

**ERRORES COMETIDOS POR LOS ALUMNOS EN LA ASIGNATURA
ESTADÍSTICA Y BIOMETRÍA DE LA CARRERA INGENIERÍA
AGRONÓMICA, UNIVERSIDAD NACIONAL DE CATAMARCA (2008)**

Rojas, Ilda Rosa y Ovejero, Diana

Cátedra de Estadística y Biometría, Facultad de Ciencias Agrarias. Universidad Nacional de Catamarca.
Avda. Belgrano y Maestro Quiroga (4700) San Fernando del Valle de Catamarca. Argentina. Tel/fax: 03833-430504.
E-mail: ildarojas@argentina.com.ar, dianaove@yahoo.com.ar

**ERRORS IN STATISTICS AND BIOMETRY MADE BY STUDENTS IN
SECOND YEAR OF AGRONOMY ENGINEERING AT THE NATIONAL
UNIVERSITY OF CATAMARCA (2008).**

SUMMARY

The students of the Statistics and Biometry course in the 2nd year of the Agronomic Engineering career at the National University of Catamarca have problems in the assimilation/learning and interpretation of statistical concepts and their practice. The need arises to analyse the errors in order to find a solution. The purpose of this study is to identify the errors, investigate the causes, and make proposals that could improve the teaching-learning process. The research carried out is of a quantitative-qualitative type, so statistical techniques and content analysis were used. Information was surveyed by means of the analysis of the first partial exam given to students in 2008.

KEY WORDS: errors, statistical concepts, teaching-learning.

RESUMEN

Debido a que se perciben, en los alumnos que cursan la asignatura Estadística y Biometría correspondiente al segundo año de Ingeniería Agronómica de la Universidad Nacional de Catamarca, problemas en la asimilación e interpretación de conceptos estadísticos y en su aplicación práctica surge la necesidad analizar los errores cometidos a fin de buscar

soluciones. Para lograrlo, se plantea como objetivo: identificar los errores, investigar sus causas y realizar propuestas que mejoren el proceso enseñanza-aprendizaje. La investigación es cuantitativa y cualitativa, utilizando técnicas estadísticas y análisis de contenidos respectivamente. El relevamiento de la información se realizó mediante el análisis del primer examen parcial tomado a los alumnos del año 2008.

PALABRAS CLAVES: errores, conceptos estadísticos, enseñanza – aprendizaje.

INTRODUCCION

Para que el estudiante sea capaz de relacionar los conceptos, ideas, leyes, proposiciones, etc. que dispone en su mapa cognitivo con las que se le están presentando como nuevas, es preciso que los mismos hayan sido adquiridos de una forma significativa y no arbitraria o literal (Moreira, 2000).

Averiguar los conocimientos previos que el alumno presenta antes de iniciar el aprendizaje y que son potencialmente relacionables con los nuevos contenidos a aprender, es uno de los factores fundamentales que influyen en el progreso de los alumnos. Un principio ampliamente asumido en psicología educativa es el enunciado por Ausubel, Novac y Hanesian (1983) quienes establecen que el factor más importante que influye en el aprendizaje es lo que el alumno ya sabe por lo tanto resulta necesario averiguar esos conocimientos y enseñar consecuentemente

El cognitivismo sostiene que la mente del alumno no es una página en blanco: el alumno tiene un saber anterior, y estos conocimientos anteriores pueden ayudar al nuevo conocimiento, pero a veces son un obstáculo en la formación del mismo. El conocimiento nuevo no se agrega al antiguo, sino que lucha contra él y provoca una nueva estructuración del conocimiento total. Los errores cometidos por los alumnos en matemática son una manifestación de esas dificultades y obstáculos propios del aprendizaje, y se acepta unánimemente que es necesaria la detección y análisis de los mismos, y su utilización positiva en una suerte de realimentación del proceso educativo (Del Puerto, Minnaard y Seminara, 2004)

Un obstáculo es una concepción que ha sido en principio eficiente para resolver algún tipo de problemas pero que falla cuando se aplica a otro. Debido a su éxito previo se resiste a ser modificado o a ser rechazado: viene a ser una barrera para un aprendizaje posterior. Se revela por medio de los errores específicos que son constantes y resistentes. Para superar tales obstáculos se precisan situaciones didácticas diseñadas para hacer a los alumnos conscientes de

la necesidad de cambiar sus concepciones y para ayudarlos a conseguirlo (Matemática Par@educ.ar, 31/08/2010)

Brousseau (1983) citado por Batanero, Godino, Green et al (1994) da las siguientes características de los obstáculos:

- Un obstáculo es un conocimiento, no una falta de conocimiento. El alumno utiliza este conocimiento para producir respuestas adaptadas a un cierto contexto que encuentra con frecuencia. Cuando se usa este conocimiento fuera de este contexto genera respuestas incorrectas. Una respuesta universal exigirá un punto de vista diferente.
- El alumno resiste a las contradicciones que el obstáculo le produce y al establecimiento de un conocimiento mejor. Es indispensable identificarlo e incorporar su rechazo en el nuevo saber.
- Después de haber notado su inexactitud, continúa manifestándolo, de forma esporádica.

Además identifica tres tipos de obstáculos:

- *Obstáculos ontogénicos*: son debidos a las características del desarrollo del alumno. Por ejemplo, para comprender la idea de probabilidad se requiere el razonamiento proporcional.
- *Obstáculos didácticos*: resultan de las elecciones didácticas hechas para establecer la situación de enseñanza. Por ejemplo, la introducción de un nuevo simbolismo tal como: $(\sum x_i)/n$ cuando los estudiantes necesitan trabajar con ejemplos concretos.
- *Obstáculos epistemológicos*: Relacionados intrínsecamente con el propio concepto y conteniendo parte del significado del mismo. Por ejemplo, las circularidades que se presentan en las diferentes definiciones del significado de la probabilidad (clásica, frecuencial) que mostraron en su día la necesidad de una definición axiomática.

Según estos mismos autores 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 “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”.

Como señalan Del Puerto, Seminara y Minnaard (2004) los errores son posibles en todo proceso de adquisición y consolidación de conocimientos. Su identificación permite demarcar los distintos componentes implicados en la comprensión de un concepto. Su análisis es una herramienta fundamental para relevar el estado de conocimiento y brinda, además, una señal de hacia dónde se debe reorientar el proceso de enseñanza-aprendizaje con el fin de mejorar los

resultados. Es una circunstancia que permite al alumno argumentar, discutir y rever sus conocimientos a fin de lograr una mejor comprensión.

