

# Reflexiones sobre la enseñanza experimental de la estadística

por

DANIEL PEÑA

Departamento de Estadística y Econometría  
Universidad Carlos III de Madrid

## RESUMEN

Este trabajo presenta algunas ideas sobre la organización de un curso experimental de estadística dirigida a científicos sociales o naturales para los que la estadística tiene un importante papel instrumental. Se resalta la necesidad de un método activo de aprendizaje y de un sistema continuo de seguimiento y control de los conocimientos adquiridos por los estudiantes. La estadística debe considerarse una asignatura experimental y enseñarse con un fuerte apoyo de un aula o laboratorio informático. Es tan importante el contenido de lo que enseñamos como el método utilizado, y debemos adecuar ambos aspectos a los objetivos explícitos de la enseñanza.

*Palabras clave:* Enseñanza de la Estadística. Trabajo en grupos. Aulas Informáticas. Probabilidad. Método de Montecarlo.

*Clasificación AMS:* 6001, 60A05, 6201.

## 1. INTRODUCCION

El contenido de la estadística como disciplina científica ha evolucionado enormemente en los últimos treinta años. Este cambio ha sido debido principalmente a la revolución en los métodos de cálculo, que ha permitido la utilización masiva de modelos estadísticos de complejidad creciente. Por ejemplo, la regresión múltiple, los modelos lineales generalizados, los modelos ARIMA de series temporales o los modelos de análisis multivariante han dejado de ser técnicas especializadas de escasa aplicación práctica para convertirse en herramientas habituales en muchos campos científicos. Por otro lado, estas facilidades de cálculo han permitido desarrollar métodos más flexibles y potentes de estimación y diagnosis (desde la estimación no paramétrica de curvas a las técnicas de remuestreo), que no requieren las hipótesis convencionales necesarias para construir los modelos clásicos.

Estos cambios en la teoría y la práctica estadística han dado origen a numerosas conferencias y artículos científicos donde se analizan críticamente los métodos actuales de enseñanza y se proponen alternativas de cambio. Véanse como muestra los trabajos de Bisgaard (1991), Box (1976), Ehrenberg (1983), Hogg *et al.* (1985), Makuch *et al.* (1990), Moore y Roberts (1989), Mosteller (1988), Singer y Willet (1990), Taner (1985), Vanderman y David (1984) y las actas de las conferencias «Making Statistics more effective in Business Schools» (1988) (1990). En nuestro país, recientemente la revista *Estadística Española* ha dedicado dos números monográficos (122, 3.º de 1989, y 123, 1.º de 1990) a un debate sobre este tema.

Un aspecto destacado en este debate es cómo orientar un curso de introducción a la estadística en áreas como la Economía, la Administración de Empresas o la Ingeniería, donde la estadística tiene un importante papel instrumental. Aunque naturalmente el programa concreto deberá variar en función de los conocimientos previos e intereses de los estudiantes, existen algunos principios generales sobre los que parece existir un amplio consenso:

- 1) La estadística es la herramienta básica para los profesionales que utilizan información cuantitativa. En consecuencia, el curso debe basarse en mostrar su relevancia para la adquisición de conocimientos en el área de especialización de los estudiantes mediante el análisis estadístico de datos reales asociados a problemas relevantes.
- 2) Las aplicaciones actuales de la estadística requieren, de forma imprescindible, el uso del ordenador. En consecuencia, este aspecto debe ocupar un lugar central en el diseño del curso, que debe dar más peso a las interpretaciones y limitaciones de los métodos presentados que a los detalles concretos de cálculo.

- 3) Los previsibles cambios en la tecnología informática en el próximo futuro van a ampliar de forma decisiva las capacidades de los métodos estadísticos. En consecuencia, es más importante desarrollar en el estudiante la capacidad de entender la variabilidad y pensar en términos estadísticos sobre problemas reales, con toda su complejidad, que enseñarle técnicas específicas para resolver problemas simples idealizados.

Estos tres principios indican que tan importante es debatir el *contenido* de los cursos como el *método* de transmisión de los conocimientos. En este artículo se presentan algunas reflexiones para el diseño de un primer curso de Estadística basado en los principios anteriores. El resto de su estructura es como sigue: en la sección 2 se presentan los objetivos del curso, que orientan el método pedagógico expuesto en la sección 3. La sección 4 analiza la utilización de la simulación en el curso como ayuda del aprendizaje. Finalmente, el artículo concluye comentando brevemente la puesta en marcha de estas ideas en la sección 5.

## 2. LOS OBJETIVOS DEL CURSO

La formulación de objetivos docentes es un aspecto crucial en el diseño de un curso: la planificación de la enseñanza, los medios utilizados y los procedimientos de evaluación van a definirse en función de estos objetivos. Su definición precisa: 1) informa al alumno del propósito del aprendizaje, y de lo que se espera de él, al término del mismo; 2) le ayuda a estudiar selectivamente, diferenciando el material básico del complementario; 3) facilita el diseño de pruebas de evaluación, que pueden establecerse sin ambigüedades; 4) permite una evaluación más precisa del contenido de la enseñanza y su adaptación para cubrir los objetivos.

Siguiendo a Bloom (1972) y Krazhwohl, Bloom y Masia (1973), los objetivos pedagógicos pueden clasificarse en objetivos de conocimiento y de actitudes. A su vez, los objetivos de conocimiento pueden subdividirse en seis categorías que corresponden a una mayor o menor utilización del potencial intelectual en relación con la materia de estudio.

