultado. Se evalúa la capacidad de recoger y registrar una información que se pueda cuantificar, utilizar algunos recursos sencillos de representación gráfica: tablas de datos, bloques de barras, diagramas lineales... y comprender y comunicar la información así expresada. Además, se comprobará que se empieza a constatar que hay sucesos imposibles, sucesos que con casi toda seguridad se producen, o que se repiten, siendo más o menos probable esta repetición".

La Consejería de Educación (2007a) remite a estos mismos contenidos cuando describe el Bloque 5 que incluye en el currículo de matemáticas con el título Tratamiento de la información, estadística y azar. En este documento se resaltan las conexiones de este bloque con los siguientes del Real Decreto 1513/2006, de 7 de diciembre: Bloque 1, Números y operaciones y Bloque 2, La medida: estimación y cálculo de magnitudes Bloque 3, Geometría. También se sugieren que sus contenidos sólo adquieren su pleno significado cuando se presentan en conexión con actividades que implican a otras áreas de conocimiento.

El documento indica que el trabajo ha de incidir de forma significativa en la comprensión de las informaciones de los medios de comunicación, y ayudar a valorar el beneficio que los conocimientos estadísticos proporcionan ante la toma de decisiones. Se da especial importancia a los contenidos actitudinales.

La principal finalidad de este núcleo temático es que las niñas y niños comiencen a interpretar los fenómenos ambientales y sociales de su entorno cercano a través de las matemáticas. Los alumnos y alumnas deben ser conscientes de los fenómenos de distinta naturaleza que suceden a su alrededor y que aparecen en los medios de comunicación. Esto ayuda a entender las matemáticas como una disciplina que ayuda a interpretar la realidad y a actuar sobre ella de forma responsable, crítica y positiva.

Se recuerda que los contenidos matemáticos implicados en este núcleo corresponden fundamentalmente a la estadística y a la probabilidad, disciplinas matemáticas entre las que existe una relación complementaria. Las múltiples aplicaciones de dichas disciplinas se extienden a todos los campos de la actividad humana. Ello ocasiona un amplio reconocimiento social, constatado por su creciente presencia en el aprendizaje de otras materias, el mercado laboral y el ambiente cultural. También se añade que los juegos de azar proporcionan ejemplos que permitirán introducir las nociones de probabilidad e incertidumbre.

Respecto a los criterios de evaluación repite los citados del Ministerio de Educación.

2. EDUCACIÓN SECUNDARIA OBLIGATORIA

Respecto a la Enseñanza Secundaria Obligatoria, el Decreto de Enseñanzas Mínimas (MEC, 2006b) indica (p. 750) que "la toma de decisiones requiere comprender, modificar y producir mensajes de todo tipo, y en la información que se maneja cada vez aparecen con más frecuencia tablas, gráficos y fórmulas que demandan conocimientos matemáticos para su correcta interpretación. Por ello, los ciudadanos deben estar preparados para adaptarse con eficacia a los continuos cambios que se generan".

Asimismo (p. 751) incide: "Debido a su presencia en los medios de comunicación y el uso que de ella hacen las diferentes materias, la estadística tiene en la actualidad una gran importancia y su estudio ha de capacitar a los estudiantes para analizar de forma crítica las presentaciones falaces, interpretaciones sesgadas y abusos que a veces contiene la información de naturaleza estadística. En los primeros cursos se pretende una aproximación natural al estudio de fenómenos aleatorios sencillos mediante experimentación y el tratamiento, por medio de tablas y gráficas, de datos estadísticos. Posteriormente, el trabajo se encamina a la obtención de valores representativos de una muestra y se profundiza en la utilización de diagramas y gráficos más complejos con objeto de sacar conclusiones a partir de ellos. La utilización de la hojas de cálculo facilita el proceso de organizar la información, posibilita el uso de gráficos sencillos, el tratamiento de grandes cantidades de datos, y libera tiempo y esfuerzos de cálculo para dedicarlo a la formulación de preguntas, comprensión de ideas y redacción de informes".