Según Mulhern (1989) citado por Rico (1995) la mayor parte de los investigadores y especialistas señalan como características generales de los errores cometidos por los alumnos las siguientes:

- Surgen, con frecuencia, de manera sorprendente en el sentido de que se han mantenido ocultos por algún tiempo para el profesor
- Son persistentes y difíciles de superar, ya que su corrección puede requerir de una reorganización fundamental de los conocimientos en el alumno.
- Pueden ser sistemáticos o por azar
 - los sistemáticos son más frecuentes y revelan los procesos mentales que han llevado al alumno a una comprensión equivocada
 - los errores por azar reflejan falta de cuidado. Son de relativa importancia
- Los errores ignoran el significado y por lo tanto las respuestas que son obviamente incorrectas no se ponen en cuestión (muchas veces los alumnos no toman conciencia del error ya que no comprenden acabadamente el significado de los símbolos y conceptos con que trabajan).

Los errores sistemáticos favorecen la creación de patrones de comportamiento equivocados en la ejecución de las tareas. Son patrones consistentes de errores. La consistencia es en dos niveles, individual: los sujetos muestran regularidad en el modo de realizar tareas y resolver problemas similares; y de carácter colectivo ya que distintas personas cometen errores semejantes en determinadas etapas de su desarrollo educativo. (Carrión Miranda, 2007)

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.

Esta investigación se hizo bajo el presupuesto que los errores cometidos por los estudiantes son, en su gran mayoría del tipo sistemático

ANTECEDENTES

Varios investigadores han abordado el estudio de los errores cometidos por los alumnos en Estadística y también en Probabilidades.

Errores y dificultades en el aprendizaje de la Estadística que se refieren al uso de representaciones gráficas y tablas de frecuencias.

La destreza en la lectura de los datos y la interpretación e integración de los mismos en una tabla o gráfico, es un componente esencial en el estudio de estadística y una exigencia en la actual sociedad tecnológica.

Elementos y competencias de lectura de gráficos estadísticos

Un aspecto de los gráficos investigado por diversos autores es la competencia en la lectura de los mismos, tarea en la que el estudiante debe realizar la traducción entre lo representado y la realidad.

En “El lenguaje de los gráficos estadísticos” (2009) los autores citan a Curcio, (1987; 1989) para mostrar los convenios de construcción y elementos de los gráficos. Al mismo tiempo mencionan que Friel, Curcio y Bright (2001) además de identificar los elementos estructurales de un gráfico estadístico, describen las competencias relacionadas con el lenguaje gráfico a saber:

- Reconocer los elementos estructurales del gráfico (ejes, escalas, etiquetas, elementos específicos) y sus relaciones. Discernir si cada elemento es o no apropiado en un gráfico particular.
- Evaluar el impacto de cada uno de estos componentes sobre la presentación de la información (por ejemplo, predecir como cambiaría el gráfico al variar la escala de un eje).
- Traducir las relaciones reflejadas en el gráfico a los datos que se representan en el mismo y viceversa.
- Saber elegir el gráfico adecuado al tipo de variable y al tipo de problema.

Niveles de lectura de gráficos

Aparte de las competencias anteriores, algunos autores definen niveles en la lectura crítica de datos. Curcio (1989) citado por Batanero, Godino, Green et al (1994) describe tres niveles distintos de comprensión de los gráficos:

- “Leer los datos”: consiste en una lectura literal del gráfico; no se realiza interpretación de la información contenida en el mismo.

- “Leer dentro de los datos”: incluye la interpretación e integración de los datos en el gráfico; demanda la habilidad para comparar cantidades y el uso de otros conceptos y destrezas matemáticas.

- “Leer más allá de los datos”: requiere que el lector realice predicciones e inferencias a partir de los datos sobre informaciones que no se reflejan directamente en el gráfico.

Arteaga, Batanero, Díaz et al (2009) indican que Friel, Curcio y Bright (2001) amplían la clasificación definiendo un nuevo nivel “Leer detrás de los datos” que consiste en valorar críticamente el método de recolección de datos, su validez y fiabilidad, así como las posibilidades de extensión de las conclusiones.

Errores en la lectura o construcción de gráficos

Además de las capacidades de lectura de los gráficos, varias investigaciones se centran en el análisis de los errores frecuentes en la producción de los mismos.

Li y Shen, (1992) citados por Arteaga, Batanero, Díaz et al (2009) encuentran que los alumnos utilizan polígonos de frecuencias con variables cualitativas, o diagrama de barras horizontales cuando lo correcto es un diagrama de dispersión; otras veces, en un mismo gráfico representan dos variables no relacionadas entre sí. Con referencia a las escalas las dificultades observadas fueron:

- Elección inadecuada de una escala para el objetivo pretendido (por ejemplo no se cubre todo el campo de variación de la variable representada).
- Omisión de las escalas en alguno de los ejes o en ambos ejes.
- Falta de especificación del origen de coordenadas.
- Divisiones insuficientes en las escalas de los ejes.

Arteaga, Batanero, Díaz et al (2009) indican que Lee y Meletiou (2003) detectan alumnos que perciben los histogramas como representación de datos aislados donde cada rectángulo se refiere a una observación particular y no a un intervalo de valores, y que comparan las alturas de los rectángulos (y no sus áreas) cuando tratan de descubrir variaciones en el histograma. En lo que respecta a los gráficos de la caja, Bakker, Biehler y Konold (2004) citados también por Arteaga, Batanero, Díaz et al (2009) advierten que los estudiantes tienen dificultad para interpretarlos posiblemente por estar basados en la mediana y los cuartiles y ser muy diferentes a otros gráficos usados por ellos.

Medidas de Resumen

Medidas de posición central

En “Significado y comprensión de las medidas de posición central” (2000) Batanero hace referencia a que Pollatsek, Lima y Well (1981) se percatan que alumnos universitarios no ponderan adecuadamente los valores al resolver los problemas y en ocasiones usan la media simple, en lugar de la media ponderada. Campbell (1974) citado por Batanero (2000) nota que el alumno tiende a situar la media en el centro del recorrido de la distribución, propiedad que es cierta para distribuciones simétricas, pero cuando la distribución es muy asimétrica la media se desplaza hacia uno de los extremos y es la moda o la mediana valor más representativo del conjunto de datos. Según Batanero (2000) esto no es siempre comprendido por los alumnos quienes invariablemente eligen la media como mejor representante de los datos sin tener en cuenta la simetría de la distribución o la existencia de valores atípicos

Batanero, en “Significado y comprensión de las medidas de posición central” (2000), señala que Carvalho (1998) analiza las producciones escritas de los alumnos cuando resuelven tareas estadísticas detectando los siguientes errores de cálculo en media, mediana y modo

- Modo: Tomar la mayor frecuencia absoluta;
- Mediana: No ordenar los datos para calcular la mediana; calcular el dato central de las frecuencias absolutas ordenadas de forma creciente; calcular la moda en vez de la mediana; equivocarse al calcular el valor central
- Media: Hallar la media de los valores de las frecuencias; no tener en cuenta la frecuencia absoluta de cada valor en el cálculo de la media.

