

## LA CONSTRUCCIÓN DEL SIGNIFICADO DE LA ASOCIACIÓN MEDIANTE ACTIVIDADES DE ANÁLISIS DE DATOS: REFLEXIONES SOBRE EL PAPEL DEL ORDENADOR EN LA ENSEÑANZA DE LA ESTADÍSTICA

Carmen Batanero, Juan D. Godino, Universidad de Granada

Antonio Estepa, Universidad de Jaén

Durante los cursos 1992 a 1998 hemos trabajado en un proyecto de investigación dirigido al estudio de las concepciones iniciales que tienen los alumnos sobre la asociación estadística y su evolución después de diversos experimentos de enseñanza usando ordenadores. En este trabajo, describimos brevemente los resultados de este proyecto, y los utilizamos como base para la reflexión sobre el papel del ordenador como recurso didáctico y como instrumento en la resolución de problemas, extendiendo las conclusiones presentadas en Batanero y cols. (1998)

### MARCO TEÓRICO

Nuestras investigaciones se han basado en el marco teórico sobre el significado y comprensión de los objetos matemáticos, en sus dimensiones personal e institucional, que se describe en Godino (1996) y Godino y Batanero (1994, 1998). El principal supuesto epistemológico subyacente es que los objetos matemáticos emergen de la actividad del sujeto en la resolución de problemas, mediatizados por los instrumentos semióticos disponibles que dependen de los institucionales en que tiene lugar dicha actividad. El significado de los objetos matemáticos se concibe como el sistema de prácticas vinculado a campos específicos de problemas, considerándose en este sistema tres tipos de elementos diferentes:

(1) Elementos *extensionales* del significado: Los diferentes problemas y situaciones prototípicas donde se usa el objeto, es decir, el campo de problema del que el objeto matemático emerge.

(2) Elementos *instrumentales/relacionales* del significado: Las diferentes herramientas semióticas disponibles para estudiar, resolver y/o representar los problemas y objetos matemáticos involucrados.

(3) Elementos *intensionales* del significado: Las diferentes propiedades características y relaciones de los objetos matemáticos con otras entidades: las definiciones, proposiciones, descripciones procedimentales, etc.

Según este modelo, la comprensión de un concepto matemático implicará la apropiación de los diferentes elementos que componen el significado institucional del concepto y, en consecuencia, tiene una naturaleza sistémica.

Al iniciar la búsqueda de bibliografía previa no encontramos trabajos dentro del ámbito de la Educación Matemática, por lo que nos basamos principalmente en las investigaciones sobre asociación en psicología, que comienzan con el trabajo pionero de Inhelder & Piaget (1955).

También se han tenido en cuenta las investigaciones de Crocker, 1981; Beyth - Maron, 1982, Chapman y Chapman (1969) y Jennigs, Amabile y Ross (1982).

## UN ESTUDIO DE CONCEPCIONES INICIALES SOBRE LA ASOCIACIÓN

### *SIGNIFICADO EXTENSIONAL DE LA ASOCIACIÓN*

Un caso particular del campo de problemas del que emerge la asociación estadística es realizar un juicio de asociación en una tabla de contingencia 2x2, como la que se muestra en el ítem 1. Para resolver este problema es preciso realizar operaciones y establecer relaciones con las diferentes frecuencias que aparecen o pueden calcularse en la tabla.

**Ítem 1:** En un centro médico se han observado a 250 personas para observar si el hábito de fumar tiene alguna relación con los trastornos bronquiales, obteniendo los siguientes resultados:

	Fuma	No fuma	Total
Padece trastornos bronquiales	90	60	150
No padece trastornos bronquiales	60	40	100
Total	150	100	250

Usando la información contenida en la tabla se podría pensar que, para esta muestra, ¿los trastornos bronquiales dependen de fumar? Razone la respuesta

Otro problema diferente es valorar la correlación existente entre dos variables cuantitativas (ítem 3), para lo cual podemos calcular la covarianza, el coeficiente de correlación o analizar la bondad del ajuste de una recta de regresión al diagrama de dispersión. Un tercer tipo de problema consiste en tratar de averiguar si una variable numérica tiene la misma distribución en dos muestras diferentes (ítem 2), que se puede resolver comparando la diferencia entre las medias o medianas, o bien, la representación gráfica o tabular de ambas distribuciones. Estos problemas y actividades son imprescindibles para construir progresivamente el concepto de asociación estadística y forman parte del significado matemático institucional del concepto dentro de un curso universitario introductorio al análisis de datos. Con más precisión, los tres tipos de problemas describen los elementos prototípicos extensionales del significado de la asociación en dicha institución.

**Ítem 2:** Al medir la presión sanguínea antes y después de haber efectuado un tratamiento médico a un grupo de 10 personas se obtuvieron los siguientes valores:

	presión sanguínea en cada mujer									
Mujer	Sra.A	Sra.B	Sra.C	Sra.D	Sra.E	Sra.F	Sra.G	Sra.H	Sra.I	Sra.J
Antes del tratamiento	115	112	107	119	115	138	126	105	104	115
Después del tratamiento	128	115	106	128	122	145	132	109	102	117

Usando la información contenida en esta tabla, ¿Piensa que la presión sanguínea en esta muestra depende del momento de que se tome antes o después del tratamiento? Razone la respuesta.

**Ítem 3:** En un estudio sociológico, se han recogido datos relativos a la tasa de natalidad y el consumo diario de proteínas animales en diferentes países representándolos en el gráfico de la figura 1 ¿Cree que la relación entre el consumo diario de proteínas animales y la tasa de natalidad en los diferentes países es directa, inversa o no existe relación entre estas variables? Razone su respuesta.

### *SIGNIFICADOS PERSONALES DE LA ASOCIACIÓN*

Nuestro proyecto de investigación comenzó con el estudio del significado personal que los estudiantes daban al concepto de asociación antes de haber estudiado el tema. Después de algunas revisiones con muestras piloto, pusimos a punto un cuestionario con ítems semejantes

a los presentados anteriormente, que propusimos a 213 alumnos de COU, antes de recibir enseñanza sobre la asociación. Se controlaron las siguientes variables de la tarea: Signo e intensidad de la asociación, relación entre las creencias previas de los alumnos sobre el contexto del problema y el tipo de asociación presentada.

