
EXPLORACIÓN DE LA NOCIÓN DE VARIACIÓN EN SITUACIONES DE AZAR

Ernesto Sánchez
Karla Trujillo
Dpto. de Matemática Educativa
CINVESTAV, México

Recibido: 12 mayo 2008 / Aceptado: 26 junio 2008

RESUMEN

Se exploran las nociones de variabilidad estadística en situaciones de azar de grupos de estudiantes mexicanos de tres niveles de escolaridad: secundaria (12-15 años), bachillerato (16-18 años) y superior (19-22 años). La motivación del estudio aquí presentado fue responder la pregunta: ¿Se puede suponer razonablemente una trayectoria en el pensamiento de los estudiantes sobre la variabilidad que se caracteriza al comienzo por un predominio de la idea de aleatoriedad, después por el de la estructura, hasta lograr una integración de ambas en una noción de variación estadística? ¿Qué consecuencias se derivan de un modelo como el anterior?

Para responder la primera pregunta, se aplicó un cuestionario de 12 preguntas a 327 estudiantes de Secundaria, 214 de Bachillerato y 74 universitarios. En el presente trabajo se analizan sólo tres preguntas, que fueron numeradas: 6, 10 y 11. Se ofrecen algunas reflexiones para responder la segunda.

Palabras clave: formación de conceptos, estadística, variación, distribuciones, educación matemática.

ABSTRACT

This paper explores the notions that some Mexican students hold about statistics variability in chance settings. The study attempts to answer the question: Is it reasonable to suppose that the students' concept of variation in chance settings pass through three stages: the perception of the randomness, the consciousness of the probabilistic structure underlying a situation in a chance context and the understanding of the relation of this structure with the empirical data? What are the implications of such a model?

In order to answer the first question, a questionnaire was administered on 327 students of middle school (12 – 15 years old), 214 high school students (16 – 18 years old) and 74 college students (19 – 22 years old) and we analyzed three of the items, numbered 6, 10 and 11. We offer some reflections to answer the second one.

Key words: concept formation, statistics, variation, distributions, mathematics education.

1. INTRODUCCIÓN Y ANTECEDENTES

Hace poco más de 10 años, Shaughnessy (1997) llamó la atención sobre la falta de estudios didácticos sobre variabilidad estadística.

Una de las primeras investigaciones empíricas sobre variación en ambientes de azar fue realizada por Shaughnessy, Watson, Moritz y Reading (1999), en la cual se formularon las siguientes preguntas: 1) ¿qué nociones tienen los estudiantes sobre la variabilidad en el muestreo? 2) ¿Cómo podemos empezar a medir sus nociones de variabilidad y dispersión?

Para responder estas preguntas se modificó una tarea sobre muestreo de un recipiente con tres tipos de gomas de mascar (Gumball), de tal forma que permitiera revelar los patrones de pensamiento de los estudiantes en variabilidad. La idea fue simplemente preguntar sobre lo que se espera obtener en muestreo repetido; supongamos un recipiente en que hay 50% de gomas rojas, 30% de gomas azules y 20% de gomas amarillas. Se extrae una muestra al azar de 10 gomas ¿Cuántas rojas esperas obtener en la muestra? La respuesta comúnmente aceptada es 5. Sin embargo, si el experimento se repite 10 veces, esperar una secuencia de 10 cincos resulta estadísticamente inaceptable pues niega la existencia de la variabilidad. Se ha observado –comentan los autores– que la comprensión de las medidas de tendencia central aumenta con la edad de los estudiantes, mientras que la de la variabilidad se comporta erráticamente a través de los diferentes grados de escolaridad. Esto es compatible con el hecho de que a lo largo del currículo de matemáticas hay un énfasis mayor en las medidas de tendencia central que en las de variabilidad.

En un experimento realizado por Watson y Kelly (2003) para explorar el pensamiento y evolución de estudiantes de grados 3 a 9 después de que estudiaron algo sobre variación, se les pidió predecir los resultados de sorteos repetidos de una ruleta con sólo dos eventos equiprobables. Se observó una disminución monótona respecto a los grados escolares en la habilidad de los estudiantes para incluir en sus predicciones una variabilidad razonable. Los autores opinan que dicha disminución está en relación inversa con su conocimiento de la probabilidad teórica.

En Watson, Kelly, Callingham, Shaughnessy (2003) se reportan los resultados de la aplicación de un cuestionario que incluye reactivos sobre variación muestral, presentación de la variación, aleatoriedad, y medidas y fuentes de variación. Para analizar los datos proponen una categorización basada en la taxonomía SOLO de Biggs y Collis. Se examina una muestra de 746 estudiantes de grados 3, 5, 7 y 9. Uno de sus objetivos era “desarrollar una escala de medida de la comprensión de la variación en el contexto de los temas de azar y manejo de datos del currículum” (p. 15). Sugieren que en futuras investigaciones se puede “Probar el cuestionario con otros estudiantes, por ejemplo, en otros ambientes culturales” (p. 19). El presente estudio toma, en parte, dicha sugerencia ya que estudia aspectos similares con estudiantes mexicanos.

