

Los errores en la resolución de situaciones problemáticas de inferencia estadística como principal obstáculo en la didáctica de la estadística

Students' errors in the resolution of problematic situations in statistical inference as a major obstacle in the didactics of statistics.

Teresita Terán y Jesica Ciminari

Universidad Nacional de Rosario, Argentina

Resumen

El objetivo de esta investigación fue evaluar los errores que cometieron los alumnos en el momento de aplicar regresión lineal simple y test de hipótesis, en la asignatura bioestadística de dictado cuatrimestral en el segundo año de la carrera de medicina veterinaria de la Universidad Nacional de Rosario, con el fin de poder considerarlos y poner mayor énfasis en la corrección de los mismos cuando se explique el tema, disminuyendo de esta manera los errores cometidos por los alumnos. Como docentes de estadística en una carrera donde ésta cumple un rol instrumental reflexionar sobre nuestras propias prácticas docentes analizando las dificultades y los errores cometidos por los alumnos, permite mejorar el proceso de enseñanza y aprendizaje, en beneficio del alumno en pos de lograr un aprendizaje significativo.

Palabras clave: Regresión lineal simple, test de hipótesis, didáctica, estadística.

Abstract

The objective of this work was to evaluate the errors committed by the students at the time of applying simple linear regression and hypothesis test, in the subject biostatistics dictated in the second year of the career of veterinary medicine of the National University of Rosario, in order to be able to consider them and place greater emphasis on correcting them when the subject is explained, thus reducing the mistakes made by the students. As teachers of Statistics in a career where it plays an instrumental role reflect on our own teaching practices analyzing the difficulties and mistakes made by students, allows to improve the teaching and learning process, for the benefit of the student in order to achieve meaningful learning.

Keywords: Simple linear regression, hypothesis test, didactics, statistics

1. Introducción

En nuestros días la estadística es una de las ciencias más útiles e influyentes en la mayoría de los campos del conocimiento, pues ofrece múltiples posibilidades de desarrollo y aplicación. Es una herramienta que abunda en literatura científica y se aplica de manera imprescindible en múltiples ámbitos de la vida científica y cotidiana. Involucra la vida de todas las personas, desde las encuestas de opinión, hasta los ensayos clínicos; de esta forma, el conocimiento de nociones básicas de estadística, se torna sumamente importante al momento de poder entender la realidad que nos rodea; sin embargo, su importancia no condice con la importancia que se le da en la educación de los jóvenes.

Resulta necesario tener en cuenta a la estadística como materia de enseñanza en la mayoría de las sociedades, y recalcar la importancia de la didáctica estadística; no sólo para que los alumnos puedan asimilar los contenidos sino que no se convierta en un

aprendizaje memorístico donde el alumno será incapaz de aplicarlo en su futura vida profesional. Pero ello requiere que el profesor sea consciente de las dificultades y errores de sus estudiantes en los temas que imparte.

Consecuentemente, el objetivo de este trabajo fue evaluar los errores que cometieron los alumnos en el momento de aplicar dos temas importantes de la inferencia estadística, la regresión lineal simple y el test de hipótesis, en la asignatura bioestadística de dictado cuatrimestral en el segundo año de la carrera de medicina veterinaria de la Universidad Nacional de Rosario, con el fin de poder considerarlos y poner mayor énfasis en la corrección de los mismos cuando se explique el tema, disminuyendo de esta manera los errores cometidos por los alumnos.

A continuación se exponen los fundamentos, método y resultados de esta evaluación.

2. Fundamentos

Gracias a la abundante investigación en didáctica de la estadística en las últimas décadas, y algunos artículos de destacados profesionales estadísticos se ha producido un cambio de paradigma en la conceptualización de la enseñanza de la estadística. En general se señala que una reforma en la enseñanza ha evolucionado a partir de los avances tecnológicos, de la identificación y precisión de características del pensamiento estadístico y de las ‘grandes ideas’ que sustentan a la estadística.