Incluye, entre otros, los siguientes contenidos dentro del Bloque 6, Estadística y probabilidad:

- *Primer Curso* (p. 753)*:*
  - "Formulación de conjeturas sobre el comportamiento de fenómenos aleatorios sencillos y diseño de experiencias para su comprobación.
  - Reconocimiento y valoración de las matemáticas para interpretar y describir situaciones inciertas.
  - Diferentes formas de recogida de información. Organización en tablas de datos recogidos en una experiencia. Frecuencias absolutas y relativas. Diagramas de barras, de líneas y de sectores. Análisis de los aspectos más destacables de los gráficos".
- *Segundo curso* (p. 755)*:*
  - Diferentes formas de recogida de información. Organización de los datos en tablas Frecuencias absolutas y relativas, ordinarias y acumuladas. Diagramas estadísticos. Análisis de los aspectos más destacables de los gráficos.
  - Medidas de centralización: media, mediana y moda. Significado, estimación y cálculo. Utilización de las propiedades de la media para resolver problemas. Utilización de la media, la mediana y la moda para realizar comparaciones y valoraciones. Utilización de la hoja de cálculo para organizar los datos, realizar los cálculos y generar los gráficos más adecuados.

- *Tercer Curso* (p. 756)*:*
  - “Necesidad, conveniencia y representatividad de una muestra. Métodos de selección aleatoria y aplicaciones en situaciones reales. Atributos y variables discretas y continuas. Agrupación de datos en intervalos. Histogramas y polígonos de frecuencias. Construcción de la gráfica adecuada a la naturaleza de los datos y al objetivo deseado.
  - Media, moda, cuartiles y mediana. Significado, cálculo y aplicaciones. Análisis de la dispersión: rango y desviación típica. Interpretación conjunta de la media y la desviación típica. Utilización de las medidas de centralización y dispersión para realizar comparaciones y valoraciones. Actitud crítica ante la información de índole estadística. Utilización de la calculadora y la hoja de cálculo para organizar los datos, realizar cálculos y generar las gráficas más adecuadas.
  - Experiencias aleatorias. Sucesos y espacio muestral. Utilización del vocabulario adecuado para describir y cuantificar situaciones relacionadas con el azar. Cálculo de probabilidades mediante la regla de Laplace. Formulación y comprobación de conjeturas sobre el comportamiento de fenómenos aleatorios sencillos. Cálculo de la probabilidad mediante la simulación o experimentación. Utilización de la probabilidad para tomar decisiones fundamentadas en diferentes contextos. Reconocimiento y valoración de las matemáticas para interpretar, describir y predecir situaciones inciertas”.
- *Cuarto curso. Opción A* (p. 758)*:*
  - “Identificación de las fases y tareas de un estudio estadístico a partir de situaciones concretas cercanas al alumnado. Análisis elemental de la representatividad de las muestras estadísticas. Gráficas estadísticas: gráficas múltiples, diagramas de caja. Uso de la hoja de cálculo. Utilización de las medidas de centralización y dispersión para realizar comparaciones y valoraciones.
  - Experiencias compuestas. Utilización de tablas de contingencia y diagramas de árbol para el recuento de casos y la asignación de probabilidades. Utilización del vocabulario adecuado para describir y cuantificar situaciones relacionadas con el azar”.
- *Cuarto curso. Opción B* (p. 759)*:*
  - “Identificación de las fases y tareas de un estudio estadístico. Análisis elemental de la representatividad de las muestras estadísticas. Gráficas estadísticas: gráficas múltiples, diagramas de caja. Análisis crítico de tablas y gráficas estadísticas en los medios de comunicación. Detección de falacias.
  - Representatividad de una distribución por su media y desviación típica o por otras medidas ante la presencia de descentralizaciones, asimetrías y valores atípicos. Valoración de la mejor representatividad en función de la existencia o no de valores atípicos. Utilización de las medidas de centralización y dispersión para realizar comparaciones y valoraciones.
  - Experiencias compuestas. Utilización de tablas de contingencia y diagramas de árbol para el recuento de casos y la asignación de probabilidades. Probabilidad condicionada. Utilización del vocabulario adecuado para describir y cuantificar situaciones relacionadas con el azar”.