Shaughnessy, Canadá y Ciancetta (2003) exploraron el pensamiento de 84 estudiantes de secundaria con tres tareas que involucraban pruebas repetidas; una de ellas es similar a la pregunta 6 del presente trabajo. Ellos observaron que los estudiantes tienden a despreciar la variabilidad en el contexto de azar y probabilidad. La razón de ello se atribuye a la forma en la cual la probabilidad es introducida a los estudiantes en la escuela.

2. MARCO DE ANÁLISIS

En este estudio se distinguen tres conceptos relacionados con la noción de *variabilidad estadística*, a saber: *aleatoriedad*, *estructura* y *variación*. Se considera aquí la *aleatoriedad* en un sentido amplio de modo que incluye otras nociones como *azar*, *incertidumbre*, *desorden*, *caos*, *dispersión* y *desviaciones*. La *estructura* incluye nociones como *regularidad*, *centros*, *tendencias* y *distribuciones*. Finalmente, la *variación* es una medida que combina la aleatoriedad y la estructura. Cuando Batanero se refiere a la naturaleza de los datos estadísticos, menciona cierta relación de los conceptos arriba mencionados con el análisis exploratorio de datos:

Para entender los principios por los que se guía el análisis exploratorio, se ha de tener en cuenta que los datos están constituidos por dos partes: la 'regularidad' o tendencia y las 'desviaciones' o variabilidad. Por regularidad entendemos la estructura simplificada de un conjunto de observaciones (en una nube de puntos es la recta a la cual se ajusta). Las diferencias de los datos con respecto a esta estructura (diferencia en nuestro caso respecto a la recta) representan las desviaciones o residuos de los datos, que usualmente no tienen por qué presentar una estructura determinada.. (Batanero, 2001; p. 27)

Las experiencias aleatorias son, por definición, impredecibles. Una pregunta como “¿Qué esperas que ocurra al lanzar una moneda? o ¿Qué crees que ocurra al lanzar un dado?” generalmente tienen como respuesta “No se sabe, porque es impredecible” o “Puede ocurrir cualquier cara de la moneda” o “cualquier cara del dado”. Son respuestas lícitas y convenientes teniendo en cuenta la “omnipresencia” de la aleatoriedad en el experimento. La respuesta a la pregunta ¿Qué crees que ocurrirá...? debe cambiar cuando se sortea una distribución no uniforme. Si ésta se conoce, hay razones para esperar algunos resultados más que otros; en algunas situaciones hay resultados que, aunque con probabilidad mayor que cero, deben considerarse imposibles, teniendo en cuenta la escala en la que se inserta la situación (Borel, 1971). Muchos procedimientos estadísticos se basan en esa consideración.

La estructura (media y distribución) de la situación o de los datos es la que permite elaborar esos procedimientos. Sin embargo, el conocimiento de la estructura de un experimento suele ser utilizado inadecuadamente por los estudiantes para intentar predecir los resultados de la experiencia; ante la pregunta ¿qué crees que pasará...? responden con el valor más probable de la distribución, aunque la probabilidad de este valor sea muy pequeña. Un estadístico responde a la pregunta ¿qué crees que ocurrirá...? con otra pregunta, a saber, ¿con qué probabilidad se quiere que ocurra la predicción? Sabe que cualquier resultado individual se debe considerar imposible o muy poco probable, pero que se puede elegir un evento no singular que proporcione una respuesta satisfactoria. Este paso es crucial en la comprensión estadística de la situación.

Por otro lado, un problema de *identificación* de la plausibilidad de un resultado consiste en decidir cuál de varias opciones podrían haber sido el resultado de un experimento determinado. Un procedimiento para resolverlos consiste en buscar razones para

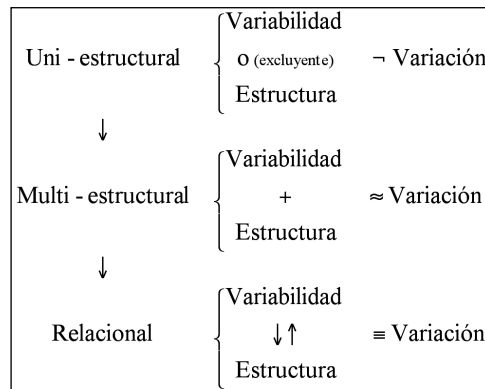
desechar ciertas opciones, por ejemplo, se podría rechazar una opción porque tenía probabilidad muy baja de haber ocurrido. En este tipo de problemas es importante la consideración de la variabilidad; ¿la opción refleja la variabilidad del experimento que generó el resultado?