El cálculo de la media se presenta como sencillo pero con frecuencia el algoritmo se aplica sin comprender su significado. En ocasiones como muestra Del Puerto, Seminara y Minnaard (2007) un problema aparentemente “simple” de cálculo de media presentó dificultades para los alumnos debido a la presencia de datos negativos y el 0.

Los términos matemáticos con que se designan los conceptos tienen un significado preciso, pero éste no siempre coincide con el asignado en el lenguaje coloquial. Russell y Mokros (1991) citado por Batanero (2000) clasifican en cuatro categorías los significados incorrectos atribuidos por los estudiantes a la palabra "media": valor más frecuente (confusión con la palabra "modo"), "valor razonable" (significado coloquial del término), "punto medio" (confusión con la mediana) y "algoritmo" (significado restringido: la media se ve sólo como el algoritmo de cálculo).

Respecto a la comprensión de la mediana Batanero (2000) refiere a Barr, (1980) para indicar que los alumnos entienden que la mediana es el centro de "algo" pero con frecuencia no

tienen en claro a qué secuencia numérica se refiere ese "algo". En "Errores y dificultades en la comprensión de los conceptos estadísticos elementales" (1994) los autores expresan que los estudiantes suelen interpretar la mediana como el valor central de los valores de la variable, de las frecuencias o incluso de la serie de datos antes de ser ordenada.

Cobo y Batanero (2000) afirman que el cómputo de la mediana es complejo porque el algoritmo de cálculo es diferente según el número de datos sea par o no y según los datos se presenten en tablas de datos agrupados o sin agrupar, y también por el hecho de que el valor obtenido es diferente, según se aplique uno u otro algoritmo. Todo esto puede generar problemas a los alumnos que están acostumbrados a un único método de cálculo y una única solución para los problemas matemáticos. Situación semejante ocurre en el cálculo de los percentiles (Del Puerto, Seminara y Minnaard, 2007)

Medidas de Dispersión

Las distribuciones de frecuencias pueden tener medias o medianas iguales y sin embargo ser diferentes en cuanto a su variabilidad es por ello que su estudio no puede limitarse al de sus medidas de tendencia central. Para Campbell (1974) citado por Batanero, Godino, Díaz et al (1994) un error frecuente es ignorar la dispersión de los datos cuando se efectúan comparaciones entre dos o más muestras o poblaciones.

Del Puerto, Seminara y Minnaard (2007) dan cuenta de que los alumnos confunden variabilidad relativa con la variabilidad absoluta comparando dos distribuciones mediante el desvío estándar, sin atender a que las medias son distintas. Establecen además que términos como: dispersión, variabilidad, desviación, variación que son claros para el docente no lo son tanto para los alumnos.

En "Errores y dificultades en la comprensión de los conceptos estadísticos elementales" (1994) se cita a Loosen y Cols (1985) para referirse a la experiencia donde estos autores detectan que el concepto intuitivo de variabilidad se equipara al de "no semejanza", es decir, cuánto varían unos valores respecto a otros, más que cuánto varían los valores respecto a un punto fijo.

En el Foro Internacional de Razonamiento, Pensamiento y Literatura Estadístico (SRTL-3) en la Universidad de Nebraska (USA) en julio de 2003, los investigadores reforzaron su creencia de que el concepto de variabilidad es un tópico complejo para entender, aprender y enseñar, y que su comprensión es un componente fundamental en el razonamiento y pensamiento estadístico (Del Puerto, Seminara y Minnaard, 2007).

Probabilidad

Ortiz, Batanero y Serrano (2001) citando a Orton (1990) expresan que hay numerosos aspectos del lenguaje que pueden afectar al aprendizaje de las matemáticas, ya que muchos estudiantes no entienden los términos que se emplean en clase como parte del vocabulario matemático. Lo que reviste un problema no son los términos en sí mismos, sino los conceptos y procesos subyacentes que se están comunicando y el significado que transmiten. Los mismos autores expresan que Pimm (1987) llama la atención sobre el uso de palabras del lenguaje cotidiano, con sentido matemático particular, lo cual puede ocurrir en el caso de los términos probabilísticos.

En “El lenguaje probabilístico en los libros de texto” (2001) los autores invocan a Rotherry (1980) para diferenciar tres categorías de palabras usadas en la enseñanza de las matemáticas:

1. Palabras específicas de las matemáticas que, normalmente, no forman parte del lenguaje cotidiano.
2. Palabras que aparecen en las matemáticas y en el lenguaje ordinario, aunque no siempre con el mismo significado en los dos contextos. Se producen innumerables confusiones cuando el profesor emplea términos “del dialecto matemático” y los alumnos lo interpretan de acuerdo al lenguaje ordinario.
3. Palabras que tienen significados iguales o muy próximos en ambos contextos.

En el caso de la probabilidad la mayoría de los vocablos pertenecen a las dos últimas categorías.

Ortiz, Batanero y Serrano (2001) señalan que en el lenguaje ordinario hay expresiones y referencias al azar y la probabilidad que no siempre coinciden con el significado que le asignan las matemáticas, lo cual dificulta aún más la enseñanza de la probabilidad.

Carmen Díaz en su trabajo “Evaluación de la falacia de la conjunción en alumnos universitarios” (2005) analiza uno de los errores en razonamiento probabilístico que se relaciona con la falta de comprensión de la probabilidad conjunta: “la falacia de la conjunción” la cual consiste en estimar para la probabilidad de la intersección de dos sucesos un valor mayor que la probabilidad simple estimada de uno de los sucesos. La investigación se realiza con finalidad diagnóstica, para conocer los posibles sesgos que presentan estudiantes y tenerlos en cuenta a la hora de la enseñanza del tema.

En “Razonamiento probabilístico en la vida cotidiana: un desafío educativo” (2006) Batanero defiende la necesidad reforzar la formación del razonamiento probabilístico en la

educación primaria y secundaria y proporcionar con ello a los alumnos un instrumento que oriente la acción ante la incertidumbre.