Se incluyen, asimismo, los siguientes criterios de evaluación relacionados con estos temas:

- *Primer curso* (p. 754)*:*
  - “Hacer predicciones sobre la posibilidad de que un suceso ocurra a partir de información previamente obtenida de forma empírica. Se trata de valorar la capacidad para diferenciar los fenómenos deterministas de los aleatorios y, en estos últimos, analizar las regularidades obtenidas al repetir un número significativo de veces una experiencia aleatoria y hacer predicciones razonables a partir de los mismos. Además, este criterio pretende verificar la comprensión del concepto de frecuencia relativa y, a partir de ella, la capacidad de inducir la noción de probabilidad.
  - Organizar e interpretar informaciones diversas mediante tablas y gráficas, e identificar relaciones de dependencia en situaciones cotidianas. Este criterio pretende valorar la capacidad de identificar las variables que intervienen en una situación cotidiana, la relación de dependencia entre ellas y visualizarla gráficamente. Se trata de evaluar, además, el uso de las tablas como instrumento para recoger información y transferirla a unos ejes coordenados, así como la capacidad para interpretar de forma cualitativa la información presentada en forma de tablas y gráficas”.
- *Segundo curso* (p. 755)*:*
  - “Formular las preguntas adecuadas para conocer las características de una población y recoger, organizar y presentar datos relevantes para responderlas, utilizando los métodos estadísticos apropiados y las herramientas informáticas adecuadas. Se trata de verificar, en casos sencillos y relacionados con su entorno, la capacidad de desarrollar las distintas fases de un estudio estadístico: formular la pregunta o preguntas que darán lugar al estudio, recoger la información, organizarla en tablas y gráficas, hallar valores relevantes (media, moda, valores máximo y mínimo, rango) y obtener conclusiones razonables a partir de los datos obtenidos. También se pretende valorar la capacidad para utilizar la hoja de cálculo, para organizar y generar las gráficas más adecuadas a la situación estudiada.
- *Tercer curso* (p. 757)*:*
  - “Elaborar e interpretar informaciones estadísticas teniendo en cuenta la adecuación de las tablas y gráficas empleadas, y analizar si los parámetros son más o menos significativos. Se trata de valorar la capacidad de organizar, en tablas de frecuencias y gráficas, información de naturaleza estadística, atendiendo a sus aspectos técnicos, funcionales y estéticos (elección de la tabla o gráfica que mejor presenta la información), y calcular, utilizando si es necesario la calculadora o la hoja de cálculo, los parámetros centrales (media, mediana y moda) y de dispersión (recorrido y desviación típica) de una distribución. Asimismo, se valorará la capacidad de interpretar información estadística dada en forma de tablas y gráficas y de obtener conclusiones pertinentes de una población a partir del conocimiento de sus parámetros más representativos.
  - Hacer predicciones sobre la posibilidad de que un suceso ocurra a partir de información previamente obtenida de forma empírica o como resultado del recuento de posibilidades, en casos sencillos. Se pretende medir la capacidad de identificar los sucesos elementales de un experimento aleatorio sencillo y otros sucesos asociados a

dicho experimento. También la capacidad de determinar e interpretar la probabilidad de un suceso a partir de la experimentación o del cálculo (regla de Laplace), en casos sencillos. Por ello tienen especial interés las situaciones que exijan la toma de *decisiones* razonables a partir de los resultados de la experimentación, simulación o, en su caso, del recuento".