Jerarquía SOLO. Con base en que la *variación* combina la *aleatoriedad* y la *estructura*, es posible organizarlas para formar un ciclo de la jerarquía SOLO (Structure Observed of the Learning Outcomes), con la cual analizar la conducta de los estudiantes frente a ciertos problemas que involucran variabilidad.

La taxonomía SOLO (Biggs & Collis, 1991) se ha estado usando en varios estudios para codificar las respuestas de estudiantes y grupos de ellos, en diferentes áreas de conocimiento, en particular, en estadística (Watson, *et al.* 2003; Reading & Lawrie, 2004). Consiste en cinco modos de funcionamiento, similares a las etapas de desarrollo de Piaget, que dan cuenta de diferentes niveles de competencias alcanzadas. Para nosotros, el aspecto importante, es que se pueden identificar niveles en cada modo. De hecho, estos niveles pueden aplicarse a los tipos de respuesta de los estudiantes a una pregunta o problema, con base en la identificación de los aspectos importantes que permiten la comprensión del tema o concepto necesario para resolver la tarea. Los siguientes niveles se pueden identificar en cada modo (siendo más importantes los últimos tres): Pre-estructural, uni-estructural, multi-estructural y relacional.

En el presente trabajo, se dice que un estudiante se centra en la *aleatoriedad*, cuando cree que en una situación aleatoria “cualquier cosa puede ocurrir”, es decir, que puede obtener cualquier dato; en estos casos se ignora que los datos generados tienen una estructura; quien así responde está en una etapa uniestructural. Se dice que un estudiante se centra en la *estructura* cuando espera que los datos se comporten de manera regular o que reflejen un rasgo sobresaliente de la distribución (moda o media), en estos casos se ignora la aleatoriedad. Alguien que se apoya sólo en la estructura también está en una etapa uni-estructural. Finalmente, se dice que un estudiante espera u observa la *variación*, si tiene en cuenta tanto la aleatoriedad como la estructura; cuando esto ocurre, la respuesta se ubica en multiestructural o relacional de acuerdo con el nivel de elaboración de la relación entre ambas componentes.

A manera de hipótesis suponemos una evolución de los sujetos en su noción intuitiva de la variación frente a conjuntos de datos, en la que ellos comienzan centrados en la aleatoriedad (uniestructural), pasan después a centrarse en la estructura (uniestructural) y luego en tratar de coordinar ambas componentes pero sin relacionarlas convenientemente (multiestructural); en una etapa subsiguiente las integran estableciendo las relaciones convenientes entre ellas (relacional). En estas dos últimas etapas emergería el concepto de variación. En el siguiente esquema se representa esta hipótesis:



Bajo estos supuestos nos hacemos las preguntas siguientes.

3. LA PREGUNTA DE INVESTIGACIÓN

¿Se puede suponer razonablemente una trayectoria en el pensamiento sobre la variabilidad de los estudiantes que se caracteriza al comienzo por un predominio de la idea de aleatoriedad, después por el de la estructura, hasta lograr una integración de ambas en una noción de variación estadística? ¿Qué consecuencias se derivan de un modelo como el anterior?

4. METODOLOGÍA

4.1. Participantes

Se aplicó un cuestionario de 12 preguntas a varios grupos de estudiantes de dos niveles educativos de la Ciudad de México; 327 estudiantes de Secundaria, 215 de Bachillerato y 74 universitarios. Los estudiantes de Bachillerato pertenecían a dos diferentes sistemas educativos; 87 estaban al final de su segundo año y nunca habían tomado ningún curso de Estadística, mientras que los otros 127 estaban a punto de terminar su tercer año y habían cursado Estadística por un semestre. Sin embargo, no se notaron diferencias significativas entre los grupos de modo que se analizaron sin señalar esas características (Sánchez y Mercado, 2006). Los estudiantes universitarios son de la carrera de Actuaría de la Facultad de Ciencias de la Universidad Nacional, unos estaban llevando Estadística I y otros ya llevaban Estadística II; con relación a las preguntas examinadas en este estudio, no hubo diferencias significativas entre ambos grupos. En conclusión, se distinguirán sólo tres clases de estudiantes de acuerdo a alguno de los tres niveles siguientes: Secundaria (12-15 años), Bachillerato (16-18), Universitario (19-22).

4.2. Instrumentos

Se diseñó un cuestionario con 12 preguntas obtenidos de la literatura sobre variación, algunos de ellos fueron levemente modificados (Watson *et al.* 2003). En este artículo, sólo se analizan 3 de esos problemas, a saber 6, 10 y 11. A continuación se presenta cada uno de ellos y se explica el tipo de respuesta que se espera.

6. *Imagina que ahora lanzas 60 veces el dado. Llena la siguiente tabla escribiendo cuántas veces crees que saldrá cada número.*

Caras del dado	Num. de veces
1	
2	
3	
4	
5	
6	
Total	60

¿Por qué consideras este resultado?