Cañizares y Batanero (1998) opinan que el dominio del cálculo de proporciones no es el único pre requisito para el cómputo adecuado de probabilidades. Ambas tareas son muy diferentes: el resultado de un problema proporcional se refiere a un acontecimiento seguro, mientras que el resultado de un problema probabilístico implica un grado de incertidumbre. En su trabajo concluyen que “... *en general el nivel de razonamiento proporcional de los alumnos es bajo. Ello podría ser un obstáculo para el aprendizaje de las probabilidades, aunque, también la enseñanza de la probabilidad podría ser un contexto muy rico que favoreciese el desarrollo del razonamiento proporcional de estos alumnos*” (pág, 12)

Variable aleatoria y Distribuciones Probabilísticas

A pesar que el concepto de variable aleatoria es fundamental en el desarrollo de distribuciones probabilísticas, distribuciones en el muestreo, pruebas de hipótesis pocos son los investigadores que se han enfocado en el análisis de las dificultades de los alumnos en su aprendizaje. Ruiz, Huerta y Batanero (2006) establecen que el concepto de variable aleatoria no es fácil de aprender ni de enseñar y citan a Oseguera (1994) para indicar que este concepto no está relacionado a las experiencias intuitivas de los alumnos en la educación formal donde con frecuencia se presenta como un preámbulo de las distribuciones probabilísticas. En su estudio exploratorio estos autores detectan en los alumnos falta de percepción de la aleatoriedad en la resolución del problema planteado, tendencia a algebrizar y no contextualizar los procesos relacionados a variables aleatorias y dificultades con la noción formal de variable aleatoria.

En lo referente a distribuciones probabilísticas las pesquisas se centran en la distribución normal y son escasas. Batanero, Tauber y Sanchez (2001) manifiestan que a pesar de la importancia de la distribución normal no han encontrado estudios sistemáticos de los errores y dificultades de aprendizaje referidos a ese tema. Expresan además que las investigaciones previas específicamente centradas sobre la comprensión de la distribución normal son reducidas en número y ponen el estudio de Piaget e Inhelder (1951) sobre el desarrollo espontáneo de la comprensión de la convergencia estocástica en los niños como primer antecedente hallado.

En “Significado y comprensión de la distribución normal en un curso introductorio de análisis de datos” (2001) los autores hacen referencia a que Huck, Cross y Clark (1986) identifican dos concepciones erróneas sobre las puntuaciones normales tipificadas en alumnos universitarios: algunos creen que todas las puntuaciones tipificadas han de tomar un valor comprendido entre -3 y +3, mientras que otros piensan que no hay límite para los valores máximo y mínimo de dichas puntuaciones. También mencionan a Hawkins, Jolliffe y Glickman

(1992) quienes describen errores en la aproximación de una distribución binomial mediante la distribución normal originados por la no diferenciación entre lo discreto y lo continuo.

Batanero, Tauber y Meyer (1999) dan cuenta de que en su investigación encontraron alumnos que consideraron normal a una variable cualitativa, otros que tomaron como normal a variables con poca dispersión y que una proporción significativa de estudiantes utilizó solamente herramientas gráficas para evaluar la normalidad sin referencia a áreas o valores críticos.

MATERIALES Y METODOS

La asignatura Estadística y Biometría posee un régimen cuatrimestral y corresponde al segundo año de la carrera Ingeniería Agronómica de la Universidad Nacional de Catamarca. Para cursarla los alumnos deben tener aprobada Matemática I y regularizada Matemática II (materias correspondientes al primer año de la carrera).

La materia está organizada sobre la base del dictado de clases teórico prácticas. La carga horaria es de seis horas y media semanales: tres horas teóricas y las restantes prácticas (dos horas en la sala de informática y una hora y media en aula). Para regularizarla, los estudiantes deben aprobar tres evaluaciones parciales y una evaluación integral, al final del dictado de la asignatura, que consiste en el análisis de un conjunto de datos donde aplican varias de las temáticas abordadas. Finalmente un examen final oral les permite aprobar la asignatura.

Con el objeto de identificar y analizar los errores más frecuentes cometidos por los alumnos se analizan los exámenes del Primer Parcial de los alumnos que cursan Estadística y Biometría en el año 2008. Los temas evaluados son: Estadística Descriptiva y Análisis Exploratorio: comprende todo aquello que hace a la terminología específica de la materia, como así también a la recopilación y ordenamiento de datos, presentación de la información, descripción y exploración de datos. Probabilidad, Variables aleatorias, Distribuciones Probabilísticas más Conocidas: Abarca los conocimientos básicos de probabilidades y variable aleatoria que son necesarios para temas de inferencia estadística. Incluye, además las distribuciones Binomial, Poisson y Normal. El Primer parcial siempre cuenta con un total de cinco ejercicios distribuidos de la siguiente manera Estadística Descriptiva y Análisis Exploratorio: uno, Probabilidad: uno, Variable Aleatoria y Distribuciones Probabilísticas más Conocidas: tres

A fin de obtener una clasificación que permita detectar cómo surgen los errores a partir de la solución correcta y poder establecer qué factores pueden haber conducido a ellos, se

efectúa un recuento del número de soluciones incorrectas y un análisis de los tipos de errores. Para la tabulación de los datos y la construcción de los gráficos se hace uso del software estadístico InfoStat (2009).

Los ítems considerados a los efectos del análisis y conforme al tema abordado son

- ❖ **Planteo del problema:** el alumno tiene claro a qué se refiere el problema?
- ❖ **Fórmulas:** la fórmula utilizada es la que corresponde? Está bien escrita?
- ❖ **Interpretación de los resultados:** interpreta correctamente el resultado obtenido? Es coherente el resultado obtenido con el problema planteado?.
- ❖ **Conclusión:** a la luz de los resultados obtenidos la conclusión es correcta?
- ❖ **Redacción:** redacta adecuadamente tanto la conclusión como la interpretación de resultados?
- ❖ **Gráficos:** el gráfico seleccionado es el correcto? Está bien construido?
- ❖ **Cálculos:** realiza correctamente los cálculos?
- ❖ **Uso de tabla:** se hace uso correcto de las tablas de probabilidad de las distribuciones estudiadas?

RESULTADOS

1. ESTADÍSTICA DESCRIPTIVA Y ANÁLISIS EXPLORATORIO

TABLA 1: Dificultades que presentan los alumnos en Estadística Descriptiva y Análisis Exploratorio

Ítems	Acción del Alumno			Total
	Responde		No Responde	
	Bien	Mal		
Planteo del problema	56 (54%)	15 (14%)	33 (32%)	104 (100%)
Interpretación	3 (3%)	63 (61%)	38 (36%)	104 (100%)
Gráficos	40 (38%)	27 (26%)	37 (36%)	104 (100%)
Fórmulas	7 (7%)	43 (41%)	54 (52%)	104 (100%)

La mayor cantidad de errores se dan en la interpretación de los resultados (61%) y en la aplicación de fórmulas (41%).