### 2.1. Objetivos de conocimiento

Los objetivos globales de conocimiento que la enseñanza de la estadística debe alcanzar en el primer ciclo universitario para economistas o ingenieros son los siguientes:

- 1) El alumno debe ser capaz de definir poblaciones que puedan ser investigadas estadísticamente. Realizar hipótesis acerca de su estructura. Razonar un procedimiento de muestreo adecuado. Describir estadísticamente los datos muestrales. Comprobar su ajuste a un modelo de probabilidad mediante un test de hipótesis. Interpretar críticamente los resultados obtenidos y señalar las consecuencias del análisis.
- 2) El alumno debe ser capaz de plantear un modelo lineal (regresión o análisis de la varianza), para estudiar las relaciones entre variables, conocer los procedimientos de estimación, ser capaz de interpretar una salida habitual de ordenador y saber aplicar los contrastes diagnósticos para juzgar la validez del modelo.
- 3) El alumno debe ser capaz de plantear formalmente un problema sencillo de decisión en condiciones de incertidumbre. Construir una función de utilidad y evaluar el beneficio esperado de recoger información adicional. Tomar una decisión justificándola críticamente mediante un estudio de sensibilidad.
- 4) El alumno debe ser capaz de presentar las conclusiones de los trabajos anteriores por escrito en una secuencia lógica: Objetivos del Estudio-Hipótesis básicas-Métodos utilizados-Análisis de los Datos-Conclusiones.

Además de estos objetivos básicos, cada unidad didáctica tendrá unos objetivos de conocimiento concretos que se plasmarán en los programas de las asignaturas en que se organice la formación estadística de los estudiantes.

## **2.2. Objetivos de actitudes**

Los objetivos de actitudes son tan importantes como los de conocimientos. Si los segundos definen qué debemos transmitir, los primeros determinan cómo debemos transmitirlo. En concreto, los objetivos de actitudes seleccionados son:

- 1) Desarrollar en el alumno una actitud científica y antidogmática ante la realidad. Esta actitud debe traducirse en un comportamiento observable de: a) diferenciar las opiniones contrastables empíricamente de las que no lo son; b) adquirir el reflejo de criticar análisis incorrectos de datos y conclusiones obtenidas sin fundamento suficiente.
- 2) Generar una actitud positiva hacia la teoría, siendo consciente de la insuficiencia de un empirismo puro para obtener conclusiones de la realidad.

- 3) Desarrollar la capacidad de comunicación verbal y escrita sobre datos estadísticos y la capacidad de argumentar en grupo sobre los mismos.
- 4) Fomentar la imaginación, la autonomía personal y la seguridad numérica de los estudiantes.

En la sección siguiente se presenta un método pedagógico diseñado para conseguir estos objetivos.

### 3. METODO PEDAGOGICO

El método pedagógico que proponemos tiene cuatro pilares básicos: 1) el trabajo del estudiante en la clase; 2) su trabajo en casa sobre ejercicios prefijados que debe después entregar; 3) el trabajo de fin de curso, que sirve de síntesis a las distintas partes de la asignatura; 4) un flujo constante de información bidireccional profesor/alumno sobre la marcha del curso.

Comenzando con el primer aspecto, la clase tradicional de una hora de duración es demasiado corta para incluir el trabajo del estudiante. En el esquema tradicional el profesor trabaja en clase, explicando la lección, y el alumno en casa, comprendiendo las explicaciones y resolviendo ejercicios de comprobación. Este esquema, que es probablemente válido para cursos avanzados con estudiantes muy motivados, es inadecuado para conseguir los objetivos de actitudes antes especificados: 1) favorece la actitud pasiva del alumno en clase; 2) es poco eficaz para generar el interés en alumnos que no tengan una motivación especial por la asignatura o un excelente profesor; 3) conduce a que una parte del tiempo de clase se pierda en exposiciones rutinarias que pueden encontrarse en cualquier libro de texto.

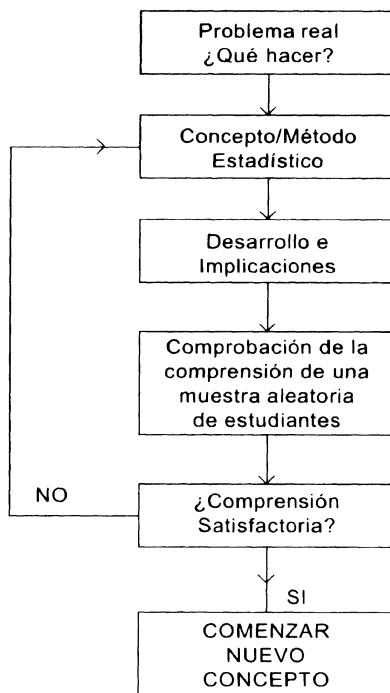
Pensamos que la función del profesor *no* es repetir la lección que los alumnos pueden encontrar escrita en libros, sino facilitar su aprendizaje explicando la relevancia de los conceptos principales y sus implicaciones, promoviendo el interés de los estudiantes mediante la presentación de aplicaciones de la estadística a problemas relevantes y desarrollando su capacidad para pensar y argumentar en términos estadísticos. Este planteamiento requiere, en contrapartida, que los alumnos estudien por sí mismos, aunque de forma dirigida y controlada, los detalles de los métodos cuyos fundamentos se han explicado en clase.

Para utilizar este enfoque recomendamos agrupar las horas de clase disponibles en módulos de dos horas. Estos módulos se impartirán según dos esquemas posibles que se describen a continuación.