Este problema pide una predicción ¿Qué crees que ocurrirá? Una respuesta razonable requiere que el estudiante ponga en juego su intuición estadística y las nociones de aleatoriedad, probabilidad y variación.

La naturaleza aleatoria del experimento significa que no se puede decir exactamente lo que ocurrirá, de manera que proponer una determinada distribución de frecuencias absolutas no es una respuesta conveniente. Si se tiene en cuenta sólo la aleatoriedad, se podría responder: “no se puede saber” o “cualquier cosa puede ocurrir”. Estas respuestas, aunque verdaderas, tampoco son convenientes a causa de que no tienen en cuenta la estructura de la situación.

Una respuesta que considera la variación y conjuga tanto la aleatoriedad como la estructura es la de proponer un rango para la frecuencia de cada cara, de manera que la probabilidad de que la frecuencia resultante quede dentro de ese rango sea grande (75%, 90% ó 95%, por ejemplo). Los estudiantes de nivel superior podrían utilizar la “regla empírica” para estimar las anteriores probabilidades.

10. *Se tiene una urna con 3 bolas, cada una tiene una letra (A, B y C). Juan saca una bola sin ver, anota en una tabla la letra que corresponde a la bola y la regresa a la urna. Juan repite el experimento 30 veces. ¿Cuál de las siguientes tablas crees que sea la que obtuvo Juan? Marca la tabla que consideres correcta.*

Letra	# de bolas
A	12
B	7
C	11
TOTAL	30

[] Tabla 1

Letra	# de bolas
A	11
B	18
C	1
TOTAL	30

[] Tabla 2

Letra	# de bolas
A	10
B	10
C	10
TOTAL	30

[] Tabla 3

En este trabajo se asume que la respuesta más razonable es la tabla 1. Se llegaría a dicha respuesta encontrando razones que llevan a descartar a las otras. El sentido común ofrece una para no creer que la tabla 3 haya ocurrido, a saber, “es muy uniforme”; como el experimento es aleatorio no se espera que el resultado presente ese patrón. Una razón más matemática lleva a no creer que en la tabla 2 haya resultado de los sorteos porque “es muy difícil que la bola C sólo haya ocurrido una vez en 30 sorteos”. Esto lleva a elegir la tabla 1 como la mejor, pues no presenta los inconvenientes que se pueden ver en las otras respuestas.

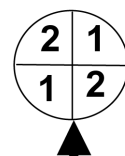
En una discusión de Kahneman y Tversky (1982) critican la elección de una opción como la que aquí se considera correcta, pero la formulación del problema es diferente, pues la pregunta es: “en muchas repeticiones ¿cuál de las tres distribuciones crees que se presentará más frecuentemente?”. En este caso, se considera que la probabilidad de ocurrir de la tabla 3 es mayor que la de la tabla 1. Los que eligen la tabla 1 se dejan llevar por la heurística de representatividad. Ésta sesga la respuesta correcta.

Lo anterior no se aplica en el caso presente, pues la situación es única, se realizó una sola vez y se debe tomar la decisión de elegir una de las tablas como la más probable de haber resultado del experimento.

Una forma de reformular el razonamiento del sentido común es como sigue. Cada tabla puede juzgarse involucrando una consideración de variabilidad. Se definen 3 eventos: “sin variación” (SV), “apariencia realista” (AR) y “valores extremos” (VE). Se elige el resultado que pertenezca a la “apariencia realista”. Este criterio es efectivo para obtener éxito en la valoración de la situación, en el sentido en que se refleja en la siguiente simulación:

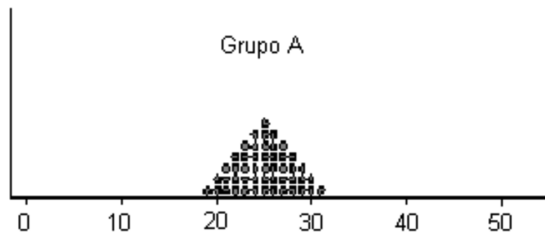
Se generan 30 extracciones al azar de la urna mencionada; la distribución resultante pertenecerá a AR, SV o VE. El programa ubica a qué evento pertenece el resultado e “inventa” dos distribuciones de modo que pertenezcan a los otros eventos. Sin indicar el resultado del sorteo, nos pide que lo adivinemos. La elección de la opción AR tiene muchas más probabilidades de éxito.

11. Tres grupos A, B y C, realizaron el experimento de girar la ruleta (50 veces por alumno), cada alumno anotó el número de veces que obtuvieron un 2, reunieron los resultados de todo el grupo y los graficaron. Sin embargo, al maestro le informaron de que uno o dos de los grupos inventaron los resultados sin hacer el experimento, mientras que los otros dos o el otro sí lo hicieron y graficaron.