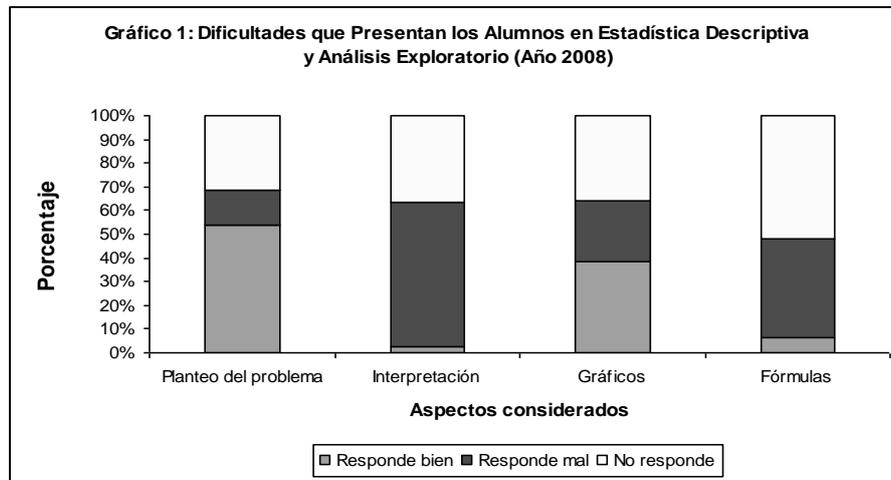


GRÁFICO 1: Dificultades que presentan los alumnos en Estadística Descriptiva y Análisis Exploratorio. (Año 2008).

El **Gráfico 1** hace notoria la dificultad que poseen los alumnos en la interpretación de resultados (97%) y la utilización de fórmulas (93%)

TABLA 2: Relación entre el Planteo del Problema y los otros Ítems ¹

Ítems	Respuestas de los alumnos	Planteo del problema	
		Bien (%)	Mal (%)
Interpretación	Bien	5	0
	Mal	88	93
	No responde	7	7
Total		100	100
Gráficos	Bien	61	26
	Mal	36	47
	No responde	3	27
Total		100	100
Fórmulas	Bien	12	0
	Mal	59	60
	No responde	29	40
Total		100	100

(1) La categoría “no responde “ del Item Planteamiento del problema no se tiene en cuenta dado que se corresponde con “ejercicio sin resolver”

De los alumnos que plantean mal el problema:

- El 93% efectúa una interpretación equivocada de los resultados.
- El 47% hace una selección inadecuada del gráfico o bien lo construye de manera incorrecta y el 26% no realiza acción alguna.

- El 60% tiene problemas con la fórmula de cálculo de las medidas de resumen solicitadas y un 40% no efectúa ningún cálculo.

Más específicamente

- Fallan en la elección del gráfico. A menudo no especifican la escala en los ejes, en otros casos la escala elegida no es la correcta y siempre olvidan el origen de coordenadas al igual que indicar la fuente de donde provienen los datos
- Si bien presentan dificultades para "leer los datos", éstas se acentúan cuando se trata de "leer dentro de los datos" (interpretación e integración de los datos en el gráfico) y "leer más allá de los datos" (predicciones e inferencias a partir de los datos sobre informaciones que no se reflejan directamente en el gráfico).
- Construyen la distribución de frecuencias pero no alcanzan a comprender el significado de la misma. Esto se pone de manifiesto al no interpretar apropiadamente el significado de las frecuencias.

Medidas de Tendencia Central

En cuanto a las medidas de tendencia central, sucede que, si la consigna es calcular las medidas de posición los errores más frecuentes se relacionan con la incorrecta aplicación de la fórmula o procedimiento para obtener el valor correspondiente (o la no aplicación) y con la gran dificultad para traducir dicho valor a los términos del problema. Si la consigna está elaborada en términos del enunciado del problema, como por ejemplo: ¿Cuál es el peso promedio de los estudiantes?, ¿Cuál es el valor de la variable que se presenta un mayor número de veces? o ¿Qué altura es superada por el 50% de las observaciones? No distinguen a qué medida de las estudiadas se hace referencia.

Medidas de dispersión

Al igual que en el caso de las medidas de posición los alumnos tienden a aplicar mal las fórmulas. Además confunden varianza con desviación estándar

2. PROBABILIDAD

En la siguiente tabla se presentan los resultados obtenidos:

TABLA 3: Dificultades que presentan los alumnos en Probabilidad

Ítems	Acción del alumno			Total
	Responde		No responde	
	Bien	Mal		
Planteo del problema	72 (69%)	23 (22%)	9 (9%)	104 (100%)
Fórmulas	58 (56%)	37 (35%)	9 (9%)	104 (100%)
Interpretación de los resultados	27 (26%)	37 (35%)	40 (39%)	104 (100%)

Conclusión	38 (37%)	41 (39%)	25 (24%)	104 (100%)
Redacción	14 (14%)	37 (35%)	53 (51%)	104 (100%)

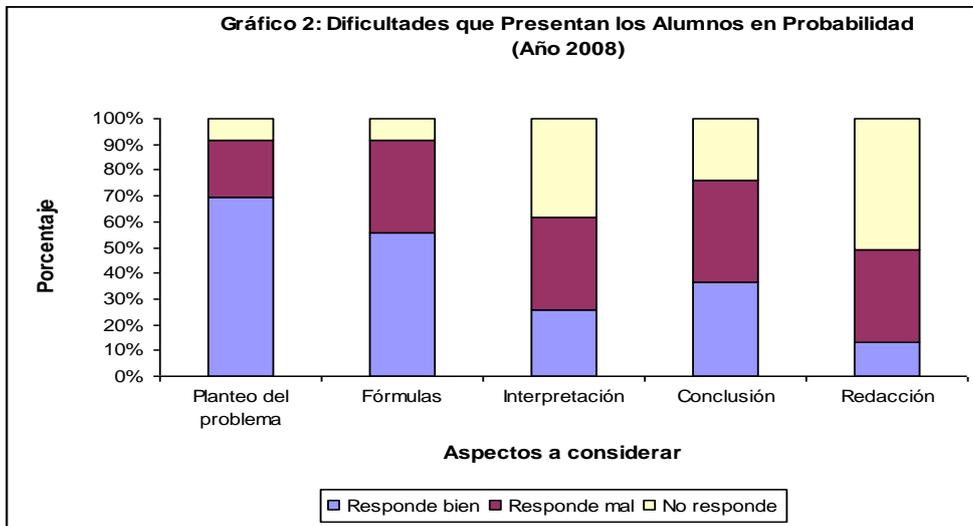


GRÁFICO 2: Dificultades que presentan los alumnos en Probabilidad. (Año 2008).

El 86% alumnos presentan problemas a la hora de redactar la interpretación de los resultados y las conclusiones obtenidas. No dejan de ser menos preocupantes los resultados obtenidos en “Interpretación” y “Conclusión” con un 74% y un 63% respectivamente, de alumnos que efectúan mal esa actividad o no la hacen.