En cada ítem analizamos el tipo de asociación percibido por los estudiantes (asociación directa, inversa o independencia). Además clasificamos las estrategias de resolución estudiada por los alumnos desde un punto de vista matemático. Esto nos permitió identificar estrategias intuitivas correctas que sugieren concepciones correctas o parcialmente correctas sobre la asociación estadística (Estepa et al., 1994; Batanero et al., 1996; Estepa, & Batanero, 1996; Estepa, & Sánchez Cobo, 1996). Algunos ejemplos se exponen a continuación:

(1) Utilizar la tendencia constante, creciente o decreciente de los puntos en los diagramas de dispersión para justificar el tipo de asociación (nula, positiva o negativa): *"Porque cuando el consumo diario de proteínas aumenta la tasa de natalidad disminuye"* (Ítem 3) Como en caso de independencia, no habrá variación conjunta, estos estudiantes muestran una concepción correcta de asociación.

(2) Utilizar las medias o los totales para comparar la distribución de una variable en dos muestras diferentes *"Porque la suma de todos los valores de la presión de la sangre antes del tratamiento es menor que la suma de valores de la presión de la sangre después del tratamiento"* (ítem 2). Aquí los estudiantes usan implícitamente la idea correcta de que una diferencia en los totales implica asociación entre las variables.

(3) Comparar la frecuencia de casos a favor y en contra de la asociación en cada valor de la variable independiente o la razón de estas frecuencias en tablas de contingencia 2xr: *"No depende, porque 3/2 de los fumadores tienen trastornos bronquiales, y hay la misma proporción en los no fumadores"* (ítem 1). Esto indica una concepción correcta, ya que la razón de posibilidades se puede utilizar para evaluar la asociación en una tabla de contingencia 2xr.

Otras estrategias de los estudiantes para resolver estos problemas eran inadecuadas, y les proporcionaban juicios incorrectos de asociación. A partir de ellas hemos descrito las siguientes concepciones erróneas sobre la asociación estadística:

(1) Concepción determinista de la asociación: Algunos estudiantes no admiten más de un valor de la variable independiente para cada valor de la variable dependiente. Cuando esto no ocurre, consideran que no hay dependencia entre las variables. En otras palabras, la relación entre las variables debe ser una función desde el punto de vista matemático. Por ejemplo: *"El tratamiento no tiene mucha influencia, ya que a algunas mujeres les aumenta la presión sanguínea, mientras que a otras les disminuye"* (Ítem 3).

(2) Concepción unidireccional de la asociación: Si se percibe la independencia solamente cuando es positiva (asociación directa), considerando la asociación inversa como independencia. El siguiente ejemplo ilustra un caso de asociación inversa interpretada como independencia y dificultades en el razonamiento proporcional. *"Personalmente creo que no hay dependencia porque si tú miras la tabla hay mayor proporción de personas con trastornos bronquiales entre los fumadores"* (ítem 1). Este tipo de concepción fue también encontrado posteriormente por Morris (1997).

(3) Concepción local de la asociación: Utilizar solamente parte de los datos proporcionados por el problema para emitir el juicio de asociación. Si la parte de datos utilizados confirma un tipo de asociación, adoptan este tipo de asociación en sus respuestas. *"Existe dependencia entre fumar y padecer trastornos bronquiales porque si observamos la tabla hay más fumadores con trastornos bronquiales que no fumadores 90>60"* (ítem 1).

(4) Concepción causal de la asociación: Algunos estudiantes solamente consideran la existencia de asociación entre variables si se puede atribuir una relación causal entre ellas. Este tipo de concepción particularmente se encontró en un problema en el que se pedía que dos jueces puntuaran a un conjunto de individuos: *"Porque un juez no puede influir en el otro. Cada uno tiene sus preferencias no puede haber mucha relación entre las puntuaciones otorgadas por cada uno"*.

### **INFLUENCIA DE LOS ENTORNOS INFORMÁTICOS EN LA ENSEÑANZA APRENDIZAJE DE LA ASOCIACIÓN**

Después de llevar a cabo el estudio sobre concepciones iniciales de los estudiantes sobre la asociación, nuestra investigación se orientó a valorar el impacto que una experiencia de aprendizaje usando ordenadores tiene en dichas concepciones iniciales. El uso de ordenadores en la enseñanza de la Estadística está recibiendo atención creciente tanto por parte de los profesores como de los investigadores como se puede constatar en Shaughnessy et al., (1996) y en la IASE Round Table Conference sobre el impacto de las nuevas tecnologías en la enseñanza-aprendizaje de la Estadística.

Es un hecho evidente la influencia que los ordenadores han tenido en el desarrollo de la Estadística como disciplina y en facilitar el acceso a la estadística a un número y diversidad cada vez mayor de usuarios, aumentando las demandas de formación básica en estadística. Los ordenadores han aumentado, por un lado, el número de contenidos estadísticos a enseñar, incluyendo el uso adecuado del software, sin el cual es hoy día impensable la realización del análisis de datos en cualquier campo de aplicación. También ha habido un cambio en los contenidos, prestándose mayor importancia a los aspectos interpretativos y conceptuales y menor a los procedimentales y algoritmos de cálculo.

La tecnología ofrece, además poderosos recursos didácticos tales como la simulación, y las representaciones gráficas, que unidos a la velocidad de cálculo y posibilidades de manipulación y exploración por parte de los estudiantes les puede ayudar a ampliar el significado de los conceptos estadísticos. El objetivo de nuestra investigación era valorar este impacto en el caso específico de la asociación estadística.

### **PLANIFICACIÓN DE LA ENSEÑANZA**

Biehler clasifica el software estadístico según sus funciones educativas. Los *micromundos* posibilitan la realización de experimentos interactivos, mediante simulaciones y visualizaciones exploratorias, que ayudan a los estudiantes a conceptualizar la estadística. Las *herramientas* permiten a los estudiantes practicar la estadística del mismo modo que lo hacen los estadísticos profesionales. En el caso particular del análisis exploratorio de datos, estas herramientas deberían facultar a los estudiantes para hacer un trabajo interactivo, exploratorio y abierto, utilizando software flexible, fácil de usar y aprender.

En nuestros dos experimentos hemos considerado *entornos informáticos de aprendizaje*, es decir, entornos instruccionales integrados que permiten al profesor y a los estudiantes trabajar con herramientas, micromundos y conjuntos de datos relativos a ciertos problemas de aplicación, así como con una selección de conceptos y procedimientos estadísticos.

El contenido de los dos cursos incluyó los conceptos básicos sobre poblaciones y muestras, organización de datos, tipos de variables estadísticas y distribuciones de frecuencias, gráficas, parámetros y estadísticos, posición central, dispersión, estadísticos de orden, asimetría, curtosis, variables estadísticas bidimensionales: tablas de contingencia, covarianza, correlación y regresión lineal. En el segundo experimento se introdujeron también los conceptos de muestreo, distribución en el muestreo, intervalos de confianza, test de hipótesis de las medias para una y dos muestras y test Chi-cuadrado.