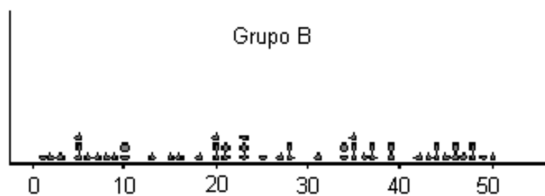
Observa las siguientes gráficas e indica en cada caso si se trata de una gráfica inventada o real y las razones que te llevan a dar tu respuesta.

a) Grupo A



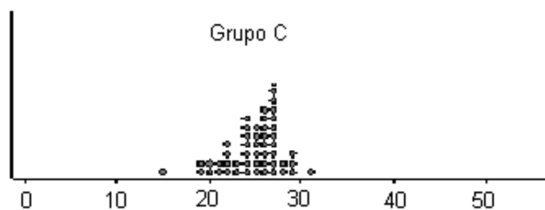
¿Los resultados del grupo A son reales o inventados?
Explica tu respuesta:

b) Grupo B



¿Los resultados del grupo B son reales o inventados?
Explica tu respuesta:

c) Grupo C:



¿Los resultados del grupo C son reales o inventados?
Explica tu respuesta.

Esta pregunta es similar a la pregunta 10 anterior. Se realizó una sola experiencia y de los resultados se debe escoger aquella que se crea que pudo haber resultado realmente de la experiencia. Se asume que la respuesta más razonable es considerar las dos pri-

meras distribuciones inventadas y la tercera real. La distribución del inciso a), refleja el esquema binomial de la distribución y debe ser descartada porque la naturaleza aleatoria del fenómeno hace difícil que las frecuencias absolutas se distribuyan tan ordenadamente. La segunda distribución no refleja el esquema binomial de la experiencia, representa más la realización de un modelo uniforme. Finalmente la tercera opción representa a un evento con las características de reflejar un esquema binomial y la aleatoriedad. Hay que mencionar que la distribución particular del inciso c) es menos probable que la del inciso a), sin embargo, al igual que en el problema 10, de lo que se trata es de evaluar cuál de las tres pudo haber sido resultado de una experiencia que ya se llevó a cabo; en este caso, el criterio de representatividad resulta efectivo.

4.3. Procedimientos

El profesor de cada grupo aplicó el cuestionario a los estudiantes de Secundaria y Bachillerato; mientras que en el caso de los universitarios uno de los autores fue quien aplicó el cuestionario.

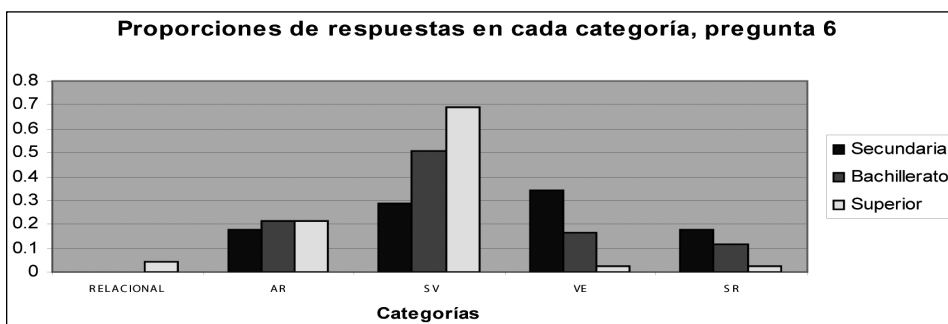
Para clasificar las respuestas se utilizó la siguiente codificación, que es similar en algunos aspectos a la utilizada por Watson *et al.* (2003) y que está basada en los ciclos jerárquicos de la taxonomía SOLO propuesta por Biggs y Collis (1991): “apariencia realista” (AR), se usa para clasificar las respuestas de la pregunta 6 si un individuo da una distribución de frecuencias no uniforme que sume 60 y cuyas frecuencias estén en el intervalo de 4 a 16 (90% de confianza). Esta misma categoría se usa si un estudiante elige la tabla 1 para responder a la pregunta 10. Finalmente, AR se utiliza para clasificar a los estudiantes que siguen una secuencia respuesta IIR en la pregunta 11, donde “IIR” significa “inventada, real, inventada”. La categoría “sin variación” (SV) se usa para clasificar las respuestas que dan una distribución uniforme en la pregunta 6 y también si se selecciona la tabla 3 en la pregunta 10; una respuesta está en esta categoría si la secuencia RII es elegida en la pregunta 11, donde “RII” significa “real, inventada, inventada”.