### 3.1. El Módulo Teórico-Práctico (Módulo A)

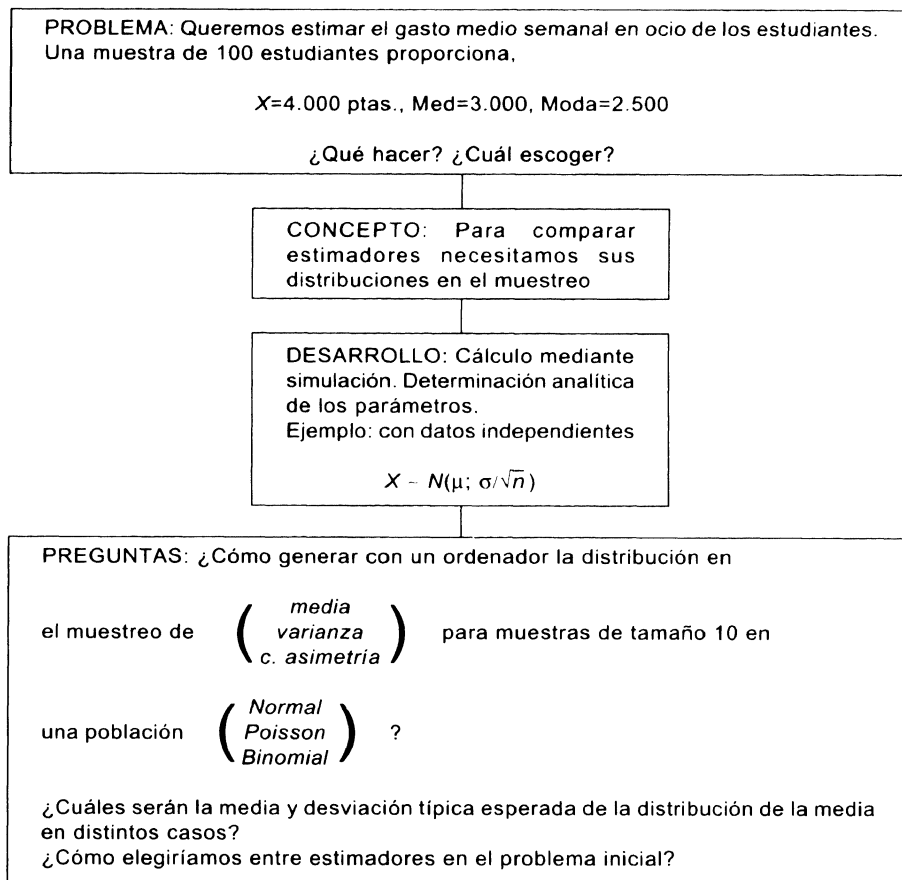
El objetivo de este módulo es presentar nuevos conceptos y métodos y explicar su utilización práctica. La primera hora del módulo se dedica a la presentación del tema, y debe comenzar siempre describiendo un problema y preguntando a los alumnos por sus ideas para resolverlo. Una vez que la clase ha comprendido la necesidad de introducir nuevas herramientas o conceptos, se pasará a su exposición y desarrollo. Finalizada ésta, es recomendable preguntar *al azar* a los alumnos para comprobar que han captado el mensaje transmitido, identificar fuentes de confusión o error y, en definitiva, comprobar el grado de comprensión de la explicación. Si ésta no es satisfactoria, se repasarán las dificultades observadas hasta asegurar la asimilación de las ideas por la mayoría de la clase. A continuación se procederá a la exposición de un nuevo concepto, y así sucesivamente. Este esquema se halla resumido en el cuadro 1. El cuadro 2 presenta un ejemplo de este método para explicar el concepto de distribución en el muestreo.

**Cuadro 1**  
ESTRUCTURA DE LA PRIMERA HORA DEL MODULO A



## Cuadro 2

### EJEMPLO DE PRESENTACION DEL MODULO A PARA EL CONCEPTO DE DISTRIBUCION EN EL MUESTREO DE UN ESTIMADOR



En la segunda hora del módulo A, la clase se divide en grupos de 4 ó 5 personas que resuelven ejercicios sobre los conceptos y métodos explicados en la primera parte. El trabajo en grupo se organiza con las características siguientes:

- 1) Se concede un tiempo limitado por ejercicio.
- 2) El profesor debe estimular las ideas creativas, *aunque sean erróneas*, y animar a los estudiantes a experimentar y comprobar por ellos mismos el resultado. Es importante llevar al ánimo de los alumnos que pensar

con autonomía implica necesariamente cometer errores y que éstos son parte también del proceso de aprendizaje.

- 3) Es imprescindible desarrollar en los estudiantes la intuición de lo que hacen. Para ello conviene pedirles que, ante cualquier problema, intenten obtener una respuesta aproximada rápida antes de realizar ningún cálculo complicado. Con esto se consigue desarrollar un sentido de «orden de magnitud» de la respuesta, y estimular el hábito de intuir una respuesta aproximada ante un problema planteado.

La verificación del trabajo de los estudiantes puede hacerse preguntando al azar a varios grupos y eligiendo de nuevo al azar a la persona que hablará en nombre del grupo.

El módulo finaliza asignando a los estudiantes ejercicios para casa que deberán entregar en la clase siguiente. Estos ejercicios tienen principalmente por objeto dirigir el estudio personal sobre los detalles de los métodos presentados y contrastar la adecuada comprensión de los mismos.

El diseño de este módulo está orientado a:

- 1) Obligar al profesor a estructurar y seleccionar la información a transmitir. Cada clase tiene unos conceptos básicos que deben ser aprendidos y el profesor debe asignar un tiempo a cada uno de ellos que dependa de su relevancia posterior y su dificultad. Esta idea de prioridades debe ser claramente transmitida a los alumnos.
- 2) Proporcionar información inmediata de la comprensión de los estudiantes. Las preguntas dirigidas a estudiantes seleccionados al azar (uno de cada diez, por ejemplo, por posición en la clase) garantiza que se recoge adecuadamente la comprensión del grupo. Preguntar hacia el vacío, tiene el riesgo de oír sólo las respuestas de los alumnos aventajados, y las preguntas a alumnos concretos pueden originar variadas suspicacias en los estudiantes y terminan reflejando los sesgos del profesor. Sin embargo, los alumnos comprenden rápidamente que las preguntas al azar tienen por objeto proporcionar información fiable al profesor de la comprensión de la clase en conjunto, lo que generalmente agradecen.
- 3) Fortalecer la confianza del estudiante a expresarse y pensar por sí mismo en términos estadísticos. Tanto en la clase como en el trabajo en grupo, el estudiante se ve forzado a expresar su opinión y a razonar en términos estadísticos ante un problema.
- 4) Acostumbrar al estudiante a cuestionarse y preguntarse, en lugar de memorizar. El estudiante debe comprender que cualquier disciplina cien-



tífica no es un conjunto de resultados cerrados e inamovibles, sino el resultado histórico concreto de un proceso de resolución de problemas de complejidad creciente. Con seguridad, el conocimiento de hoy será total o parcialmente obsoleto mañana, o tendrá dimensiones distintas de las que hoy percibimos.