Al realizar el cruce de la variable Planteo del Problema con los demás Ítems considerados se obtiene:

TABLA 4: Relación entre el Planteo del Problema y los otros Ítems ⁽¹⁾

Ítems	Respuestas de los Alumnos	Planteo del Problema	
		Bien	Mal
Fórmulas	Bien	79	4
	Mal	21	96
	No responde	0	0
Total		100	100
Interpretación	Bien	38	0
	Mal	29	70
	No responde	33	30
Total		100	100
Conclusión	Bien	50	9
	Mal	35	70
	No responde	15	21
Total		100	100
Redacción	Bien	19	0
	Mal	34	57
	No responde	47	43

Total		100	100
--------------	--	------------	------------

(1) La categoría “no responde “ del Ítem Planteamiento del problema no se tiene en cuenta dado que se corresponde con “ejercicio sin resolver”

De los alumnos que plantean mal el problema el 96% escribe o aplica mal la fórmula, el 30% no interpreta los resultados y un 70% lo hace de manera equívoca. El 21% no saca conclusiones mientras que el 70% arriba a conclusiones erróneas, mientras que el 100% presentó problemas al momento de expresar la conclusión y/o escribir la interpretación de los resultados

De manera más específica, algunos alumnos:

- Obtienen valores de probabilidad fuera del intervalo [0,1] lo que muestra que poseen dificultad para entender el concepto de probabilidad y de su relación con el de proporción
- No logran diferenciar entre sucesos independientes y sucesos mutuamente excluyentes.
- Confunden la aplicación del Teorema de la Suma de Probabilidades con el de la Multiplicación de Probabilidades.
- No advierten cuando un suceso está condicionado por otro.
- Les cuesta enunciar con palabras lo que está expresado en forma simbólica y viceversa.

3. VARIABLES ALEATORIAS Y DISTRIBUCIONES PROBABILÍSTICAS

TABLA 5: Dificultades que presentan los alumnos en Variables Aleatorias y Distribuciones Probabilísticas

Ítems	Acción del Alumno			Total
	Responde		No responde	
	Bien	Mal		
Planteo del problema	77 (74%)	5 (5%)	22 (21%)	104 (100%)
Fórmulas	68 (65%)	13 (13%)	23 (22%)	104 (100%)
Cálculos	56 (54%)	25 (24%)	23 (22%)	104 (100%)
Uso de tablas	46 (44%)	33 (32%)	25 (24%)	104 (100%)

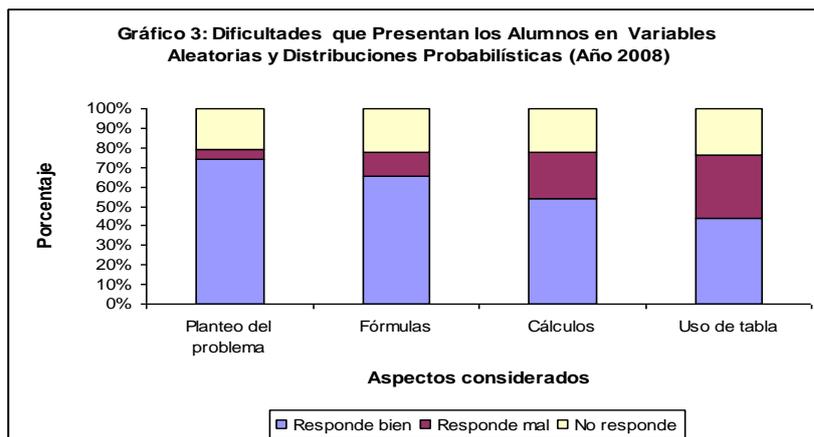


GRÁFICO 3: Dificultades que presentan los alumnos en Variables Aleatorias Distribuciones Probabilísticas (Año 2008).

Sólo el 27% de los alumnos muestra dificultades en el planteamiento del problema. Para el 56%, el manejo de las tablas resulta conflictivo o problemático.

Efectuando el cruzamiento entre Planteo del Problema con los demás ítems considerados se obtiene:

TABLA 6: Relación entre el Planteo del Problema y los otros Ítems ⁽¹⁾

Ítems	Respuestas de los Alumnos	Planteo del problema	
		Bien	Mal
Fórmulas	Bien	87	20
	Mal	13	60
	No responde	0	20
Total		100	100
Cálculo	Bien	71	20
	Mal	29	60
	No responde	0	20
Total		100	100
Uso de tabla	Bien	59	20
	Mal	40	40
	No responde	1	40
Total		100	100

(1) La categoría “no responde” del Ítem Planteamiento del problema no se tiene en cuenta dado que se corresponde con “ejercicio sin resolver”

Del total de alumnos que desarrollan un planteo erróneo del problema, el 80% exhibe dificultades en la utilización de fórmulas. Igual porcentaje se observa en lo referido a complicaciones en el uso de las tablas y a inconvenientes al efectuar los cálculos

Más específicamente

- No saben expresar simbólicamente ni de manera lógica la secuencia a seguir cuando calculan probabilidades que no aparecen de manera directa en las tablas.

- Presentan grandes dificultades en identificar los valores que asume la variable cuando el evento cuya probabilidad se desea calcular está expresado en términos de "...al menos..", "...a lo sumo..."
- Al calcular probabilidades en la Distribución Normal los alumnos tienden a confundir el valor de la variable con el que asume la función de distribución en ese punto.
- No detectan fácilmente los casos en los que puede aplicarse la aproximación de la Distribución Binomial a la distribución normal y si lo hacen no aplican correctamente la corrección correspondiente.
- Obtienen probabilidades mayores que 1
- Resulta un verdadero dilema para algunos alumnos el dar respuesta al interrogante: Cuál es la variable en estudio? en los ejercicios propuestos, particularmente los de Distribución Binomial y Distribución Poisson .

DISCUSIÓN

Como se desprende del análisis de los de los exámenes del Primer Parcial, los alumnos de Estadística y Biometría año 2008 cometen errores muy semejantes a los encontrados por los autores citados en la introducción de este trabajo, presentando también similitudes en las dificultades.

Según Batanero (2000) una de las grandes dificultades con la que se enfrentan los alumnos es la originada por el hecho de que los términos matemáticos con que se designan los conceptos estadísticos tienen un significado preciso, pero éste no siempre coincide con el asignado en el lenguaje coloquial. Para el caso particular de los alumnos cuyos exámenes son objeto de este estudio hay que añadir a eso la pobreza que evidencia su léxico. Todo ello se traduce en un gran obstáculo al momento de adquirir los conocimientos estadísticos

Una persona culta debiera poder leer críticamente los gráficos estadísticos que encuentra en la prensa, Internet, medios de comunicación, y trabajo profesional. (Gal, 2002 citado por Arteaga, Batanero, Diaz et al, 2009). Los alumnos considerados no poseen la competencia de reconocer informaciones cuantificables que han sido expresadas gráficamente a pesar de tener, teóricamente, *los* conocimientos matemáticos suficientes para ello. Una causa podría ser que durante su educación obligatoria no se incentivaron actividades que les permitieran a) interpretar datos presentados en gráficas, b) resolver problemas en los que interviniera la lectura de los mismos. Además el hecho de que en un mismo tipo de gráfico (por ejemplo, un gráfico simple de barras) puede ser usado para representar diferentes objetos

matemáticos, tales como frecuencias absolutas, relativas, porcentajes y frecuencias acumuladas, medias u otros resúmenes estadísticos es toda una cuestión.