La categoría “valores extremos” (VE) se usa para clasificar las respuestas a la pregunta 6 en las cuales al menos un elemento esté fuera del rango 4 -16 (El evento ‘la frecuencia de cada cara está entre 4 y 6’ tiene 90% de probabilidad de ocurrir). También se clasifican en esta categoría las respuestas que seleccionan la tabla 2 en la pregunta 10 y si la secuencia IRI se ofrece como respuesta en la pregunta 11, donde IRI significa “inventada, real, inventada”. La categoría (SR) significa “sin respuesta” o “inconsistente”.

Finalmente, se define la categoría “relacional” para ubicar las respuestas que establecen una relación entre la aleatoriedad y la estructura, por ejemplo, dando un rango de valores alrededor de la media 10 en el problema 6. No se aplica a los problemas 10 y 11 porque no hay opción que establezca una estrecha relación entre la aleatoriedad y estructura.

5. RESULTADOS

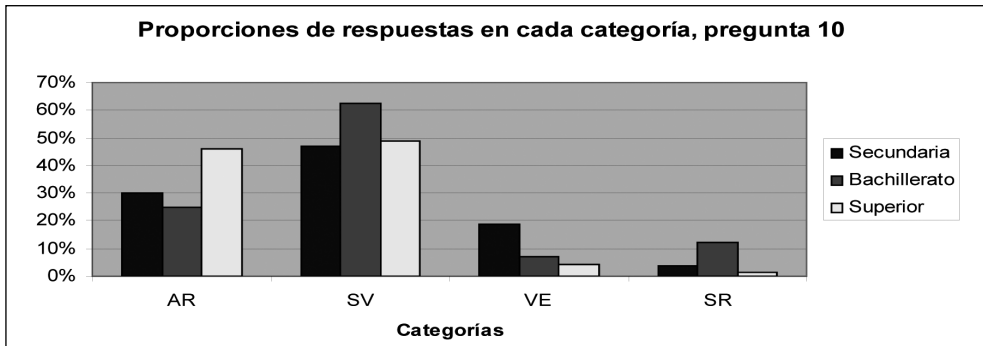
Las respuestas a los cuestionarios se han organizado con el propósito de comprobar la hipótesis de que la noción de variación comienza con la percepción del desorden y después con la noción de estructura. Si está hipótesis se cumple, se esperaría que los estudiantes de la secundaria respondan o elijan en mayor proporción el tipo de respuestas que hemos clasificado como “valores extremos”. Además, se esperaría que los estudiantes de bachillerato se inclinen por el tipo de respuestas que se ha clasificado como “sin variación” y que los estudiantes universitarios se movieran hacia el tipo de respuesta que clasificamos como “apariencia realista” y relacional.



Gráfica 1. Proporciones de respuesta, pregunta 6

En la gráfica 1 un rasgo a notar es que las respuestas clasificadas como relacionales sólo se presentaron en el nivel universitario y con muy poca frecuencia. La proporción de respuestas en la categoría de “apariencia realista” es casi constante a lo largo de los grados de escolaridad, alrededor de 20%. Es interesante observar que cuanto mayor sea el grado de escolaridad, la proporción de respuestas en la categoría “sin variación” aumenta y en la categoría “valores extremos” disminuye.

La proporción de respuestas que fueron clasificadas como “relacional” es de 4%. Las respuestas clasificadas en “apariencia realista” fueron propuestas casi en la misma proporción por los tres grupos –Secundaria, Bachillerato y universitario– (18%, 21% y 22%, respectivamente). Las respuestas clasificadas “sin variación” fueron propuestas con mayor frecuencia por estudiantes universitarios 69%, luego los de Bachillerato 50%, y finalmente los de nivel de Secundaria (29%). La proporción de las clasificadas en “valores extremos” fue de 34% en Secundaria, 16.3% en Bachillerato y 3% en el nivel universitario.

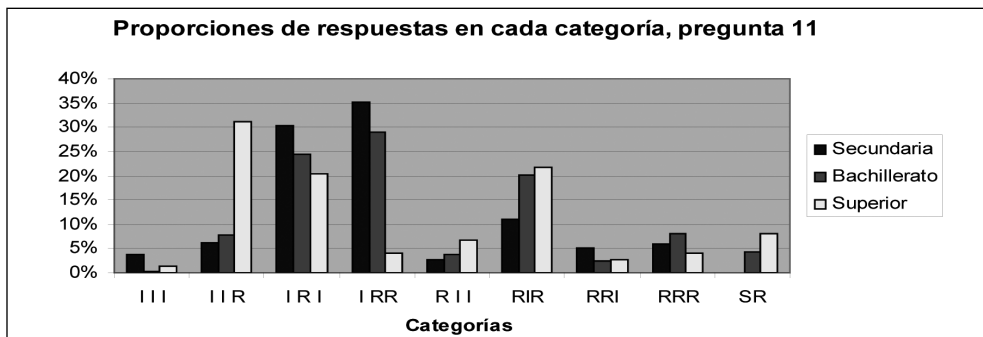


Gráfica 2. Proporciones de respuesta, pregunta 10

Se puede apreciar que la gráfica 2, que se refiere a la pregunta 10 y en las que se propusieron opciones a elegir, tiene una distribución similar a la de la pregunta 6, pero ahora la frecuencia en que se eligió la opción “Sin variación” por el grupo de Secundaria fue de 47% mientras que la frecuencia de elección de los de Bachillerato fue de 62.6% y en el nivel universitario de 49%.