### **3.2. El Módulo de Laboratorio (Módulo B)**

Este módulo tiene dos objetivos principales. El primero es mostrar la utilidad de la estadística con ayuda de la informática para resolver problemas reales. El segundo es enseñar una destreza específica, que incluye el manejo de métodos estadísticos mediante el ordenador. En general, la primera parte del módulo se desarrolla en la clase, mientras que la segunda debe hacerse en un laboratorio de cálculo o aula informática. No obstante, con grupos pequeños ambas partes pueden fusionarse y realizarse íntegramente en el aula informática. El diseño de este módulo está basado en las experiencias de Rafael Romero y sus colaboradores en la Universidad Politécnica de Valencia, que han obtenido con él excelentes resultados (véase Peña, Prat y Romero, 1990).

La clase comienza con la presentación de un problema, que puede analizar datos o evaluar procedimientos estadísticos. Si se presentan datos es imprescindible que correspondan a una situación real y que el alumno reciba información suficiente para comprender el interés de su estudio y el contexto en que se presenta el problema. Si se evalúan procedimientos estadísticos (un método de estimación, un contraste de hipótesis, etc.) es conveniente referirlos a un problema real donde aparezca de manera natural la necesidad de esta comparación.

Una vez presentado el problema se discutirá con los estudiantes el procedimiento a emplear y se realizará el análisis mediante el uso de un paquete informático, cuyos resultados se proyectarán con una pantalla de cristal líquido.

Es importante que el paquete a utilizar sea adecuado a la formación previa de los alumnos, de manera que su uso resulte una ayuda, en lugar de una dificultad adicional. También conviene que disponga de buena capacidad gráfica, y que sea flexible, para permitir análisis no convencionales de los datos. De los paquetes que conocemos hemos utilizado MINITAB y STATGRAPHICS. El primero es más adecuado para estudiantes con buena formación matemática, y el segundo más simple de utilizar y mucho mejor en su capacidad gráfica. En conjunto, pensamos que el segundo es más apropiado para un curso introductorio, aunque indudablemente tiene sus limitaciones, especialmente para las simulaciones que comentamos en la sección siguiente.

Si el número de estudiantes y la capacidad del aula informática lo permite, este trabajo debe realizarse individualmente. En general, en la mayoría de los casos habrá que dividir a los estudiantes en grupos. En nuestra experiencia, para que el trabajo sea efectivo el tamaño máximo del grupo debe ser de tres estudiantes.

En la segunda parte de la clase los alumnos repiten en un aula informática el análisis efectuado por el profesor en clase. Posteriormente, los estudiantes que lo deseen pueden realizar análisis alternativos de los datos. Después de esta sesión, y como ejercicio para casa, se pide a los estudiantes que efectúen un análisis similar con un conjunto distinto de datos y preparen un breve informe con sus conclusiones.

Este módulo está orientado a:

- 1) Asegurar el aprendizaje de los estudiantes del manejo del ordenador para un análisis concreto.
- 2) Mostrar un caso real de utilización de métodos estadísticos.
- 3) Estimular la curiosidad del estudiante por repetir el mismo análisis con otros datos.
- 4) Desarrollar la capacidad de síntesis ante un análisis estadístico.

### **3.3. El trabajo de fin de curso**

El tercer componente del método pedagógico es el trabajo de fin de curso. En este trabajo los alumnos deben aplicar a un problema real elegido por ellos los métodos estudiados durante el curso. El formato del trabajo que deben presentar se encuentra en el Apéndice I. En nuestra experiencia, la realización de este trabajo es una parte importante del proceso de aprendizaje, sirve para unificar los distintos temas tratados en el curso y contribuye a la generación de una actitud positiva hacia la utilidad de la estadística. Recientemente, Radke-Sharpe (1991) ha argumentado convincentemente respecto a la importancia de esta práctica en los cursos introductorios de estadística.

El trabajo correspondiente a un curso básico cuatrimestral de introducción a la Estadística debe basarse en un análisis de datos y en el ajuste a ellos de un modelo probabilístico de acuerdo con el primer objetivo global de conocimientos de la sección 2.

Los objetivos segundo y tercero especifican claramente la orientación del trabajo para cursos que cubran estos aspectos. Peña (1991) presenta ejemplos

de trabajos correspondientes a los dos primeros objetivos para alumnos de ingeniería. Hunter (1977) ha presentado numerosos ejemplos de diseño de experimentos realizados por alumnos.

Este trabajo de fin de curso es voluntario y sirve para mejorar la nota. En nuestra experiencia los estudiantes aprenden mucho con su realización y es frecuente que perciban el goce intelectual que supone el descubrimiento de lo inesperado en la investigación científica.

### **3.4. El control de aprendizaje**

Finalmente, el cuarto componente del método pedagógico es el flujo constante de información alumno/profesor sobre la marcha del curso. Todo el esquema de presentación del material está orientado a favorecer este mecanismo.

Es bien conocido que un buen diseño de un sistema requiere de forma imprescindible un sistema de control adecuado para corregir las desviaciones respecto al objetivo. Los mecanismos de control del aprendizaje en este método son:

- 1) Las preguntas a los estudiantes en clase. Este es un método muy eficaz ya que permite corregir inmediatamente falsas interpretaciones y obliga a los estudiantes a pensar y cuestionar lo explicado. Una herramienta adicional que hemos utilizado con éxito, y que ha sido defendida convincentemente por Mosteller (1988), es pedir a los estudiantes que pongan por escrito anónimamente, después de cada módulo teórico-práctico, qué aspectos de la exposición han resultado más confusos y les han despertado más dudas. Esta miniencuesta requiere tres minutos después de cada clase y puede ser muy eficaz para reorientar la clase siguiente y establecer una comunicación eficaz entre profesor y estudiantes.
- 2) Los ejercicios realizados en clase y en el aula informática. Cada actividad encomendada a los estudiantes se controla y proporciona información directa sobre su asimilación.
- 3) Los ejercicios para casa, que aseguran el trabajo adicional del estudiante sobre lo explicado.
- 4) El trabajo fin de curso.
- 5) El examen final.