La planificación de la enseñanza incluyó la delimitación de los objetivos de aprendizaje, organización de una secuencia instruccional de contenidos, y la selección de conjuntos de datos apropiados para contextualizar el conocimiento estadístico. Adoptamos una "perspectiva multivariante", aunque únicamente se enseñaran, a nivel formal, técnicas univariadas o bivariadas. Por lo tanto, los estudiantes exploraron ficheros de datos utilizando un paquete de software interactivo de ordenador. Se llevaron a cabo 21 sesiones en el primer experimento y 40 en el segundo. En siete de las sesiones del primer experimento y en 20 del segundo los estudiantes trabajaron en el laboratorio de estadística, resolviendo problemas, cuyas soluciones requerían el análisis de diferentes conjuntos de datos proporcionados por el profesor o recogidos por los propios alumnos. En la confección de los problemas de las sesiones prácticas se tuvieron en cuenta las variables de tarea relativas a la resolución de problemas estadísticos que se exponen en Godino y col. (1991). En el resto de las sesiones, se hacía una introducción de los conceptos estadísticos y se resolvían problemas relacionados con ellos.

### ***PRIMER EXPERIMENTO: IDENTIFICACIÓN DE CONCEPCIONES RESISTENTES Y ANÁLISIS DEL PROCESO DE APRENDIZAJE.***

La muestra de este experimento estaba integrada por diecinueve estudiantes de 20 años matriculados en un primer curso universitario de análisis exploratorio de datos y estadística descriptiva. Los estudiantes trabajaban con el paquete estadístico PRODEST, que había sido desarrollado por el equipo de investigación algunos años antes. Aunque este software es limitado, si se compara con paquetes estadísticos más modernos, incluye todas las herramientas necesarias para desarrollar un curso de estadística descriptiva a nivel universitario.

Para evaluar los cambios en las concepciones de los estudiantes, éstos cumplimentaron dos versiones equivalentes de un cuestionario como pre-test y post-test. Encontramos una mejora general en las estrategias que utilizaban, así como la persistencia después del proceso de enseñanza, en algunos estudiantes, de las concepciones erróneas unidireccional y causal sobre la asociación estadística. Todos estos resultados se han descrito con detalle en Batanero et al., (1997).

Adicionalmente se observó, durante las sesiones de laboratorio, a una pareja de estudiantes para poder describir y estudiar su proceso de aprendizaje. Un miembro del equipo de investigación observaba a la pareja de estudiantes mientras trabajaban en la resolución de los problemas, recogiendo las respuestas escritas a los distintos problemas y registrando los debates de los alumnos en cintas de cassette, así como las intervenciones del profesor en el trabajo de esta pareja de alumnos. También se grabó en disco las interacciones de los alumnos con el ordenador y la pareja de estudiantes fue entrevistada al comienzo y al final del experimento y su interacción con el ordenador quedó también registrada en un fichero que fue posteriormente impreso para disponer de una traza de sus estrategias en la resolución de los problemas.

Cuando estudiamos con detalle todas las observaciones anteriores encontramos algunas dificultades, que se repetían con regularidad, relacionadas con la asociación. Algunas de ellas se superaban, bien por los propios alumnos, bien con la ayuda del profesor, al final del proceso de enseñanza, si bien algunas volvían a aparecer de vez en cuando. Otras veces la dificultad no se superaba, a pesar del debate entre los alumnos o las intervenciones del profesor. Ocasionalmente, el profesor no apreciaba la confusión de los estudiantes.

### ***ELEMENTOS INTENSIONALES DEL SIGNIFICADO DE LA ASOCIACIÓN***

A continuación describimos el proceso de aprendizaje de estos alumnos, destacando algunos elementos intensionales clave del significado matemático de la asociación (Godino & Batanero, 1998) y el proceso de comprensión por parte de los estudiantes de estos elementos de significado a lo largo del tiempo de aprendizaje.

*1. La comparación de dos o más muestras, con objeto de estudiar la posible relación entre dos variables, debe efectuarse en términos de frecuencias relativas.* En la primera sesión los alumnos comienzan comparando frecuencias absolutas de la distribución de una variable en dos muestras. Este error es advertido por el profesor al final de la clase, pero se presenta de nuevo en las sesiones 2, 3 y 5. A partir de ahí los estudiantes parecen haberlo superado.

*2. La posible existencia o no de diferencias en la distribución de una variable entre dos o más muestras se deduce a partir de la comparación de toda la distribución de la variable en cada una de las muestras y no de una parte de la misma.* Los estudiantes, sin embargo, comienzan con la comparación de valores aislados, al estudiar las dos muestras. Por ejemplo, en la primera sesión, los estudiantes solamente comparan los valores de máxima y mínima frecuencia en ambas muestras; aunque estas diferencias apuntan a la existencia de posible asociación, este modo de proceder es insuficiente para cuantificar la intensidad de la misma. Esta dificultad vuelve a aparecer en las sesiones 2, 3 y 5, desapareciendo en las sesiones posteriores.

*3. A partir de una misma frecuencia absoluta pueden deducirse dos frecuencias relativas condicionales diferentes, según la variable que se emplee como condición. El papel de condición y condicionado en la frecuencia relativa condicional no es intercambiable.* Numerosos autores como Falk (1986) señalan la dificultad de interpretación de una

probabilidad condicional porque los alumnos no diferencian a veces el papel jugado por la condición y el condicionado, con lo que pueden confundir  $P(A|B)$  con  $P(B|A)$  o no llegar a discriminarlas. Muchos alumnos, en este estudio, mostraron confusiones similares a esta en el estudio de las concepciones previas, que siguieron manifestándose durante el proceso de instrucción y que ha seguido manifestándose al finalizar la instrucción. Apareció en la sesión 5 y se superó con la ayuda del profesor. No apareció en el resto de las sesiones.

4. *Dos variables son independientes si la distribución condicional de una de ellas no cambia cuando se varían los valores de la otra variable.* Hasta llegar a la sesión 5, los estudiantes no descubren que una condición para la independencia es la invarianza de las distribuciones relativas condicionales, cuando varía el valor de la variable condicionante.