Con referencia a la pregunta 11, los datos se han organizado como sigue: indicamos con III la opción “Inventada-Inventada-Inventada”, es decir, que se consideró que la gráfica del inciso a) es inventada, y las de los b) y c) también. De manera análoga se representa como IIR las respuestas “Inventada-Inventada-Realista” y las siguientes 6 opciones: IRI, IRR, RII, RIR, RRI, RRR.

Con relación a la hipótesis que se ha formulado sobre la evolución de la noción de variación, se esperaría que los estudiantes de Secundaria aceptaran como realista las respuestas que clasificamos como “valores extremos”, es decir, que respondan con mayor frecuencia $*R* = \{IRI, IRR, RRI, RRR\}$. En cambio, de los estudiantes de Bachillerato, responder $R^{**} = \{RRR, RRI, RIR, RII\}$ indicaría que tienden a darle más importancia a la estructura. Sin embargo, los resultados no son tan claros y presentan algunos matices que comentaremos enseguida.



Gráfica 3. Proporciones de respuesta, pregunta 11

Se sugiere poner atención en las respuestas con frecuencias mayores a 10%, a saber, IIR, IRI, IRR y RIR. La primera se considera la más razonable y, consecuentemente, es elegida con mayor frecuencia por los estudiantes universitarios. Las respuestas IRI e IRR son propuestas con mayor frecuencia por los estudiantes de Secundaria; no obstante, no son pocos los de Bachillerato que también propusieron lo mismo. Es remarcable que haya 24% y 21% de estudiantes de Bachillerato y universitarios respectivamente que proponga IRI; en cambio, la proporción de universitarios que propone IRR es baja mientras que en los de Secundaria y Bachillerato sigue siendo alta. Finalmente, los universitarios presentan mayor frecuencia (22%) en la respuesta RIR, seguida por los de Bachillerato (20%) y con sólo 11% de los de Secundaria.

6. DISCUSIÓN Y CONCLUSIONES

En este estudio se puede observar que los porcentajes de respuestas obtenidos en nuestra exploración son consistentes con la hipótesis de que se puede trazar una trayectoria del pensamiento de los estudiantes acerca de la variabilidad. Los más jóvenes son proclives a percibir principalmente la aleatoriedad en los fenómenos de azar, mientras que los mayores llevan su atención hacia la estructura. Sólo algunos de éstos manifiestan un nivel relacional de pensamiento sobre la variación. Es probable que el avance hacia una comprensión de la variación sólo pueda ser el resultado de una instrucción mejor diseñada de la que tradicionalmente se ofrece; esto es claro en el caso de México, pues los programas del Bachillerato general sólo proponen estudiar la desviación estándar como una fórmula para calcular.

Convendría discutir ¿cuáles son las razones que explican el hecho de que se perciba en primer lugar la aleatoriedad y posteriormente haya un cambio radical hacia la estructura? En este estudio se coincide con otros autores (Watson & Kelly, 2003; Shaughnessy & Ciancetta, 2003) en que es una consecuencia de los cursos de probabilidad y estadística. En particular, se ha sostenido que en estos cursos se pone mayor énfasis en la enseñanza de las medidas de tendencia central que en la enseñanza de la variación. Sin embargo, aunque es un hecho que así ocurre, surge la pregunta ¿por qué ha ocurrido eso durante tanto tiempo? ¿puede ser de otra manera?

Los tres momentos que se han destacado: aleatoriedad, estructura y variación pueden caracterizarse por pasos importantes: Piaget e Inhelder (1975) sostuvieron que la idea de azar y aleatoriedad es una construcción de los individuos que se puede lograr sólo después de un proceso de maduración intelectual y se produce alrededor de los 12 años. Los alumnos de Secundaria están cerca de esa edad, por ello perciben con fuerza esta componente. Piaget e Inhelder señalaron que a partir de la integración de la totalidad de resultados posibles gracias a las operaciones combinatorias, los individuos descubren las distribuciones y construyen la idea de probabilidad (p. 234); es entonces que pueden predecir resultados. Los conocimientos que los estudiantes adquieren en los cursos de probabilidad fortalecen esta construcción que, sin embargo, resulta muy limitada pues niega la aleatoriedad y se aplica de manera determinista “ocurrirá 10, 10, 10, 10, 10”.