En lo referido a las “medidas de resumen”, se observa en los estudiantes dificultades para comprender las ideas de resumen de un conjunto de datos o representante de los datos, y como advierte Mokros y Russell (1995) citado por Batanero (2000) hasta que no conciben el conjunto de datos como un todo no podrán alcanzar tal comprensión. Se detecta que los alumnos pueden: a) seleccionar correctamente la medida de resumen a utilizar y errar al aplicar la fórmula correspondiente b) no cometer el error anterior pero obtener valores no consistentes con las observaciones c) efectuar correctamente la selección de la medida y la aplicación de la fórmula, y alcanzar un valor razonable sin embargo no saber interpretarlo en términos del problema planteado. Todo esto indicaría que no hay una real comprensión de los conceptos subyacentes en la definición de la medida

Los problemas que se evidencian en el tema probabilidad tendrían como causa el hecho de que los alumnos necesitan del razonamiento proporcional el cual, según se indicó en la introducción resulta un “tópico difícil”. La “falsas intuiciones” que los alumnos traen consigo al empezar la enseñanza sería otra causa. Por último, el poco tiempo disponible y la escasez de conocimientos previos de los estudiantes hacen bastante complejo el llevar a cabo un estudio significativo de la probabilidad.

Se sabe que el concepto de variable aleatoria no es ni fácil de aprender ni de enseñar. Se necesita de la comprensión de fenómenos aleatorios, sucesos aleatorios, operaciones de eventos y Probabilidad y si no se han afianzado estos saberes el concepto de variable aleatoria no podrá ser asimilado (comprendido). Todas estas cuestiones se convierten en obstáculos a la hora de abordar la temática de las distribuciones probabilísticas de variables aleatorias discretas (Binomial y Poisson) y de las variables aleatorias continuas (Normal). Un aspecto interesante, es el hecho de que los alumnos que se apoyan en las representaciones gráficas de la distribución Normal en la resolución de ejercicios de cálculo de Probabilidades obtienen mejores resultados que los que no lo hacen. Esto hace especular (Ramos y Espinel, 2003) citado por Espinel y Ramos (2007) que la visualización gráfica proporciona al estudiante una mayor comprensión de los conceptos y del desarrollo metodológico. Al abordar la aproximación de una distribución Binomial mediante la distribución normal los alumnos no llegan a distinguir lo discreto y de lo continuo provocando esto la aplicación mecánica de la corrección por continuidad. Los inconvenientes observados en los estudiantes en cuanto a la aproximación de una distribución Binomial mediante la distribución Poisson están referidos principalmente a no reconocer que la variable toma un número finito de valores en el caso Binomial, mientras en el otro puede asumir infinitos (numerables).

CONCLUSION

Como soluciones al problema, desde un punto de vista general, y tal como lo expresa Batanero en “Razonamiento probabilístico en la vida cotidiana: un desafío educativo” (2006) es necesario reforzar la formación del razonamiento probabilístico en la educación primaria y secundaria ya que como afirma Fischbein (1975), citado por Batanero, Tauber y Sánchez, (2001) la intuición probabilística, y en consecuencia la estadística basada en aquella, no se adquieren espontáneamente.

Resulta imprescindible conocer el tipo de error que comenten los alumnos en cada contenido desarrollado, para así buscar la estrategia adecuada. A tal fin se considera que los contenidos y actividades de cada unidad didáctica se deben adaptar, ampliar o variar a fin de contemplar la diversidad de errores y dificultades que pueden presentar los alumnos. Sin embargo es vital que el estudiante participe activamente en el proceso de superación de sus propios errores; ya no es suficiente decirle cuál es el camino correcto o cuál es la solución debe ser el alumno quien advierta que su saber es insuficiente o no es el adecuado, pues de lo contrario continuará recurriendo a él

Está considerado altamente beneficioso para despertar el interés del alumno por la temática abordada y para que no sólo adquiriera una serie de conocimientos, sino también que desarrolle procedimientos autónomos de pensamiento al llevar a cabo actividades donde éste pase por las distintas fases: plantear el problema, recoger, organizar y analizar los datos, y extraer conclusiones en relación al problema planteado. Cabe destacar que el año 2008 fue el segundo año en que se puso en práctica el Trabajo Práctico Especial “Análisis de Datos Utilizando InfoStat” con datos originados y recogidos por los alumnos (años anteriores la cátedra proveía los datos). En la primera clase práctica en aula los estudiantes son medidos, pesados y responden una pequeña encuesta. Con los datos generados deben efectuar un estudio comparativo entre los niveles de alguna de las variables categóricas de la encuesta para una variable cuantitativa de las relevadas aplicado las distintas técnicas estudiadas. La presentación del trabajo para su evaluación es al final del dictado de la signatura.

Estadística descriptiva, análisis exploratorio son temas que tienen aplicación directa en este Trabajo Práctico Especial.

Para temas que involucran probabilidades se propone llevar a cabo actividades en las que el alumno pueda resolver situaciones problemáticas basadas en contextos de la vida real ya que la comprensión de estadísticas, datos, azar y probabilidad se deriva de aplicaciones del mundo real. Probabilidades es un tema de aplicación parcial en el Trabajo Práctico Especial

En cuanto a Variables aleatorias y Distribución de Probabilidades además de diseñar tareas de aplicación de esta temática en entornos de interés para el estudiante, se sugiere integrar las nuevas tecnologías (como en el caso de Estadística Descriptiva) utilizando el software estadístico InfoStat; al mismo tiempo se propone el uso de los applets, pequeños programas interactivos escritos en Java, que permiten resolver problemas mediante la aplicación de una técnica Estadística o visualizar un concepto determinado. Como expresa Espinel y Ramos (2007, pág 8) "... su uso puede resultar beneficioso ya que mediante éstos se pueden corregir algunos de los errores más comunes que cometen los alumnos. Así por ejemplo: Los applets dedicados al cálculo de probabilidades como representan la distribución de probabilidad estudiada, permiten identificar las áreas con las probabilidades, y los puntos críticos con los valores en la recta, lo que conlleva un manejo adecuado de las tablas de distribución".