5. *En la determinación de la asociación entre dos variables, éstas juegan un papel simétrico. Por el contrario, en el estudio de la regresión, las variables desempeñan un papel asimétrico. Hay dos rectas de regresión diferentes, según cual de las dos variables actúe como variable independiente.* El hecho de que en la correlación no se distinga entre la variable explicativa y la variable explicada, mientras que en la regresión esta diferencia sea esencial (Moore, 1995) provocó gran confusión entre los estudiantes. Cuando necesitaron seleccionar la variable explicativa para calcular la línea de regresión en las sesiones 5, 6 y 7, no supieron qué variable elegir. Por ejemplo, para calcular la línea de regresión del peso sobre la altura, los estudiantes se desconcertaron por el hecho de que existía mutua dependencia entre las dos variables, debatieron largamente sin llegar a una solución aceptable para ellos. El profesor no se dio cuenta del problema y finalmente los estudiantes calcularon la línea de regresión eligiendo la variable explicativa al azar. Al final del período de enseñanza estos estudiantes aun no habían descubierto que se pueden calcular dos líneas de regresión diferentes.

6. *Una correlación positiva indica dependencia directa entre las variables.* Aunque en la sesión 6, los alumnos pudieron interpretar la magnitud del coeficiente de correlación, no discutieron el tipo de asociación (directa o inversa). Al final de la sesión aunque llegan a indicar que "al aumentar una variable la otra aumenta" no identifican este hecho con la idea de relación directa entre las variables. Nunca llegan a emplear la idea de "relación o dependencia directa".

7. *Una correlación negativa indica dependencia inversa entre las variables.* En la sesión 6, los alumnos se sorprenden al encontrar por primera vez un coeficiente de correlación negativo, hasta el punto de preguntar al profesor si ello es posible. Asimismo, aparece la duda en la comparación de dos coeficientes de correlación negativos, ya que, en este caso, un número menor corresponde a mayor intensidad en la asociación. Así, el conocimiento adquirido sobre el orden de los números negativos dificulta ahora la comprensión del signo negativo del coeficiente de correlación; se convierte en obstáculo epistemológico para dicha comprensión. Consideramos que ello se debe al fenómeno de inversión en la relación de orden (González y cols., 1990). En realidad, aunque ayudados a veces por el profesor han observado que el signo negativo del coeficiente de correlación se corresponde con una pendiente negativa en la recta de regresión y que al crecer los valores de  $x$  disminuyen los de  $y$ , en el resto de la sesión, no llegan a utilizar el término "dependencia inversa". No llegan a diferenciar los dos tipos de asociación al término del aprendizaje.

8. *El valor absoluto del coeficiente de correlación es indicativo de la intensidad de la asociación.* Aunque en las primeras actividades los alumnos asocian un alto valor del coeficiente con una dependencia fuerte, hasta la sesión 6 no identifican, en principio, la idea de "intensidad de asociación" con el coeficiente.

### **SEGUNDO EXPERIMENTO: AMPLIACIÓN DE LOS ELEMENTOS DE SIGNIFICADO INSTRUMENTALES/REPRESENTACIONALES**

Treinta y seis estudiantes participaron en el segundo experimento, trabajando con algunos procedimientos del paquete estadístico *Statgraphics*, que incluye nuevos procedimientos instrumentales/representacionales y un entorno más dinámico y versátil para estudiar la asociación. Podemos clasificar las herramientas disponibles por su nivel de reducción de datos, su carácter numérico o gráfico y su enfoque analítico (descriptivo o inferencial). Cada una de las herramientas proporcionan diferentes resúmenes de los datos y, en consecuencia, diferentes significados de la asociación.

1. *Resúmenes numéricos en el nivel descriptivo:* Tablas de contingencia (v.g. ítem 1) y sus diferentes tipos de frecuencias. Tablas de frecuencias unidimensionales de distribuciones condicionales y sus estadísticos, correlación y coeficiente de determinación y parámetros de la línea de regresión.

2. *Resúmenes numéricos en el nivel inferencial:* Intervalos de confianza para las medias o la diferencia de medias. Test de hipótesis para las medias o medianas. Test Chi-cuadrado de asociación entre dos variables.

3. *Representaciones gráficas de distribuciones condicionales unidimensionales:* Gráfico del tronco, gráfico de barras, gráfico de la caja, gráfico de sectores, histogramas, curva empírica de distribución, funciones de densidad.

4. *Gráficos bidimensionales:* Histogramas tridimensionales, gráfico de mosaicos y diagrama de dispersión tridimensional.

En *Statgraphics*, además de la posibilidad de seleccionar gráficos y variables, se pueden tener diversas representaciones gráficas y tabulares simultáneamente en 1 pantalla, y manipular ciertas características de los gráficos como anchura, formato, escala, etc. El fichero *Statfolio* nos permite elaborar informes de investigación con resultados parciales que se incluyen en el texto al mismo tiempo que el análisis se está llevando a cabo.

Por otro lado, el hecho mismo de realizar los cálculos con ayuda del ordenador, introduce cambios en los elementos intensionales de significado que el alumno debe usar en la resolución de problemas, como veremos en la resolución que hace M. Luisa, del siguiente problema, durante una de las prácticas.

**Problema:** En el fichero MFF20 (referido a datos sobre capacidad lectora de niños de 6 años), ¿Cuál es el número máximo de errores en comprensión para el 90 por ciento de alumnos?.

Respuesta escrita de M. Luisa: "He usado la opción *STATS* del menú principal del programa y luego he elegido *DESCRIPTIVE METHODS* y *PERCENTILES*. Puse el nombre de la variable en la ventana de entrada de datos, escribiendo: MFF20. errorpalab. Después



puse 90 en el cuadro de porcentaje y obtuve 5. El número máximo de errores para el 90 por ciento de los niños es 5".

Tras esta concisa respuesta, podemos seguir el procedimiento de la estudiante, que resolvió correctamente el problema, eligiendo el método estadístico adecuado (cálculo de percentiles). Vemos que la alumna no necesita conocer el algoritmo de cálculo de percentiles, en el cual debemos diferenciar entre datos agrupados y no agrupados y entre número par e impar de datos. El algoritmo se ha "encapsulado", transformándose en una herramienta disponible (una opción del programa), un operador que se aplica a un vector de datos (la variable errorpalab) y asigna a cada rango (en este caso 90) un valor numérico de la variable.

Se añaden los siguientes pasos que no existen en un algoritmo tradicional:

1. Escoger STATS en el menú general, lo que supone diferenciar entre procedimientos estadísticos y procedimientos gráficos o de gestión de ficheros.

2. Elegir DESCRIPTIVE METHODS, dentro del menú STATS, reflexionando que necesitamos un método descriptivo de análisis de datos.