- *Cuarto curso opción A y B* (p. 758 y 760).
  - "Elaborar e interpretar tablas y gráficos estadísticos, así como los parámetros estadísticos más usuales correspondientes a distribuciones discretas y continuas, y valorar cualitativamente la representatividad de las muestras utilizadas. Se trata de valorar la capacidad de organizar la información estadística en tablas y gráficas y calcular los parámetros que resulten más relevantes con ayuda de la calculadora o la hoja de cálculo. En este nivel se pretende, además, que tengan en cuenta la representatividad y la validez del procedimiento de elección de la muestra y analicen la pertinencia de la generalización de las conclusiones del estudio a toda la población.
  - Aplicar los conceptos y técnicas de cálculo de probabilidades para resolver diferentes situaciones y problemas de la vida cotidiana. Se pretende que sean capaces de identificar el espacio muestral en experiencias simples y en experiencias compuestas sencillas, en contextos concretos de la vida cotidiana, y utilicen la regla de Laplace, los diagramas de árbol o las tablas de contingencia para calcular probabilidades. Se pretende, además, que los resultados obtenidos se utilicen para la toma de decisiones razonables en el contexto de los problemas planteados".

Respecto a la Consejería de Educación de la Junta de Andalucía (2007b), los temas de estadística y probabilidad se incluyen en Educación Secundaria en el Bloque 6: "Interpretación de fenómenos ambientales y sociales a través de las matemáticas y de las estadísticas y la probabilidad". Se citan como contenidos de este Bloque los correspondientes a los Bloques 5, *Funciones y gráficas y 6, Estadística y probabilidad* del Decreto del Ministerio (MEC, 2006b), que se ponen en este currículo en relación directa.

Entre las orientaciones, se indica que "la estadística y la probabilidad también están presenta hoy día en las diferentes materias, así como en los medios de comunicación, en los que aparecen datos que es necesario interpretar. Además de obtener conclusiones a partir de los datos expuestos en un gráfico o en una tabla es necesario conocer los procesos previos a su representación. Abordar cuestiones de planificación para la recogida de la información, utilizar técnicas de recuento y manipulación de los datos, así como estudiar la forma para agruparlos, son tareas tan importantes como los cálculos que con ellos puedan realizarse y su posterior interpretación" (p. 55).

Se resalta la relación de este núcleo temático con los siguientes (MEC, 2006b): Bloque 1, Contenidos comunes, de 1.º a 4.º; Bloque 2, Números, de 1.º a 4.º; Bloque 3, Álgebra; y Bloque 4, Geometría, de 1.º a 4.º Asimismo se sugiere la conveniencia de relacionarlo con Ciencias sociales, geografía e historia, Ciencias de la naturaleza y Biología y geología, en el 4.º curso. Las orientaciones metodológicas y criterios de evaluación son similares a los sugeridos en por el Ministerio de Educación.

3. BACHILLERATO

La enseñanza de las Matemáticas en el Bachillerato comprenden los cursos 1º y 2º, y el Decreto 1467/2007, de 2 de noviembre, por el que se establece la estructura del bachillerato y se fijan sus enseñanzas mínimas (MEC, 2007) fija los siguientes contenidos en relación con el tema que nos interesa:

- *Matemáticas I, modalidad de Ciencias y Tecnología* (p. 45449)*:*
  - "Distribuciones bidimensionales. Relaciones entre dos variables estadísticas. Regresión lineal.
  - Estudio de la probabilidad compuesta, condicionada, total y a posteriori.
  - Distribuciones binomial y normal como herramienta para asignar probabilidades a sucesos".
- *Matemáticas aplicadas a las ciencias sociales I, modalidad Humanidades y Ciencias Sociales* (p. 45475)*:*
  - "Estadística descriptiva unidimensional. Tipos de variables.
  - Métodos estadísticos. Tablas y gráficos. Parámetros estadísticos de localización, de dispersión y de posición.
  - Distribuciones bidimensionales. Interpretación de fenómenos sociales y económicos en los que intervienen dos variables a partir de la representación gráfica de una nube de puntos. Grado de relación entre dos variables estadísticas. Regresión lineal. Extrapolación de resultados.
  - Asignación de probabilidades a sucesos. Distribuciones de probabilidad binomial y normal".
- *Matemáticas aplicadas a las ciencias sociales II, modalidad Humanidades y Ciencias Sociales* (p. 45476):
  - "Profundización en los conceptos de probabilidades a priori y a posteriori, probabilidad compuesta, condicionada y total. Teorema de Bayes.
  - Implicaciones prácticas de los teoremas: Central del límite, de aproximación de la Binomial a la Normal y Ley de los Grandes Números.
  - Problemas relacionados con la elección de las muestras. Condiciones de representatividad. Parámetros de una población. Distribuciones de probabilidad de las medias y proporciones muestrales.
  - Intervalo de confianza para el parámetro p de una distribución binomial y para la media de una distribución normal de desviación típica conocida.
  - Contraste de hipótesis para la proporción de una distribución binomial y para la media o diferencias de medias de distribuciones normales con desviación típica conocida".

Otro aspecto a resaltar de los Decretos, son los criterios de evaluación que se contemplan, entre los que se pretende para el

- *Matemáticas I, modalidad de Ciencias y Tecnología* (p. 45450)
  - "Asignar probabilidades a sucesos correspondientes a fenómenos aleatorios simples

y compuestos y utilizar técnicas estadísticas elementales para tomar decisiones ante situaciones que se ajusten a una distribución de probabilidad binomial o normal. En este criterio se pretende medir la capacidad para determinar la probabilidad de un suceso, utilizando diferentes técnicas, analizar una situación y decidir la opción más conveniente. También se pretende comprobar la capacidad para estimar y asociar los parámetros relacionados con la correlación y la regresión con las situaciones y relaciones que miden.

- Realizar investigaciones en las que haya que organizar y codificar informaciones, seleccionar, comparar y valorar estrategias para enfrentarse a situaciones nuevas con eficacia, eligiendo las herramientas matemáticas adecuadas en cada caso. Se pretende evaluar la madurez del alumnado para enfrentarse con situaciones nuevas procediendo a su observación, modelado, reflexión y argumentación adecuada, usando las destrezas matemáticas adquiridas. Tales situaciones no tienen por qué estar directamente relacionadas con contenidos concretos; de hecho, se pretende evaluar la capacidad para combinar diferentes herramientas y estrategias, independientemente del contexto en que se hayan adquirido".

- *Matemáticas aplicadas a las ciencias sociales I, modalidad Humanidades y Ciencias Sociales* (p. 45475)*:*

  - "Utilizar las tablas y graficas como instrumento para el estudio de situaciones empíricas relacionadas con fenómenos sociales y analizar funciones que no se ajusten a ninguna fórmula algebraica, propiciando la utilización de métodos numéricos para la obtención de valores no conocidos. Este criterio está relacionado con el manejo de datos numéricos y en general de relaciones no expresadas en forma algebraica. Se dirige a comprobar la capacidad para ajustar a una función conocida los datos extraídos de experimentos concretos y obtener información suplementaria mediante técnicas numéricas.
  - Distinguir si la relación entre los elementos de un conjunto de datos de una distribución bidimensional es de carácter funcional o aleatorio e interpretar la posible relación entre variables utilizando el coeficiente de correlación y la recta de regresión. Se pretende comprobar la capacidad de apreciar el grado y tipo de relación existente entre dos variables, a partir de la información gráfica aportada por una nube de puntos; así como la competencia para extraer conclusiones apropiadas, asociando los parámetros relacionados con la correlación y la regresión con las situaciones y relaciones que miden. En este sentido, más importante que su mero cálculo es la interpretación del coeficiente de correlación y la recta de regresión en un contexto determinado.
  - Utilizar técnicas estadísticas elementales para tomar decisiones ante situaciones que se ajusten a una distribución de probabilidad binomial o normal. Se pretende evaluar si, mediante el uso de las tablas de las distribuciones normal y binomial, los alumnos son capaces de determinar la probabilidad de un suceso, analizar una situación y decidir la opción más adecuada".