De esta manera el profesor dispone de abundante información respecto a la marcha de cada estudiante y éste de numerosas señales respecto a lo exigido para aprobar. En consecuencia, es esperable que un método como el aquí descrito lleve asociado una alta tasa de aprobados entre los alumnos que lo sigan.

#### **4. LOS METODOS GRAFICOS Y LA SIMULACION COMO AYUDA AL APRENDIZAJE DEL CALCULO DE PROBABILIDADES**

Los cursos tradicionales de Estadística son famosos entre los estudiantes de muchas Universidades como los más aburridos e inútiles que han recibido. Además, son generalmente considerados como particularmente incomprensibles.

En la sección anterior hemos presentado un esquema metodológico que conduce, tanto en nuestra experiencia como en la de nuestros colegas (véase Peña, Prat y Romero, 1990), a una alta motivación e interés por parte de los participantes. Sin embargo, la dificultad de comprensión de muchos de los conceptos estadísticos requiere un comentario aparte. Muchos profesores han argumentado que esta dificultad radica en que los estudiantes no sólo deben comprender conceptos muy abstractos, sino ser capaces de relacionarlos inmediatamente con la realidad. Véase Watts (1991) para una discusión reciente de estas ideas.

La dificultad de comprensión de los conceptos básicos de la inferencia estadística es un problema real, como lo demuestra el largo período de maduración que ha necesitado la humanidad hasta asimilarlos. Por ejemplo, mientras los conocimientos de álgebra y cálculo que recibe un estudiante de primer ciclo provienen en gran parte de los siglos XVIII y XIX, la inferencia estadística ha sido creada íntegramente en el siglo XX, y muchas de las ideas y técnicas que aparecen en los libros introductorios han sido comprendidas y aplicadas en los últimos cuarenta años.

En consecuencia, es conveniente partir de esta dificultad objetiva y diseñar métodos para contrarrestarla. De las tres partes fundamentales de un curso básico de introducción a la estadística (Descriptiva, Probabilidad e Inferencia), la parte de cálculo de probabilidades es generalmente la más difícil, especialmente para alumnos con poca vocación matemática.

Los métodos gráficos y la simulación de problemas con el ordenador pueden ser una ayuda valiosa para su comprensión, como indicamos a continuación.

#### 4.1. Diagramas probabilísticos

Los conceptos de independencia entre sucesos y probabilidad condicionada, incluyendo el teorema de Bayes, son abstractos y difíciles para los estudiantes de cursos introductorios. Mi experiencia es que estas ideas pueden entenderse mejor utilizando un diagrama simple donde las probabilidades de los sucesos se representan por rectángulos dentro de un cuadrado de área unidad que representa la probabilidad del suceso seguro.

Estos diagramas se diferencian de los diagramas de Venn en: 1) representan *probabilidades* en lugar de sucesos; 2) utilizan rectángulos en lugar de contornos circulares. Las ventajas de modificar los diagramas de Venn utilizando rectángulos han sido señaladas por Berrondo (1990). Los diagramas probabilísticos que proponemos recogen esta sugerencia, pero aplicada a probabilidades en lugar de a sucesos.

En un diagrama probabilístico la probabilidad de cada suceso se asocia al área de un rectángulo. El suceso seguro,  $P(E)$ , se representa por un cuadrado de lado unidad; dada una clase completa de sucesos  $(A_1, \dots, A_n)$  con  $A_i \cap A_j = \emptyset$  y  $\bigcup_{i=1}^n A_i = E$ , las probabilidades  $P(A_i)$  se representan por rectángulos de base  $P(A_i)$  y altura unidad. Si consideramos ahora otro suceso,  $B$ , podemos considerar los casos siguientes:

- 1)  $B$  es independiente de todos los sucesos  $A_i$ , es decir,

$$P(BA_i) = P(B) P(A_i) \quad \forall i$$

entonces el suceso  $B$  se representa por un rectángulo de lado  $P(B)$ , y, como indica la figura 1, su intersección con cada suceso  $A_i$  proporciona automáticamente un área igual a la probabilidad del suceso conjunto  $(B A_i)$ . Naturalmente,

$$P(B) = \sum P(BA_i) = P(B) \sum P(A_i)$$

- 2)  $B$  y  $A_i$  son dependientes, para algún suceso  $A_i$ . Entonces,

$$P(B A_i) = P(B|A_i)P(A_i)$$

y la probabilidad del suceso conjunto se obtiene multiplicando cada base,  $P(A_i)$ , por una altura distinta  $P(B|A_i)$ , que puede ser mayor o menor a  $P(B)$ . Geométricamente, es obvio en la figura 2 que  $P(B)$  representa una «media» ponderada de la probabilidad  $P(B|A_i)$

$$P(B) = \sum P(B|A_i)P(A_i)$$

**Figura 1**  
 DIAGRAMA PROBABILISTICO PARA SUCESOS  $B$  Y  $A$ , INDEPENDIENTES

	$P(A_1)$	$P(A_2)$	$P(A_3)$	$P(A_4)$	$P(A_5)$
$P(B)$		$P(BA_2)$	$P(BA_3)$		

**Figura 2**  
 DIAGRAMA PROBABILISTICO PARA SUCESOS  $A_i$  y  $B$  DEPENDIENTES

	$P(A_1)$	$P(A_2)$	$P(A_3)$	$P(A_4)$	$P(A_5)$
$P(B)$		$P(BA_2)$	$P(BA_3)$		

La probabilidad  $P(B|A_i)$  vendrá dada por el área del suceso  $(B \cap A_i)$  cuando el suceso  $A_i$  tiene área unitaria. Supongamos que tenemos una partición de  $A_j$ ,  $B_j A_i$ ,  $j=1, \dots, m$ , de manera que  $A_i = \bigcup_{j=1}^m (B_j A_i)$ . Entonces para obtener  $P(B_j|A_i)$  tendremos que dividir  $P(B_j A_i)$  por  $P(A_i)$  para que la suma de las probabilidades sea la unidad.