3. Escoger PERCENTILES, dentro de ONE VARIABLE NUMERICA que supone otros dos niveles de menús de opciones y requiere la identificación del tipo de variable (numérica), número de variables (una) y el procedimiento a usar (percentiles) dentro de los métodos descriptivos.

4. Identificar la variable MFF20-errorpalab dentro del fichero y asignarla en la ventana de entrada de datos. Requiere un conocimiento de la forma de operar el programa y la interpretación de la estructura del fichero de datos, así como relacionar lo anterior con el enunciado del problema.

5. Identificar el rango (90), es decir el valor necesario de los parámetros para ejecutar correctamente el procedimiento.

6. Interpretar los resultados.

Con un algoritmo tradicional se requieren llevar a cabo una serie de operaciones. Suponiendo que operamos con los datos brutos, esto es, sin realizar agrupaciones previas de los mismos. Estos pasos serían los siguientes:

1. Calcular el número  $n$  de datos.

2. Calcular el 90 por ciento de  $N$ ,  $Z=90 \times N/100$ . El valor de  $Z$  indica el lugar de la observación que contiene el valor del percentil, en la serie ordenada de datos. Es necesario calcular una proporción y recordar el significado del rango del percentil, como lugar que indica un porcentaje del número de datos.

3. Ordenar los datos por valores crecientes de la variable. Esto implica recordar que los percentiles son estadísticos de orden que indica la posición relativa de los valores de los datos y no el orden en que los datos fueron recogidos.

4. Localizar el elemento que ocupa el valor  $Z$  y hallar el valor de la variable que corresponde al elemento  $Z$ . Supone recordar que el percentil es el valor de la variable del dato que ocupa una posición dada y no la posición en sí misma. Si  $Z$  no corresponde exactamente

a un elemento, por haber una indeterminación, encontrar los dos elementos anterior u posterior a Z y calcular su media aritmética. Implica comprender el caso de indeterminación y el significado de una desigualdad.

Todos estos pasos sirven para reforzar las propiedades intensionales del concepto de percentil, propiedades que son irrelevantes si se usa el algoritmo informático. Como vemos, el significado de los percentiles se ve notablemente afectado por los nuevos útiles disponibles.

### ***SIGNIFICADOS PERSONALES DE LA ASOCIACIÓN DESPUÉS DE LA ENSEÑANZA***

La diversidad de los conocimientos estadísticos previos de los estudiantes en el segundo experimento era amplia, porque se trataba de un curso de libre configuración y los estudiantes inscritos pertenecían a diferentes Facultades: Educación, Psicología, Económicas, Caminos ... Durante el curso realizamos algunas evaluaciones mediante tareas de papel y lápiz y proyectos individuales realizados por los alumnos. Al final del curso se llevó a cabo un examen utilizando un nuevo conjunto de datos para valorar el significado final que los estudiantes tenían de la asociación estadística. Cada estudiante trabajó en solitario en un ordenador y sus soluciones se registraron en un disco, utilizando los ficheros "*Statfolios*", que incluyen los cálculos y gráficos realizados, junto con los comentarios y soluciones. El examen consistió en el análisis de un nuevo conjunto de datos relativo a las puntuaciones de 48 alumnos de una asignatura de Educación Física sobre el que se les planteaba a los alumnos algunos problemas:

**Problema 1.** Razone si en este conjunto de datos practicar deporte depende del sexo.

**Problema 2.** ¿Existe relación entre practicar deporte y número de pulsaciones después de 30 flexiones?

**Problema 3.** El profesor quiere valorar la eficacia de sus clases de Educación Física. ¿Cree que ha habido mejora en el tiempo empleado en recorrer 30 metros en diciembre con relación a diciembre?

**Problema 4.** ¿Cree que el número de pulsaciones después de 30 flexiones depende del número de pulsaciones en reposo?

La diferencia principal entre los problemas es el tipo de variables implicadas: Dos variables cualitativas (problema 1), dos variables cuantitativas (problema 4) y una variable de cada tipo (problemas 2 y 3). Otra variable importante es la intensidad de la asociación: independencia (problema 4), dependencia débil (problema 2), asociación moderada (muy significativo el valor de la t en el problema 3), (significativo valor de la Chi-cuadrado en el problema 1).

Los estudiantes deben identificar estas diferencias (discriminar entre diferentes elementos extensionales de significado), seleccionando entre las diferentes herramientas disponibles (elementos instrumentales/representacionales del significado) las que son más adecuadas para resolver el problema: Tablas de contingencia, test Chi-cuadrado, gráfico de mosaicos o comparación de gráficos de barras en el problema 1, comparación de histogramas, curva empírica de distribución, funciones de densidad, gráfico de la caja o diferentes estadísticos, intervalos de confianza o contrastes de hipótesis en los problemas 2 y 3 y estudiar el coeficiente de correlación, la nube de puntos o los parámetros de la recta de regresión en el problema 4. Finalmente los elementos intensionales del significado se utilizarían para interpretar las "salidas" de los diferentes programas y para realizar un juicio exacto de asociación.

A partir de los ficheros *statfolio* categorizamos las soluciones y procedimientos de los estudiantes. La mayor parte de los estudiantes dan un juicio de asociación correcto, aunque con frecuencia obtienen soluciones correctas utilizando procedimientos inadecuados al tipo de problema, no relacionando correctamente los elementos de significado extensionales (problemas) e instrumentales (útiles para resolverlos).

Además, entre las herramientas adaptadas a los problemas, los estudiantes no siempre seleccionan las que un estadístico profesional elegiría para realizar el análisis pedido. En consecuencia, las soluciones de los estudiantes no siempre coinciden con las soluciones "tipo". Por ejemplo, la mejor solución para el problema 1 sería utilizar el test Chi-cuadrado para analizar la diferencia de proporciones de varones y hembras que practica deporte, que da un valor significativo. Siete estudiantes, por el contrario, calculan en su lugar el coeficiente de correlación, que proporciona un valor de 0.28, que es muy pequeño, y en consecuencia, interpretan que no hay relación entre las variables. Otro ejemplo es el problema 3, donde un estudiante intenta visualizar la relación utilizando un diagrama de dispersión que no proporciona una solución al problema.

Esta selección de un procedimiento correcto pero no óptimo indica la falta de flexibilidad para cambiar de una a otra representación de la asociación. Por ejemplo, los estudiantes S11, S23, S26, S35, S36 resuelven tres problemas utilizando el coeficiente de correlación; el estudiante S12 resuelve tres problemas comparando gráficos de barras; el estudiante S10 resuelve los cuatro problemas comparando representaciones gráficas de distribuciones marginales y el estudiante S27 resuelve todos los problemas comparando tablas de contingencia de frecuencias dobles. Otro ejemplo es no tomar como variable explicativa la que permite una más simple interpretación del análisis, ya que tres estudiantes llegan a una conclusión no adecuada cuando calculan frecuencias relativas condicionales de la variable *practicar deporte* con la variable *número de pulsaciones*, que proporciona una mala visualización de la relación.