Según Batanero (2000) las razones de este interés hacia la enseñanza de la Estadística han sido repetidamente señaladas por diversos autores, desde comienzos de la década de los ochenta. Las diferentes razones que se mencionan en este artículo son las siguientes:

- El conocimiento estadístico contribuye al desarrollo personal, y al razonamiento crítico. El estudiante se habitúa a valorar la evidencia para tomar sus decisiones. Además, adquiere un sentido de los métodos y razonamientos que permiten transformar los datos para resolver problemas y efectuar predicciones.
- La estadística es parte de la cultura hoy día; por lo que se habla de cultura estadística. Todos los ciudadanos necesitan la capacidad de lectura e interpretación de tablas y gráficos estadísticos que con frecuencia aparecen en los medios informativos. Para orientarse en el mundo actual, ligado por las telecomunicaciones e interdependiente social, económica y políticamente, es preciso interpretar una amplia gama de información sobre los temas más variados.
- Es un instrumento en la vida profesional ya que en muchas profesiones se precisan unos conocimientos básicos del tema. La estadística es indispensable en el estudio los fenómenos complejos, en los que hay que comenzar por definir el objeto de estudio, y las variables relevantes, tomar datos de las mismas, interpretarlos y analizarlos.
- Ayuda a comprender otros temas del curriculum, tanto de la educación obligatoria como posterior.

La autora sugiere que aunque la estadística cuenta con una axiomática satisfactoria, sigue siendo sujeto de numerosas discusiones sobre la interpretación de conceptos básicos, como la independencia o la probabilidad. Ello se debe a los problemas filosóficos que se mezclan con los conceptos estadísticos. Por ello el profesor debe conocer esta problemática. De otro modo no podrá comprender algunas dificultades de sus estudiantes, quienes necesitan materializar en ejemplos concretos los conceptos y

modelos matemáticos. Afortunadamente, la investigación sobre la enseñanza de la estadística ha crecido mucho en los últimos años, debido a su importancia, ampliamente reconocida, en la formación general del ciudadano.

Cuando realizan un curso de estadística, algunos estudiantes dan un significado particular a los conceptos que tratamos de enseñarles; hablamos entonces de concepción incorrecta (Godino, Batanero y Font, 2004). La problemática que se plantea al profesor es que algunas de estas concepciones sirven para resolver problemas sencillos y entonces el profesor no se da cuenta de que son incorrectas. Pero al pasar a un problema más complejo se cometen errores. En estas circunstancias se habla de la existencia de un obstáculo cognitivo que puede explicar la existencia de errores y dificultades especiales. Brousseau (1983) describe las siguientes características de los obstáculos, resumidas también en Godino, Batanero y Font (2004):

- Un obstáculo no es sinónimo de falta de conocimiento. Por el contrario es un conocimiento que el alumno usa para resolver problemas familiares, dentro de un contexto. Cuando se usa este conocimiento fuera de este contexto genera respuestas incorrectas.
- Aunque se den cuenta de su error, el obstáculo tiene un carácter persistente, pues el alumno tiene dificultad para superarlo y después de haber notado su inexactitud, continúa manifestándolo, de forma esporádica.
- Generalmente en la historia de la matemática se ha presentado el mismo obstáculo similar en el desarrollo de los conceptos.

Un obstáculo es una concepción que ha sido en principio eficiente para resolver algún tipo de problema pero que falla cuando se aplica a otro. Viene a ser una barrera para un aprendizaje posterior. Se revela por medio de los errores específicos que son constantes y resistentes.

Otras dificultades experimentadas por los estudiantes se deben a una falta del conocimiento básico necesario para una comprensión correcta de un concepto o procedimiento dado. Rico (1995) establece que, al estudiar los errores, de acuerdo con las dificultades encontradas por los alumnos, se debiera reconocer que los errores también son función de otras variables del proceso educativo: el profesor, el currículo, el entorno social en que se enmarca la escuela, el medio cultural y sus relaciones, así como las posibles interacciones entre estas variables.