El Teorema de Bayes se obtiene también trivialmente por consideraciones geométricas. Supongamos que se desea calcular  $P(A_i|B)$  en la figura 2. Por definición, ésta es la probabilidad  $P(A_i B)$  corregida para que  $P(B)$  sea la unidad. Entonces  $P(A_i|B)$  será la fracción que  $P(A_i B)$  representa sobre el total de  $P(B)$ , es decir,

$$P(A_i|B) = \frac{P(A_i B) P(B)}{\sum P(A_j|B) P(B)}$$

que es el Teorema de Bayes.

Para ilustrar las ventajas de estos diagramas consideremos el ejemplo de las tres puertas, que ha sido recientemente objeto de una cierta polémica en *The American Statistician* (véase Morgan *et al.*, 1991). El problema es el siguiente: un concursante tiene que elegir entre tres puertas, detrás de una de las cuales se encuentra un premio. Supongamos que el concursante elige la puerta 1. A continuación el presentador le muestra que en la puerta 2 no está el premio y le ofrece la posibilidad de reconsiderar su decisión. ¿Debe cambiar de puerta el concursante y elegir la puerta 3?

Para resolver el problema, sea  $R_i$  el suceso el premio está en la puerta  $i$ , y  $B_i$  el suceso el presentador muestra la puerta  $i$ . Tendremos que calcular las probabilidades  $P(R_1|B_2)$  y  $P(R_3|B_2)$  y compararlas. Es claro que, *a priori*,  $P(R_i)=1/3$ . Las probabilidades conjuntas de los sucesos  $(B_2 R_i)$  son

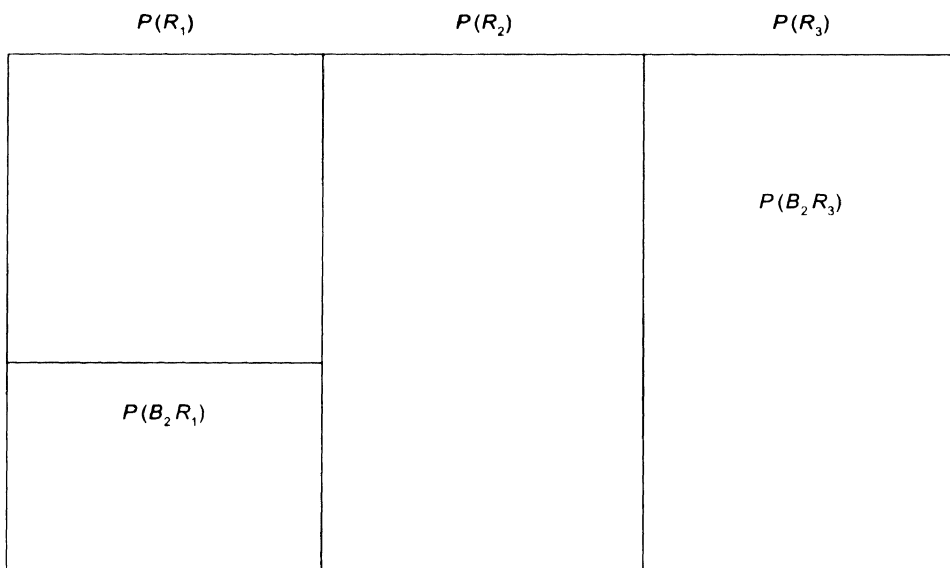
$$P(B_2 R_i) = P(B_2|R_i) P(R_i)$$

y por tanto necesitamos las condicionadas  $P(B_j|R_i)$ . Suponiendo que cuando el premio está en la puerta 1 el presentador muestra al azar cualquiera de las dos puertas 2 y 3,  $P(B_2|R_1)=0,5$ . Por otro lado,  $P(B_2|R_2)=0$  y  $P(B_2|R_3)=1$ , si suponemos que el presentador *siempre* debe dar una segunda oportunidad al concursante. Por tanto,  $P(B_2 R_1)=1/6$ ;  $P(B_2 R_2)=0$ ;  $P(B_2 R_3)=1/3$ . Estas probabilidades se representan en la figura 3. Es claro que  $P(B_2)=0,5$  y que

$$P(R_1|B_2) = \frac{P(R_1 B_2)}{P(B_2)} = \frac{1}{3}$$

mientras que  $P(R_3|B_2)=2/3$ , y siempre es mejor cambiar de puerta, ya que se dobla la probabilidad de ganar.

**Figura 3**  
**DIAGRAMA PROBABILÍSTICO PARA EL PROBLEMA DE LAS TRES**  
**PUERTAS CUANDO EL CONCURSANTE ELIGE LA 1 Y LE MUESTRAN**  
**LA 2 VACIA**



Este resultado, que resulta obvio con el diagrama probabilístico, es a primera vista contraintuitivo y más difícil de entender sin representaciones gráficas.

En resumen, el diagrama probabilístico presenta las ventajas siguientes:

- 1) Permite ilustrar de manera gráfica las nociones de probabilidad condicionada y conjunta de sucesos, así como la noción de independencia probabilística.
- 2) Hace trivial la obtención del teorema de la probabilidad total o el teorema de Bayes, que resultan de sumar áreas o estandarizar para distintos espacios muestrales.
- 3) Desarrolla una intuición geométrica de los valores de las probabilidades que puede evitar errores aritméticos groseros en su manipulación.
- 4) Presenta las probabilidades como áreas, lo que facilita la asociación posterior entre probabilidades para variables continuas y áreas bajo la función de densidad.