Otros estudiantes interpretan incorrectamente los resultados de un procedimiento correctamente seleccionado y su análisis muestran una falta de comprensión de elementos intensionales del significado de la asociación. Las principales dificultades fueron las siguientes:

- a) Confundir la frecuencia relativa condicional con la frecuencia doble (9 estudiantes) o con las frecuencias marginales (94 estudiantes) en las tablas de contingencia;
- b) Usar solamente una distribución marginal (1 estudiante en el problema 1) o comparar las distribuciones marginales de las dos variables en estudio (1 estudiante en el problema 4);
- c) Comparar el coeficiente de correlación de las variables en estudio con una variable diferente, debido a la falta de comprensión de los parámetros a utilizar como entrada en los programas (1 estudiante en el problema 4);
- d) Usar sus propias teorías previas en lugar de los datos que tienen que utilizar (problema 4);
- e) Interpretar la asociación de manera determinista (1 estudiante en el problema 3);

f) Utilizar el coeficiente de correlación en la misma variable en dos muestras relacionadas, con el objeto de estudiar la diferencia en esas dos muestras (1 estudiante en el problema 3) o el coeficiente de determinación (1 estudiante en el problema 3);

g) Comparar frecuencias absolutas en gráficos de barras, en lugar de utilizar frecuencias relativas (1 estudiante en el problema 2).

En general, observamos que los estudiantes prefieren resúmenes numéricos a los gráficos, especialmente en los problemas 1 a 3. Esto es debido al hecho de que cada gráfico disponible requiere su propia interpretación, que no siempre los estudiantes dominan. Además, los estudiantes han preferido resúmenes numéricos porque la idea de distribución es difícil para ellos, como se muestra en la confusión entre los diferentes tipos de frecuencias. Finalmente, señalamos el escaso uso de procedimientos inferenciales, posiblemente la comprensión de estos conceptos requiere un mayor periodo de enseñanza, antes de que los estudiantes decidan utilizarlos en la resolución de problemas.

### ***ANÁLISIS DETALLADO DEL CASO DE JUAN***

Además de la prueba final que hemos analizado, disponemos para cada alumnos de una prueba de evaluación inicial de ideas elementales de estocástica, el registro de las soluciones de las 8 relaciones de ejercicios realizadas, 2 pruebas escritas, y un proyecto personal de análisis de datos que realizaron 7 de los alumnos voluntariamente. Partiendo de esta información, en esta sección analizamos los conocimientos adquiridos por un estudiante (Juan) de los que realizaron el proyecto, así como su capacidad final de análisis de datos y los factores que han condicionado su aprendizaje.

Juan es un estudiante de primer curso de Psicopedagogía (4º año de estudios universitarios), que había estudiado antes la mayor parte de los contenidos estadísticos del curso, aunque no había manejado ordenadores ni software estadístico. Ha mostrado gran interés por la asignatura.

#### *1) Capacidad de planteamiento de problemas*

Fueron escasas las actividades en las que los estudiantes debían formular sus propias cuestiones sobre los ficheros de datos y no han recibido una atención suficiente durante el curso. En el caso de Juan podemos valorar esta dimensión a partir del proyecto personal de análisis de datos realizado sobre un tema elegido por él mismo: "*Hábitos de estudio de alumnos universitarios y su relación con la valoración que hacen de las calificaciones de sus profesores*". Juan ha mostrado importantes dificultades en articular su problema en el proyecto y en formular cuestiones específicas en términos de las variables que incorporó en el cuestionario. Manifestó un planteamiento confuso del problema a investigar, deficiente identificación de las variables y pobreza en el tratamiento y discusión de los resultados. Formula preguntas de carácter escolar y rutinario, como las siguientes: "*¿Cuántos alumnos tienen una nota media superior a 6? ¿Cuál es el percentil 25 del total?*"

Debemos tener en cuenta que el proyecto fue realizado de modo muy personal, sin el concurso del profesor, a excepción de alguna consulta en la preparación del cuestionario. La recogida de datos y la realización del informe se hizo al final del curso, en periodo de

exámenes finales. El análisis de los datos se realizó en horas extraordinarias, en las que el aula de informática estaba atendida por un becario, haciendo uso del horario de tutorías para las consultas al profesor. Estas circunstancias, unido a la propia dificultad de cualquier proyecto real de análisis de datos, son variables explicativas de las importantes carencias observadas en el comportamiento de Juan en este aspecto.

### *2) Grado de dominio del recurso informático y de los instrumentos estadísticos*

Juan ha alcanzado un grado alto de dominio del software utilizado y de las opciones específicas estudiadas, interesándose incluso por el manejo de procesadores de textos en horas de entrada libre al aula de informática. Hemos apreciado dificultades de comprensión de las representaciones estadísticas, gráficas y tabulares. Así, respecto del gráfico de la caja da la siguiente explicación en uno de los ejercicios propuestos:

*"El gráfico de la caja es un diseño pensado para informarnos de los siguientes estadísticos:*

*La media, que nos da la línea central que atraviesa el gráfico; la mediana que se representa por la línea perpendicular central que cruza la media; el rango máximo o mínimo viene definido por los salientes (bigotes) que salen del rectángulo; el recorrido intercuartílico que va desde el cuartil superior al cuartil inferior, es lo que se vería como un rectángulo delimitado por los bordes del rectángulo donde se expresan los cuartiles".*

Hay aspectos del gráfico de la caja que sí han sido comprendidos: *"En este gráfico se puede ver la simetría de la muestra si nos fijamos en el interior del rectángulo de la caja. Si el espacio entre la mediana y uno de los cuartiles es mayor que la distancia de la mediana al otro cuartil se puede decir que existe una asimetría".* Hay que hacer notar, sin embargo, la imprecisión de atribuir tal asimetría a la muestra, y no a la distribución de frecuencias de la variable estadística correspondiente.