3. Antecedentes

Batanero (2000) afirma que la comprensión de un concepto no puede reducirse a conocer las definiciones y propiedades, sino que debe abarcar: a) el reconocimiento de los problemas donde debe emplearse el concepto, las notaciones y palabras con que se lo denota y en general todas sus representaciones, b) la habilidad operatoria en los diferentes algoritmos y procedimientos relacionados con el concepto, y c) la capacidad de argumentar y justificar propiedades, relaciones y soluciones de problemas. Asumiendo que los errores de los estudiantes están basados en conocimientos y experiencias previas, pudiendo tener diferentes causas que los motivan (dificultades didácticas, epistemológicas, cognitivas, de actitudes, entre otras), su detección y análisis se convierte en uno de los objetivos de la investigación.

La regresión lineal simple y los test de hipótesis son contenidos de estadística que ya forman parte del currículum de enseñanza obligatoria. Su inclusión en dicho nivel posibilita el tratamiento estadístico de datos bivariados y el análisis de la relación que puede existir entre las variables consideradas. La investigación previa ha señalado numerosos errores en los dos temas.

Respecto al contraste de hipótesis, la principal dificultad es comprender los conceptos de nivel de significación y valor p ; además muchos estudiantes suponen que el contraste de hipótesis es un tipo de prueba probabilística que finaliza con el cálculo de la probabilidad de una hipótesis. (Vallecillos y Batanero, 1997). También se confunden el papel de las hipótesis nula y alternativa o las regiones crítica y de aceptación, se producen errores al manejar los cálculos relacionados y en la interpretación de los resultados (Batanero, Vera y Díaz, 2012).

En relación con el estudio de la correlación y la regresión, Sánchez-Cobo, Estepa y Batanero (2000) definen diferentes concepciones incorrectas sobre la correlación, como la concepción lineal (admitir solo la relación lineal) o causal (confundir correlación y causalidad). Por su parte, Estepa (2008) describe la dificultad que tienen los estudiantes para pasar de una representación de los datos bivariantes a otros (por ejemplo, de los datos al gráfico o tabla y viceversa).

4. Método

Como se ha indicado, el objetivo de este trabajo fue evaluar los errores que cometieron los alumnos en los temas de la regresión lineal simple y el test de hipótesis, dentro de la asignatura bioestadística de la carrera de medicina veterinaria de la Universidad Nacional de Rosario, con el fin de poder considerarlos y poner mayor énfasis en la corrección de los mismos cuando se explique el tema, disminuyendo de esta manera los errores cometidos por los alumnos. Es de destacar que los errores no aparecen por azar, sino que surgen en un marco conceptual consistente, basado sobre conocimientos adquiridos previamente, y todo proceso de instrucción es potencialmente generador de errores, debido a diferentes causas, algunas de las cuales se presentan inevitablemente.

Se analizaron 275 evaluaciones finales correspondientes a los períodos 2014, 2015 y 2016 y 2017 de Bioestadística, donde se incluía en la parte práctica los temas Regresión Lineal Simple y Test de Hipótesis.

Luego de realizar un análisis de contenido de la situación problemática correspondiente a ambos temas se observó un total de 132 errores clasificados en nueve grupos para el tema de regresión lineal simple y 145 errores cometidos por los alumnos en el tema test de hipótesis, que también se clasificaron en nueve grupos.

La clasificación de los errores encontrados en el tema regresión lineal simple se realizó a partir de encontrar las siguientes situaciones:

- Intercambio en las variables X e Y : como sabemos en regresión lineal simple existe una variable independiente (X) y una dependiente (Y). Se ha observado que se tomaba en forma invertida las dos variables dadas en el ejercicio. Interpretación errónea de los coeficientes de regresión (b_0 y b_1): cada uno de los coeficientes en regresión lineal tienen su significado particular. Aquí se ha encontrado varios tipos de errores con respecto a los coeficientes. En algunos casos no han sabido interpretarlos en términos del problema, sólo en forma teórica; en otros se han invertido las interpretaciones y algunos otros no han

sabido interpretarlos directamente, cualquiera de estos errores se han incluido en esta categoría.