Un tercer momento se caracteriza por la adquisición de un pensamiento que sintetiza la aleatoriedad con la estructura. En las tareas de predicción, un individuo en este nivel renuncia a predecir un resultado determinado sin, por ello, regresar a que “cualquier resultado puede ocurrir”. El conocimiento de la estructura de la situación le permite reducir el universo de lo posible a un evento razonable desechando eventos cuya probabilidad sea casi nula.

En las tareas de identificación de resultados plausibles (10 y 11) fue más fácil identificar la opción que considera la variabilidad que en la tarea de predicción (6). Esto se explica porque aquellas no requieren de la renuncia de la idea de que predecir es dar un resultado determinado, pues al haberse realizado las experiencias el resultado está determinado y sólo se tienen que buscar razones para desechar las opciones.

En conclusión, concebir la variación implica conocer la estructura (la media, la distribución uniforme, etc.) y tener en cuenta la aleatoriedad, dispersión o irregularidad del fenómeno. Estos dos aspectos se integran proponiendo un intervalo alrededor de los valores esperados cuya longitud se determina de acuerdo a la probabilidad con la que se quiere hacer la predicción, y que, por lo tanto, se establezca de antemano.

Este modelo de la evolución del pensamiento hacia la variación que se hace evidente en las situaciones de predicción de azar puede ser útil para examinar el progreso de los razonamientos de los estudiantes en su comprensión de fenómenos sociales y naturales distintos de los de juego.

Para el futuro, los autores de este estudio encuentran conveniente continuar un programa de investigación sobre variación que contenga los siguientes aspectos:

1. Realizar estudios empíricos en el aula para entender cómo y en qué aspectos la instrucción puede promover un desarrollo óptimo de las nociones de variación de los estudiantes.
2. Estudiar el pensamiento de los profesores sobre la variación y la instrucción que requieren para superar sus dificultades.
3. Explorar cómo utilizar la tecnología para propiciar el tránsito de las concepciones uni-estructurales de la variación a las multi-estructural y relacional.

Este trabajo fue financiado por CONACYT (Consejo Nacional de Ciencia y Tecnología) de México, bajo el proyecto: 45063-H.

REFERENCIAS BIBLIOGRÁFICAS

- Batanero, C., (2001). *Didáctica de la Estadística*. Granada, España: Departamento de Didáctica de las Matemáticas, Universidad de Granada.
- Biggs, J.B.; Collis, K.F., (1991). Multimodal learning and the quality of intelligence behavior. H. A. Rowe (Ed.). *Intelligence, reconceptualization and measurement*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Borel, E., (1971). *Las probabilidades y la vida*. Barcelona, España: Oikos-Tau S.A. Ediciones

- Kahneman, D. & Tversky, A. (1982) Subjective probability: A judgment of representativeness. En D. Kahneman; P. Slovic; A. Tversky, *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Piaget, J.; Inhelder, B. (1975). *The origin of the Idea of Chance in Children*. New York, USA: W W Norton & Company Inc.
- Reading, C.; Lawrie, C. (2004). Using SOLO to analyze group responses. In Hoines, M; Fuglestad, A. *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education*, Bergen. Norway.
- Sánchez, E.; Mercado, M. (2006) Notions of variability in chance settings. En Novotná, J., Moraová, H., Krátká, M. y Stehlíková, N. (Eds.). *Proceedings 30th Conference for the International Group for the Psychology of Mathematics Education*. Prague, Czech Republic. Vol. 5, pp. 25-32
- Shaughnessy, J.M., (1997). Missed Opportunities in Research on the Teaching and Learning of Data and Chance. En F. Biddulph; K. Carr (eds.) *People in Mathematics Education*, Mathematics Education Research Group of Australia, Waikato, New Zealand, Vol. 1, pp. 177 –197.
- Shaughnessy, M.; Watson, J.; Moritz, J.; & Reading, Ch., (1999), School Mathematics Students' Acknowledgement of Statistical Variation. Paper presented in the *77th Annual NCTM Conference*, San Francisco, CA. USA.
- Shaughnessy, J.M., Canada, D., & Ciancetta, M. (2003). Middle school students' thinking about variability in repeated trials: A cross-task comparison. En N.A. Pateman, B.J. Dougherty, & J.T. Zilliox (Eds.), *Proceedings of the 27th annual conference of the International Group for the Psychology of Mathematics Education held jointly with the 25th conference of PME-NA* (Vol. 4, pp. 159–165). Honolulu, HI: Center for Research and Development Group, University of Hawaii
- Watson, J., Kelly, B. (2003). Statistical variation in chance settings. En N. A. Pateman; B. J. Dougherty; J. Zilliox. *Proceedings of the 2003 Joint Meeting of PME and PMENA*. Honolulu, Hi. USA.
- Watson, J., Kelly, B., Callingham, R., Shaughnessy, M. (2003). The measurement of school students' understanding of statistical variation. *International Journal of Mathematical Education Science and Technology*, 34, No. 1, pp. 1 –19.