### *3) Aspectos intensionales (conceptos y sus propiedades)*

Entre las respuestas a cuestiones planteadas en las relaciones de ejercicios hemos encontrado dificultades como las siguientes:

- No reconoce que una distribución de frecuencias debe venir dada por el conjunto de valores y sus respectivas frecuencias. Así, por ejemplo al interpretar una tabla de contingencia explica: *"Las distribuciones marginales de la variable sexo son: 33.3 y 67.7. Las distribuciones condicionales de 'deporte' según 'sexo=chico' son: 20.0, 45.0 y 35.0"*. Vemos que da las frecuencias absolutas, prescindiendo de informar de los valores de la variable correspondiente.
- Respecto a la idea de asociación hemos detectado una concepción local de la misma, ya que basa un juicio de asociación en la frecuencia de una única casilla en las tablas de contingencia.

### *4) Elementos procesuales y afectivos*

El proyecto personal realizado por Juan sugiere la existencia de limitaciones importantes de la capacidad de expresión verbal y gráfica, explicación y validación alcanzadas, componentes para las cuales no ha sido posible organizar situaciones didácticas específicas. Por el contrario, como hemos indicado mostró un gran interés y motivación, alcanzando un alto grado de dominio del software.

## REFLEXIONES FINALES

En la apertura de la Conferencia organizada por el IASE en Granada sobre las nuevas tecnologías Hawkins (1997) realiza una reflexión sobre las expectativas creadas por los ordenadores, respecto a su papel como instrumento didáctico en la enseñanza y aprendizaje, comparando los resultados obtenidos con dichas expectativas. Los resultados de nuestros experimentos pueden servir para centrar la discusión sobre algunos de los "mitos" resaltados por Hawkins:

*¿Facilita el ordenador la enseñanza de la estadística?*

Hay una gran diferencia entre las posibilidades teóricas de la tecnología como recurso didáctico y las posibilidades en un planteamiento concreto. Con frecuencia el software o el hardware puede ser inapropiado para los estudiantes. En nuestra experiencia, una barrera importante la pone el idioma, añadido a la lentitud de operación de los ordenadores en el aula de informática. El nulo o escaso conocimiento previo de los alumnos sobre el uso básico del ordenador hizo necesario dedicar parte del tiempo al aprendizaje del manejo básico del sistema operativo, y otros contenidos no estadísticos.

Otra dificultad se debe al hecho de que el software no se adapta a las necesidades docentes. Faltan posibilidades que el profesor necesitaría, por ejemplo, el uso de intervalos de diferente amplitud en los histogramas, y sobran muchos otros procedimientos no necesarios en un curso particular. La estructura de los menús del software no siempre es la más adecuada, porque algunas opciones de uso creciente se encuentran "escondidas" en un menú de tercer o cuarto nivel.

*¿Es suficiente introducir un ordenador en la clase de estadística, para realizar una innovación en nuestros métodos de enseñanza?*

Los ordenadores reducen sustancialmente el tiempo que antes se dedicaba al aprendizaje de los algoritmos de cálculo de los estadísticos y, a primera vista, podría pensarse que este tiempo puede dedicarse a actividades interpretativas y a profundizar en el estudio conceptual. Por otro lado, se añaden nuevos contenidos, tales como nuevas representaciones gráficas, la organización y estructuración de los datos y el uso del software estadístico.

Nuestra experiencia muestra que esto es claramente insuficiente para que la enseñanza sea realmente innovadora y que tampoco podemos olvidar por completo las fórmulas y las actividades tradicionales con "papel y lápiz". Es necesario un gran trabajo previo de diseño de las situaciones didácticas, incluyendo los ficheros de datos adecuados, y la evaluación de las concepciones previas de los alumnos. Todo el proceso de integración de los problemas, conceptos teóricos, discusiones colectivas, trabajo individual y colectivo plantea un gran número de problemas didácticos, pendientes aún de investigación.

*¿Facilitan los ordenadores la comprensión de los conceptos estadísticos?*

Hemos mostrado claramente las dificultades de los alumnos, para el caso concreto de la asociación estadística, debido a que los estudiantes necesitan captar y relacionar tres tipos diferentes de elementos, para dominar y utilizar el concepto en la resolución de problemas:

(1) *Elementos intensionales del significado*: tales como la diferencia entre dependencia funcional y estadística, relación entre correlación y regresión, las diferentes frecuencias relativas que se pueden extraer de una tabla de contingencia, el papel de las variables dependiente e independiente, parámetros en la ecuación de regresión e interpretación del coeficiente de correlación.

(2) *Elementos instrumentales/representacionales del significado*; El uso de diferentes herramientas para tratar o representar el concepto, como los gráficos de mosaico, diagramas de dispersión, tablas de doble entrada o series paralelas de gráficos de la caja, diagramas acumulativos o gráficos de barras.

(3) *Elementos extensionales de significado*: Las diversas situaciones-problema cuya solución necesita el estudio de la asociación, de las cuales hemos descrito una muestra en este trabajo.

La dificultad que supone la comprensión de todos estos elementos se muestra particularmente cuando se compara el significado de la asociación presentado en los experimentos de enseñanza (significado institucional en la institución particular de un curso introductorio de análisis exploratorio de datos) y el significado personal que los estudiantes finalmente han adquirido, donde solamente han adquirido parte del significado pretendido, y algunas de las concepciones incorrectas sobre la asociación permanecen.

*¿Mejoran los ordenadores el uso que se hace de la estadística?*

El análisis de datos es una actividad de alta cualificación aún a un nivel exploratorio, requiriendo una gran variedad de conocimientos sobre los problemas y conceptos fundamentalmente relacionados con los procedimientos estadísticos, gráficos, numéricos, descriptivos e inferenciales (Batanero y Truran, en prensa). Esto requiere la capacidad de seleccionar los mejores instrumentos para analizar y representar datos, flexibilidad en los cambios de los procedimientos de selección, adecuada interpretación de los resultados (elementos intensionales) y la habilidad para relacionarlos con los problemas (elementos extensionales). Aun en el caso de que nuestros estudiantes obtengan soluciones correctas a los problemas, podemos observar sus dificultades en cada paso de los procesos descritos.

Como se desprende de los dos experimentos, enseñar a un grupo de estudiantes con conocimientos y actitudes muy diversos esta actividad tan compleja, más allá de la rutina o tareas elementales no es una tarea sencilla y ciertamente requiere mayor tiempo del que se dedica a un curso introductorio de estadística. Los conocimientos didácticos por parte de los profesores tampoco son garantía suficiente para el aprendizaje.

El problema se acentúa porque el software estadístico es hoy día de acceso generalizado, incluso para los usuarios sin conocimientos estadísticos sólidos, quienes creen que el hecho de ser capaces de manejar un programa y realizar un gráfico u otro análisis elegido muchas veces, simplemente porque "está de moda", ya les capacita para el análisis de sus datos. Los ordenadores no pueden pensar, y por tanto, no son capaces de distinguir cuándo un método de análisis es o no adecuado a un problema particular. Incluso cuando el método fuese adecuado, la interpretación de los resultados no viene proporcionada por los ordenadores.