- Mal planteo del modelo de regresión: en esta categoría se incluyó aquellos casos donde no se ha planteado el modelo de regresión lineal simple o lo han hecho en forma errónea.
- Interpolación fuera del campo de la variable independiente: se les pide a los alumnos que una vez obtenido el modelo de regresión lineal simple estimado, realicen predicciones tomando algún valor de la variable independiente que ellos consideren interesante. Se ha observado que toman valores de la variable que se encuentran fuera del campo de la variación de la misma.
- Interpretación incorrecta en términos del problema: la interpretación de la recta de regresión lineal simple se hace bien en forma teórica pero en el momento de hacerlo en términos del problema, o bien no lo hacen, o lo hacen en forma incorrecta.
- Error en la elección de la distribución de probabilidad: en los test de hipótesis correspondientes para los parámetros del modelo de regresión lineal β_0 y β_1 , se toma en forma errónea la distribución de probabilidad de la variable pivotal.
- Mal interpretación de la conclusión: esta categoría también está asociada a los test de hipótesis correspondientes para los parámetros del modelo de regresión lineal β_0 y β_1 . Si bien el test de hipótesis se resuelve de manera correcta, en el momento de interpretar la conclusión no lo interpretan en términos del problema o bien lo hacen de manera incorrecta.
- Error al graficar la recta de regresión: en todos los ejercicios de regresión lineal simple, en uno de los puntos del ejercicio se le pide a los alumnos que una vez obtenida la recta de regresión la grafiquen. Se ha encontrado que no saben graficarla o bien la grafican de manera incorrecta, invirtiendo los ejes o tomando mal las unidades de medidas en los mismos.
- Error en analizar los datos como Correlación: al dar un ejercicio con dos variables sin especificar cuál es la variable dependiente y cual la variable independiente, es frecuente que interpreten al ejercicio como un análisis de correlación en lugar de regresión lineal simple. Lo cual se toma como erróneo porque se marca de manera adecuada que la variable independiente toma valores puntuales que son asignados por el investigador.

La clasificación de los errores encontrados en el tema test de hipótesis se realizó a partir de encontrar las siguientes situaciones:

- Planteo incorrecto de las hipótesis nula y alternativa: En esta categoría se han incluido aquellos casos donde se observa que se plantea incorrectamente la hipótesis nula, o la alternativa, o ambas. Se han observado casos donde invierten las hipótesis, casos donde se plantean con la estadística y no con el parámetro y otros donde no lo realizan directamente.
- Confusión entre parámetro y estadístico: aquí se han incluido los casos donde toman en forma equivocada los datos proporcionados de los parámetros y de las estadísticas, no pudiendo distinguir cuál es en cada caso. Esto lleva a calcular de manera errónea los cálculos posteriores del ejercicio.

- Elección y/o escritura incorrecta del estadístico de prueba: aquí como bien la categoría lo indica se incluyeron los casos donde se elige de manera incorrecta el estadístico de prueba a emplear, o bien, lo eligen de manera correcta pero hay algún error dentro del mismo.
- Error en los cálculos: se han incluido todos los casos donde se observa que se resuelve de manera correcta lo solicitado por el ejercicio pero no llegan al resultado correcto sólo por un error de cálculo.
- Búsqueda errónea del valor de tabla: en esta categoría se han observado dos situaciones. Una es que buscan el valor en la tabla correcta pero de manera errónea, y el segundo es que utilizan otra tabla directamente.
- Error en la elección de la distribución de probabilidad: aquí sucede que eligen de manera incorrecta la distribución de probabilidad de la variable pivotal, o bien, la eligen correctamente pero no escriben de manera correcta sus parámetros y grados de libertad correspondientes.
- Incorrecta regla de decisión del test de hipótesis: es esta categoría se han incluido casos donde escriben de manera incorrecta la regla de decisión teórica y casos donde la escriben bien pero en el momento de aplicarla con los resultados obtenidos lo hacen de manera incorrecta.
- Interpretación errónea en términos de la situación problemática: Se arma esta categoría porque se observa que el desarrollo del test de hipótesis se hace en forma correcta pero en el momento de concluir en términos de problema no hacen o bien, no interpretan coherentemente de acuerdo al ejercicio planteado.
- Mal empleo de la región de rechazo, error de tipo I y II: en esta categoría se han incluido varias situaciones que se han observado. Se elige la región de rechazo opuesta, dibujan mal la región de rechazo en el gráfico, interpretan de manera errónea el error de tipo I, o II, o ambos.