Variables Aleatorias y Distribuciones probabilísticas tienen una aplicación natural en "Análisis de Datos Utilizando InfoStat".

Se supone que si la presentación del Trabajo Práctico Especial se hiciera de manera alternada en el sentido de que a medida que se van desarrollando las clases se fueran aplicando las técnicas y exponiendo y explicando los resultados, la efectividad de esta actividad sería mayor.

Se cree conveniente insistir, en los casos que sea factible, en visualizaciones gráficas ya que éstas brindan al alumno una mayor comprensión de los conceptos y del desarrollo metodológico orientándolos en la interpretación correcta del problema y en la extracción de conclusiones apropiadas

La base debe ser la participación activa del alumno. Se aprende con lo que se vive, lo que se recrea y lo que se reinventa y no simplemente con lo que se lee y se escucha.

BIBLIOGRAFIA

- ARTEAGA, P., BATANERO, C., DÍAZ, C. Y CONTRERAS, J.M. (2009). El lenguaje de los gráficos estadísticos. *Revista Iberoamericana de Educación Matemática*, Junio de 2009, Número 18, páginas 93 – 104. ISSN: 1815-0640. http://www.fisem.org/descargas/18/Union_018_012.pdf

- AUSUBEL, D. P.; NOVAK, J. D. Y HANESIAN, H. (1983). Psicología educativa: un punto de vista cognitivo. México, Editorial Trillas.
- BATANERO, C. (2000). Significado y comprensión de las medidas de posición central. *UNO*, 2000, 25, 41-58. <http://www.ugr.es/~batanero/ARTICULOS/isboa.pdf>
- BATANERO, C. (2006). Razonamiento probabilístico en la vida cotidiana: un desafío Educativo. En P. Flores y J. Lupiáñez (Eds.), *Investigación en el aula de Matemáticas. Estadística y Azar*. Granada: Sociedad de Educación Matemática Thales. ISBN: 84-688-0573-4. CD ROM. <http://www.ugr.es/~batanero/publicaciones%20index.htm>
- BATANERO, C., GODINO, J.D., GREEN, D.R., HOLMES, P. Y VALLECILLOS, A. (1994). Errores y dificultades en la comprensión de los conceptos estadísticos elementales. [Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics Education in Science and Technology*, 25(4), 527-547] <http://www.ugr.es/~batanero/didactica%20de%20la%20estadistica.htm>
- BATANERO, C., TAUBER, L. Y MEYER, R., (1999). From Data Analysis to Inference: A Research Project on Students' Understanding of The Normal Distribution. International Statistical Institute, 52nd Session 1999. <http://www.stat.auckland.ac.nz/~iase/publications/5/bata0054.pdf>
- BATANERO, C., TAUBER, L. Y SÁNCHEZ, M.V. (2001). Significado y comprensión de la distribución normal en un curso introductorio de análisis de datos . *Cuadrante*, 10 (1), 59-92. <http://www.ugr.es/~batanero/ARTICULOS/Cuadrante.pdf>
- CAÑIZARES, M.J. Y BATANERO, C. (1998). Influencia del razonamiento proporcional y de las creencias subjetivas en la comparación de probabilidades. *UNO*, 14, 99-114. <http://www.ugr.es/~batanero/publicaciones%20index.htm>
- CARRIÓN MIRANDA, V. (2007) Análisis de errores de estudiantes y profesores en expresiones combinadas con números naturales. *Unión Revista Iberoamericana de educación matemática*. Septiembre de 2007, Número 11, páginas 19-57. ISSN: 1815-0640. http://cimm.ucr.ac.cr/ciaem/articulos/educacion/materiales/Analisis%20de%20errores%20de%20estudiantes%20y%20profesores%20en%20expresiones%20con%20n%C3%BAsmeros%20naturales.*Miranda%20Carri%C3%B3n,%20Vicente.*Union_011_007.pdf
- COBO, B. Y BATANERO, C. (2000). La mediana ¿Un concepto sencillo en la enseñanza secundaria? . *UNO*, 23, 85-96. <http://www.ugr.es/~batanero/ARTICULOS/MEDIANA.pdf>
- DEL PUERTO, S., MINNAARD, C., Y SEMINARA, S. (2004). Errores en el aprendizaje de las Matemáticas. *Elementos de Matemática*. Publicación Didáctico Científica de la Universidad CAECE. 1.ª parte, 19 (74), pp. 5-18. 2.ª parte, 19 (75), pp.17-32. <http://www.rieoei.org/deloslectores/1285Puerto.pdf>

- DEL PUERTO, S., SEMINARA, S. y MINNAARD, C. (2007). Identificación y análisis de los errores cometidos por los alumnos en Estadística Descriptiva. *Revista Iberoamericana de Educación* N° 43/3 (ISSN: 1681-5653). <http://www.rieoei.org/expe/1729Puerto.pdf>
- DÍAZ, C. (2005). Evaluación de la falacia de la conjunción en alumnos universitarios. *Suma* 2005, 48, 45-50. <http://www.ugr.es/~batanero/publicaciones%20index.htm>
- DI RIENZO, J.A., CASANOVES, F., BALZARINI, M.G., GONZALEZ, L., TABLADA, M., ROBLEDO, C.W. InfoStat versión 2009. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina.
- ESPINEL, M.C., RAMOS, R.M. (2007). Algunas alternativas para la mejora de la enseñanza de la inferencia estadística en secundaria. *Números*, Revista didáctica de las matemáticas N° 67 http://www.sinewton.org/numeros/numeros/67/ideas_03.php
- MOREIRA, M. A. (2000). Aprendizaje Significativo: teoría y práctica. Ed. Visor. Madrid.
- ORTIZ, J.J., BATANERO, C. Y SERRANO, L. (2001). El lenguaje probabilístico en los libros de texto. *Suma*, 38, 5-14. <http://www.ugr.es/~batanero/publicaciones%20index.htm>.
- RICO, L. (1995). "Errores y dificultades en el aprendizaje de las Matemáticas", cap. 3. : 69-108, en KILPATRIK, J.; GÓMEZ,P., y RICO, L.: Educación Matemática. Grupo Editorial Iberoamérica, Méjico. <http://cumbia.ath.cx:591/pna/Archivos/RicoL94-148.PDF>.
- RUIZ HERNÁNDEZ, B., ALBERT HUERTA, J. AND BATANERO, C. (2006). An exploratory study of students' difficulties with random variables. http://www.ime.usp.br/~abe/ICOTS7/Proceedings/PDFs/InvitedPapers/6B2_RUIZ.pdf
- Tipos de Obstáculos. Matemática Par@ educ.ar. Aportes para la enseñanza en el Nivel Medio. http://aportes.educ.ar/matematica/popup/tipos_de_obstaculos.php