## 4.2. El método de Montecarlo

La simulación de problemas con ordenador mediante el método de Montecarlo puede utilizarse para ilustrar los puntos más difíciles e ilustrar los conceptos impartidos en las clases teóricas. En particular, este método puede utilizarse para:

- 1) Resolver aproximadamente problemas de probabilidades elementales calculando frecuencias relativas de sucesos, lo que permite al estudiante «experimentar» con los conceptos y ver su plasmación material en casos concretos.
- 2) Simular distribuciones y estudiar sus propiedades al variar sus parámetros. En especial, este método permite comprobar el teorema central del límite y desarrollar la intuición del alumno respecto a las aproximaciones a la normal.
- 3) Evaluar el efecto de transformaciones en las variables sobre su distribución.

Además de estas ventajas, el método de Montecarlo es muy útil en la parte de inferencia para:

- 1) Simular distribuciones en el muestreo. Este concepto, básico en la inferencia clásica, es generalmente difícil de asimilar por los alumnos. Sin embargo, la experiencia demuestra que es relativamente fácil comprenderlo cuando el alumno ha generado por sí mismo estas distribuciones mediante Montecarlo.
- 2) Comparar estimadores. Es frecuente que los alumnos de un curso básico finalicen con una vaga noción de estadísticos  $t$  e intervalos de confianza de dos sigma, pero sin una clara idea de las hipótesis requeridas para su uso ni de la robustez y fiabilidad de los procedimientos empleados. El método de Montecarlo permite a los estudiantes experimentar desde el principio con distintas distribuciones y juzgar en casos concretos la eficacia de distintos procedimientos de inferencia.
- 3) Comparar contrastes de hipótesis, tanto por su potencia como por su robustez.

El método de Montecarlo permite al alumno inquieto comprobar por sí mismo los resultados explicados en clase y potencia el pensamiento autónomo ante un problema. Su utilización presenta, además, la ventaja de proporcionar una herramienta general para resolver directamente problemas estadísticos más sofisticados (véase, por ejemplo, Simon y Bruce, 1991). Por otro lado, refuerza

la utilidad de la teoría estadística, cuando ésta puede aplicarse, para obtener soluciones exactas ante problemas cuya solución aproximada requiere un gran tiempo de cálculo.

## 5. CONCLUSIONES

El diseño y puesta en marcha de un curso experimental en estadística exige una considerable inversión inicial en preparación de material docente, programas e infraestructura. Por otro lado, requiere profesorado con experiencia en análisis de datos reales, acceso a un aula informática convenientemente dotada y clases de un máximo de 70/80 estudiantes, que puedan dividirse en dos grupos para acudir al aula informática. Es conveniente disponer para la clase de un ordenador portátil con una pantalla de cristal líquido, y la formación de grupos en ella requiere aulas con mobiliario desplazable. Finalmente, es muy necesario contar con apoyo institucional para resolver los problemas de infraestructura y logística que un cambio de este tipo genera.

Un curso de estas características puede ser inviable si el número de alumnos es excesivo, falta motivación en el profesorado para mejorar la enseñanza, o se carecen de los medios informáticos mínimos. Además, un curso de este tipo requiere una continua adaptación y puesta al día, que debe basarse en las encuestas de los estudiantes y en la evaluación y crítica de los resultados académicos obtenidos por éstos.

La experiencia demuestra que la enseñanza de la Estadística puede ser una experiencia atractiva, divertida y fructífera para alumnos y profesores. Un requisito imprescindible para que esto ocurra es que los alumnos decidan involucrarse seriamente en el proceso de aprendizaje, lo que requiere que comprendan la utilidad de la estadística para su formación. En primer lugar, como ciudadanos en un mundo donde la argumentación y la toma de decisiones con métodos estadísticos tiene un papel creciente; en segundo, como herramienta imprescindible en la adquisición de conocimientos en su campo específico; en tercero, como lenguaje de comunicación entre distintos campos científicos y apoyo básico del trabajo empírico interdisciplinario. Este convencimiento puede resultar de un acto de fe al escuchar al profesor, pero será probablemente más sólido y duradero si es adquirido a través de experiencias prácticas con problemas reales.

## AGRADECIMIENTOS

Mis primeras experiencias docentes con la filosofía aquí expuesta fueron realizadas en la Escuela de Organización Industrial, a mediados de los años setenta, impulsadas y dirigidas por el entonces Director, Juan Ramón Figuera. He aprendido mucho sobre la enseñanza de la estadística del excelente trabajo pedagógico realizado por Albert Prat, Universidad Politécnica de Barcelona; George E. Box, University of Wisconsin-Madison; Rafael Romero, Universidad Politécnica de Valencia, y George C. Tiao, The Graduate School of Business of the University of Chicago. Los comentarios de M. Berrondo me han ayudado mucho para ver las ventajas del diagrama probabilístico. Estoy muy agradecido a un evaluador anónimo de la Revista por su cuidadosa lectura y sus valiosos comentarios.

## A P E N D I C E I

### TRABAJOS DURANTE EL CURSO

Para realizar el trabajo, el alumno deberá presentar un proyecto del mismo al profesor de su grupo, y tras la aprobación de dicho proyecto por el profesor, seguir las pautas de control que éste le marque y entregarlo en el plazo que se indique. *No se admitirá ningún trabajo que no haya sido supervisado y controlado por los profesores.*

Los trabajos que se presenten durante el curso deberán seguir el formato siguiente:

Primera página: Título del trabajo, nombre del alumno, número y grupo.

A continuación, el texto se dividirá en las secciones siguientes:

1. *Objetivos del trabajo*

Se indicará de forma clara el objetivo que se pretende alcanzar con el trabajo realizado y el interés del mismo. También se revisarán, en su caso, trabajos previos que se conozcan sobre el tema estudiado.