5. Resultados y discusión

A continuación, en la Tabla 1 se presenta los tipos de errores más frecuentes y los porcentajes de los alumnos correspondiente al tema regresión lineal simple y en la Tabla 2 los tipos de errores más frecuentes y sus porcentajes correspondiente al tema test de hipótesis. Cabe destacar, que un alumno puede cometer más de un error.

Tabla 1: Clasificación de los errores cometidos por los alumnos en regresión lineal simple

Clasificación de los errores	% de alumnos
Intercambio en las variables X e Y	34,2
Interpretación errónea de los coeficientes de regresión (b_0 y b_1)	28,9
Mal planteo del modelo de regresión	26,3
Interpolación fuera del campo de la variable independiente	16,7
Interpretación incorrecta en términos del problema	12,3
Error en la elección de la distribución de probabilidad	10,5
Mal interpretación de la conclusión	3,2
Error al graficar la recta de regresión	2,5
Error en analizar los datos como Correlación.	1,5

Tabla 2: Clasificación de los errores cometidos por los alumnos en test de hipótesis.

Clasificación de los errores	% de alumnos
Planteo incorrecto de las hipótesis nula y alternativa	24,2
Confusión entre parámetro y estadístico	18,9
Elección y/o escritura incorrecta del estadístico de prueba	8,3
Error en los cálculos	3,7
Búsqueda errónea del valor de tabla	11,3
Error en la elección de la distribución de probabilidad	4,5
Incorrecta regla de decisión del test de hipótesis	21,2
Interpretación errónea en términos de la situación problemática	33,5
Mal empleo de la región de rechazo, error de tipo I y II	7,5

Como se puede observar en la Tabla 1, los errores más significativos cometidos por los alumnos en la resolución de los problemas de regresión lineal simple fueron: intercambio de las variables X y Y, error que también fue encontrado en la investigación de Sánchez-Cobo et al. (2000). Fue también muy frecuente la interpretación errónea de los coeficientes de regresión (b_0 y b_1), mal planteo del modelo de regresión, interpolación fuera del campo de la variable independiente e interpretación incorrecta en términos del problema.

En cuanto al tema test de hipótesis los errores con mayor porcentaje de incidencia cometidos por los alumnos son: el planteo de las hipótesis, la diferenciación entre parámetro y estadística, la regla de decisión adoptada y la interpretación incorrecta en términos del problema.

En el planteo de las hipótesis nula y alternativa se observa una gran dificultad de los alumnos, sobre todo en la diferenciación entre estas hipótesis, al igual que ocurrió en el trabajo de Vallecillos y Batanero (1997). Generalmente las hipótesis nula y alternativa son intercambiadas, lo que conlleva la provocación de errores de tipo procedimental, ya que al tomar como nula la hipótesis alternativa implica, en la práctica, tomar como nula una hipótesis compuesta, lo que bloqueará o impedirá a los estudiantes continuar con el problema.