- *Matemáticas aplicadas a las ciencias sociales II, modalidad Humanidades y Ciencias Sociales* (pp. 45476- 7):

  - "Asignar probabilidades a sucesos aleatorios simples y compuestos, dependientes o independientes, utilizando técnicas personales de recuento, diagramas de árbol o tablas de contingencia. Se trata de valorar tanto la competencia para estimar y calcular probabilidades asociadas a diferentes tipos de sucesos como la riqueza de procedimientos a la hora de asignar probabilidades a priori y a posteriori, compuestas o condicionadas. Este criterio evalúa también la capacidad, en el ámbito de las ciencias sociales, para tomar decisiones de tipo probabilístico que no requieran la utilización de cálculos complicados.

- Diseñar y desarrollar estudios estadísticos de fenómenos sociales que permitan estimar parámetros con una fiabilidad y exactitud prefijadas, determinar el tipo de distribución e inferir conclusiones acerca del comportamiento de la población estudiada. Se pretende comprobar la capacidad para identificar si la población de estudio es normal y medir la competencia para determinar el tipo y tamaño muestral, establecer un intervalo de confianza para μ y p, según que la población sea Normal o Binomial, y determinar si la diferencia de medias o proporciones entre dos poblaciones o respecto de un valor determinado, es significativa. Este criterio lleva implícita la valoración de la destreza para utilizar distribuciones de probabilidad y la capacidad para inferir conclusiones a partir de los datos obtenidos.
- Analizar de forma critica informes estadísticos presentes en los medios de comunicación y otros ámbitos, detectando posibles errores y manipulaciones tanto en la presentación de los datos como de las conclusiones. Se valora el nivel de autonomía, rigor y sentido crítico alcanzado al analizar la fiabilidad del tratamiento de la información estadística que hacen los medios de comunicación y los mensajes publicitarios, especialmente a través de informes relacionados con fenómenos de especial relevancia social".

Al desarrollar los núcleos de contenidos propuestos la Junta de Andalucía (2008) remite a estos contenidos, que deben desarrollarse teniendo en cuenta cuatro núcleos transversales:

1. La resolución de problemas.
2. Aprender de y con la historia de las matemáticas.
3. Introducción a los métodos y fundamentos matemáticos.
4. Modelización matemática, dando sugerencias de posibles contenidos históricos.

Respecto a la Historia, para el de bachillerato de Ciencias y Tecnología, la Consejería de Educación indica que se pueden trabajar, entre otros, los siguientes aspectos (p. 171): "Los inicios del cálculo de probabilidades desde Pacioli a Gauss y su influencia en las distribuciones de probabilidad. Las formulaciones actuales dadas por Borel y Kolmogorov. La progresión de la estadística durante el siglo XX con la aplicación de la probabilidad".

También se sugieren los siguientes aspectos históricos al desarrollar los núcleos de contenidos propuestos en el Real Decreto 1467/2007 para el currículo de Matemáticas Aplicadas a las Ciencias Sociales I y II (p. 203): "Historia de la Estadística y la Probabilidad: los orígenes de los censos desde la Antigüedad a nuestros días. Consideración de la estadística como ciencia: aportaciones deAchenwall, Quételect y Colbert. Los orígenes de la Probabilidad: Pacioli, Tartaglia, Pascal, Bernoulli, De Moivre, Laplace y Gauss. Las relaciones actuales entre Estadística y probabilidad: Pearson. Estadística descriptiva: Florence Nightingale".

Los criterios de evaluación son similares a los del Ministerio.