2. *Los datos*

Se especificará la fuente de datos, y el procedimiento de recogida de los mismos. A continuación se presentarán en una tabla los datos mismos, a no ser que su tamaño sea muy grande, en cuyo caso se presentarán en un apéndice. Conviene comentar las limitaciones o defectos posibles del procedimiento de recogida de la información y cualquier comentario aclaratorio que se considere necesario.

3. *Análisis descriptivo de los datos*

Esta sección incluirá gráficos de los datos (histogramas, diagrama de barras, etc.), medidas características (medias, desviaciones típicas, etc.) y su interpretación. Conviene explorar si los datos pueden dividirse en grupos más homogéneos (según el tipo de datos, por día de la semana, sexo, localización geográfica, etc.), en cuyo caso se estudiarán separadamente estos grupos de menor variabilidad interna.

4. *Construcción del modelo*

Se indicarán las hipótesis básicas en que se fundamente el modelo (o modelos, si hay varias poblaciones), las estimaciones y contrastes de hipótesis pertinentes y los resultados del análisis estadístico. Debe incluirse siempre una sección sobre validación y diagnóstico del modelo construido.

5. *Conclusiones*

Se indicarán las conclusiones principales obtenidas y sus limitaciones, así como las extensiones posibles del trabajo que podrían aclarar los interrogantes no resueltos.

6. *Referencias*

Se listarán aquí, por orden alfabético, los libros o artículos citados en el texto. El formato a utilizar es el habitual en los textos y referencias manejados.

## REFERENCIAS

- ASA (1980): «Preparing statisticians for Careers in industry». Report of the ASA section on Statistical Education Committee on Training of Statisticians for Industry. *The American Statistician*, 34, 2, pp. 65-75.
- BERRONDO, M. (1990): «Vers une syntax des diagrammes de Venn». 2nd World Congress of Bernoulli Society. Uppsala. Suecia.
- BISGAARD, S. (1991): «Teaching Statistics to Engineers». *The American Statistician*, 45, 4, pp. 274-283.
- BLOOM, B. S., y otros (1972): «Taxonomía de los objetivos de la Educación. Ambito del Conocimiento». Ed. Marfil.
- Box, G. E. P. (1976): «Science and Statistics». *Journal of the American Statistical Association*, 71, p. 356.

- EHRENBERG, A. S. C. (1983): «We must Preach what is practiced». *The American Statistician*, 37, pp. 248-250.
- HOGG, R., *et al.* (1985): «Statistical Education for Engineers: An Initial Task Force Report». *The American Statistician*, 39, 3, pp. 168-175.
- HUNTER, W. (1977): «Some Ideas about Teaching Design of Experiments with 2 Examples of Experiments conducted by Students». *The American Statistician*, 31, pp. 12-17.
- KRATHWOHL, D. R.; BLOOM, B. S., y MASIA, B. B. (1973): «Taxonomía de los Objetivos de la Educación. Ambito de la Creatividad». Ed. Marfil.
- MAKING STATISTICS MORE EFFECTIVE IN SCHOOLS OF BUSINESS (1988): «Proceedings of the Third Annual Conference». University of Wisconsin-Madison.
- MAKING STATISTICS MORE EFFECTIVE IN SCHOOLS OF BUSINESS (1990): «Proceedings of the Fifth Annual Conference». The University of Kansas.
- MAKUCH, W. M., y otros (1990): «A Statistical Computing Curriculum to Meet Industrial Needs». *The American Statistician*, 44, 1, pp. 42-48.
- MOORE, T. L., y ROBERTS, R. A. (1989): «Statistics at Liberal Arts Colleges». *The American Statistician*, 43, pp. 80-85.
- MORGAN, J. P., *et al.* (1991): «Let's Make a Deal: The Player's Dilemma». *The American Statistician*, 45, 4, pp. 284-289.
- MOSTELLER, F. (1988): «Broadning the Scope of Statistics and Statistical Education». *The American Statistician*, 42, 2, pp. 93-99.
- PEÑA, D. (1991): «Estadística, Modelos y Métodos». Alianza Universidad Textos, n.ºs 109 y 110.
- PEÑA, D.; MOLINA, C., y CORDERO, M. (1987): «La enseñanza del diseño de Experimentos en Ingeniería: algunas experiencias prácticas». *Questiio*, 11, 1, pp. 121-130.
- PEÑA, D.; PRAT, A., y ROMERO, R. (1990): «La Enseñanza de la Estadística en las Escuelas Técnicas». *Estadística Española*, 32, pp. 147-200.
- RADKE-SHARPE, N. (1991): «Writting As a Component of Statistics Education». *The American Statistician*, 45, 4, pp. 292-293.
- SIMON, J. L., y BRUCE, P. (1991): «Resampling: A Tool for Every day Statistical Work». *Chance*, 4, 1, pp. 22-32.
- SINGER, J. D., y WILLET, J. B. (1990): «Improving the Teaching of Applied Statistics:

- Putting the Data Back into Data Analysis». *The American Statistician*, 44, 3, pp. 223-230.
- TANNER, M. A. (1985): «The use of Investigations in the Introductory Statistics Course». *The American Statistician*, 39, 4, pp. 306-310.
- VANDERMAN, S., y DAVID, H. T. (1984): «Statistics for Quality and Productivity: A new graduate-level Statistics course». *The American Statistician*, 38, 4, pp. 235-243.
- WATTS, D. G. (1991): «Why is Introductory Statistics Difficult to learn? And What Can We Do to Make it Easier?». *The American Statistician*, 45, 4, pp. 290-291.

## ON TEACHING EXPERIMENTAL STATISTICS

### SUMMARY

This article presents some thoughts about how to organize an introductory statistics course for Economics, Business or Engineering students. The need for an active way of teaching, experiments carried out by the students, and continuous monitoring of the classroom learning process is stressed. It is also argued that statistics should be considered as an experimental course with heavy computer support. The design of the course should be the result of a clear set of pedagogical objectives. Examples of these ideas are presented.

*Key words:* Introductory Statistics. Computers. Probability. Montecarlo Methods.

*AMS Classification:* 6001, 60A05, 6201.