En cuanto a la dificultad que presentan los alumnos en distinguir de forma correcta un parámetro de un resumen estadístico, identificado igualmente en Vallecillos y Batanero, se puede decir que aunque los estudiantes distinguen perfectamente la media de la muestra y la de la población, manifiestan una falta de comprensión de los distintos papeles que juegan, en el contraste de hipótesis, el estadístico muestral y el parámetro poblacional, confundiendo en muchas oportunidades su simbología.

Se observa una falta de apreciación de la media muestral como variable aleatoria y de su distribución. Por último, los porcentajes que se observan en la incorrecta regla de decisión y conclusión del test de hipótesis junto con su interpretación en contexto se deben en general a todo lo analizado en forma precedente y también a que los alumnos muchas veces resuelven el ejercicio de test de hipótesis de manera mecánica, sin interpretar ni entender que es lo que se está queriendo probar, con lo cual una vez tomada la decisión, no saben cómo concluir ni hacerlo en términos del problema, sólo se sabe si se rechazó o no la hipótesis nula.

6. Reflexiones finales

La identificación de los puntos que presentan dificultad y de los errores detectados en el proceso de enseñanza–aprendizaje (incluso después del mismo) son preocupaciones básicas de la educación estadística. A partir de su conocimiento es posible diseñar actividades didácticas útiles para el profesor, que sean adecuadas para superar las dificultades y favorezcan la comprensión. Según Batanero (2001), la comprensión personal de un concepto es la captación de su significado.

Dado que el significado de un objeto se concibe en relación con los contextos institucionales, la comprensión de un concepto por un sujeto, en un momento y circunstancias dadas, implicará la adquisición de los distintos elementos que componen su significado institucional. Por ello, se considera que el análisis de enseñanza, como caracterización del significado institucional del concepto, se encuentra íntimamente relacionado con el aprendizaje. De esta manera, el alumno conoce o comprende un determinado concepto si hay un ajuste entre el significado institucional y el personal que elabora.

Como docentes de estadística en una carrera donde ésta cumple un rol instrumental, reflexionar sobre nuestras propias prácticas docentes analizando las dificultades y los errores cometidos por los alumnos, permite mejorar el proceso de enseñanza y aprendizaje, en beneficio del alumno tendiendo a que disminuyan los errores que comete en pos de lograr un aprendizaje significativo.

Referencias

- Batanero, C. (2000) ¿Hacia dónde va la educación estadística? *Blaix*, 15, 2-13.
- Batanero, C. (2001). *Didáctica de la estadística*. Granada: Grupo de Investigación en Educación Estadística.
- Batanero, C., Godino J., Green, D., Holmes P. y Vallecillos A. Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics Education in Science and Technology*, 25(4), 527-547
- Batanero, C., Vera, O. y Díaz, C. (2012). Dificultades de estudiantes de psicología en la comprensión del contraste de hipótesis. *Números*, 80, 91-101
- Brousseau, G. (1983). Les obstacles épistémologiques et les problèmes en mathématiques. *Recherches en Didactique des Mathématiques*, 4(2), 165-198.
- Estepa, A. (2008). Interpretación de los diagramas de dispersión por estudiantes de Bachillerato. *Enseñanza de las Ciencias*, 26(2), 257-270.
- Godino, J., Batanero, C. y Font, V. (2003). *Fundamentos de la enseñanza y aprendizaje de la Matemática para maestros*. Universidad de Granada. Disponible en: <http://www.ugr.es/local/jgodino/edumat-maestros/>
- Rico, L. (1995). Errores y dificultades en el aprendizaje de las Matemáticas. En L. Rico y P. Gómez (Ed.) *Educación Matemática*. México: Grupo Editorial Iberoamérica.
- Sánchez-Cobo, F. T., Estepa, A. y Batanero, C. (2000). Actividades de traducción en la estimación de la correlación. *Enseñanza de las Ciencias*, 18(2). 297-310
- Vallecillos, A., y Batanero, C. (1997). Aprendizaje y enseñanza del contraste de hipótesis: Concepciones y errores. *Enseñanza de las Ciencias*, 15(2), 180-197.