*¿Podremos guiarnos por los resultados de la investigación para progresar en el uso de los ordenadores en la enseñanza?*

Este es, quizás el punto más crítico, puesto que la investigación sobre la enseñanza de la estadística es aún muy incipiente y se concentra preferentemente en el estudio de las concepciones de los alumnos sobre conceptos elementales. No hay apenas investigación sobre las concepciones de los alumnos respecto a conceptos estadísticos avanzados, probablemente porque este tipo de conceptos ni se han incluido en el currículo de secundaria hasta muy recientemente. Más escasos son todavía los estudios de experimentos de enseñanza con o sin ordenadores.

Por otro lado, evaluar este tipo de experimentos es muy laborioso por la cantidad de datos generados y la ausencia de modelos previos para el análisis e integración de los mismos. La misma evaluación del trabajo de los alumnos con el ordenador plantea problemas de investigación específicos, porque es difícil transferir los resultados de la investigación sobre evaluación usando métodos tradicionales. Además la variedad de parámetros a tener en cuenta en el análisis de un experimento hace que estos sean difícilmente reproducibles en otros contextos, con otro software u otro tipo de alumnos. Todo ello muestra la necesidad de proseguir la investigación y reforzar las conexiones de ésta con la práctica docente.

**Agradecimientos:** Esta investigación se ha llevado a cabo dentro del proyecto PB96-1411 (Dirección General de Enseñanza Superior, M.E.C. Madrid)

## REFERENCIAS

- Batanero, C., Estepa, A., Godino, J. D. y Green, D. R. (1996): Intuitive strategies and preconceptions about association in contingency tables. *Journal for research in Mathematics Education*, 27(2), 151-169.
- Batanero, C., Estepa, A., Godino, J. D. (1997): Evolution of student's understanding of statistical association in a computer based teaching environment. In J. B. Garfield, & G. Burrill (Eds.), *Research on the Role of Technology in Teaching and Learning Statistics* (pp. 191-205). Voorburg, The Netherlands: International Statistical Institute.
- Batanero, C. y Godino, J. D. (1998): Understanding graphical and numerical representations of statistical association in a computer environment. *Proceedings of the V International Conference on Statistical Education*. Singapore
- Batanero, C., Godino, J. D. y Estepa, A. (1998): Building the meaning of association through data analysis activities. Research Forum. *22 conference of the International Group for the Psychology of Mathematics Education*. Stellembosch, Sudáfrica.
- Batanero, C. y Truran, J. (en prensa). Some insights from research on advanced stochastic thinking. *International Statistical Review*.
- Beyth-Marom, R. (1982): Perception of correlation re-examined. *Memory and Cognition*, 10(6), 511-519.
- Bielher, R. (1994). Software tools and mathematics Education: the case of statistics. En C. Keitel and R. Ruthnen (Eds.) *Learning for computers: mathematics education & technology* (pp. 68-100). Berlin: Springer Verlag.
- Bielher, R. (1997). Software for learning and for doing statistics. *International Statistical Review*, 65(2), 167-190.



- Chapman, L. J. & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology*, 74, 271-280.
- Crocker, J. (1981). Judgement of covariation by social perceivers. *Psychological Bulletin*, 90 (2), 272-292.
- Estepa, A. y Batanero, C. (1996). Judgements of correlation in scatter plots. Student's intuitive strategies and preconceptions. *Hiroshima journal of Mathematics Education*, 4, 21-41.
- Estepa, A. y Sánchez-Cobo, F. (1996). Association Judgements in the comparison of two samples. En L. Puig y A. Gutiérrez (Eds.) *Proceeding of the 20th. Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 337-344). Universidad de Valencia.
- Estepa, A., Green, D. R., Batanero, C. y Godino, J. D. (1994): Judgements of association in contingency tables. An empirical study of student's strategies and preconceptions. En J. P. Ponte y J. F. Matos (Eds.), *Proceedings of the XVIII International. Conference on the Psychology of Mathematics Education* (v.2, pp 312-319). Universidad de Lisboa.
- Falk, R. (1986). Conditional probabilities: insights and difficulties. En R. Davidson y J. Swift (Eds.), *Proceedings of the second conference on teaching statistics* (pp. 292-297). Voorburg: International Statistical Institute.
- J. B. Garfield & G. Burrill (1997) (Eds.), *Research on the Role of Technology in Teaching and Learning Statistics*. Voorburg, The Netherlands: International Statistical Institute.
- Godino, J. D., Batanero, C. y Estepa, A. (1991). Task variables in Statistical Problem Solving Using Computer. In J. P. Ponte, J. F. Matos, & D. Fernandes (Eds.). *Problem Solving and New Technologies. Research in Contexts and Practice*. Berlin: Springer Verlag.
- Godino, J. D. y Batanero, C., (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques*, 14 (3), 325-355.
- Godino, J. D. y Batanero, C., (1998). Clarifying the meaning of mathematical objects as a priority area of research in mathematics education. En J. Kilpatrick y A. Sierpiska (Eds.), *Mathematics Education as a research domain: A search for identity* (pp.177-195). Dordrecht : Kluwer A. P.
- González, J. L. y otros (1990). *Números enteros*. Madrid: Síntesis.
- Hawkins, A. (1997). Myth-Conceptions. En J. Garfield y G. Burrill (Eds.), *Research on the Role of Technology in Teaching and Learning Statistics*. Voorburg, The Netherlands: International Statistical Institute.
- Inhelder, B. y Piaget, J. (1955). *De la logique de l'enfant à la logique de l'adolescent*. Paris: Presses Universitaires de France.
- Jennins, D. L., Amabile, M. T. y Ross, L. (1982). Informal covariation assesment: Data based versus theory-based judgements. In D. Kahneman, P. Slovic & A. Tversky (Eds.). *Judgement under uncertainly: Heuristics and biases* (pp. 211-230). New York. Cambridge University Press.
- Moore, D. S. (1995). *The basic practice of statistics*. New York: Freeman.
- Morris, E. (1997). An investigation of student's conceptions and procedural skills in statistical topic correlation. Centre for Information Technology in Education. The open University. Cite Report N° 230.
- Shaughnessy, J. M., Garfield, J. y Greer, B. (1996). Data handling. En A. Bishop et al. (Eds.), *International handbook of mathematics education* (v.1, pp. 205-237). Dordrecht: Kluwer, A. P.