

# Estadística y Sociología

por  
**MANUEL GARCIA FERRANDO**  
Universidad Complutense de Madrid

## 1. INTRODUCCION

La enseñanza de la estadística en los departamentos y Facultades de Sociología ha presentado, desde sus inicios, no pocas dificultades, derivadas en gran medida de la escasez de conocimientos matemáticos que suelen exhibir los estudiantes que intentan obtener un título universitario en esta disciplina científico social. Algo similar ocurre con los estudios de Antropología y de Ciencia Política (1), ya que no pocos de los estudiantes que se matriculan en programas de estas materias científico-sociales, lo hacen pensando que se trata de conocimientos más próximos a las Humanidades que a las Ciencias Físico-Naturales, por lo que sus expectativas de enfrentarse con el estudio serio de la Estadística suelen ser muy bajas.

Esta situación ha hecho que tradicionalmente, los estudiantes de sociología, por su escasa formación matemática, no hayan podido aprovecharse plenamente de los enormes avances que se han producido en los últimos años en el campo de la estadística matemática y de la econometría.

---

(1) Los estudios de Sociología, Ciencia Política y Antropología Social se cursan en España en las Facultades de Ciencias Políticas y Sociología, con programas de estudios que tienen una parte común para las tres licenciaturas y otra parte específica para cada una de ellas. La Estadística y las Técnicas de investigación social, son materias comunes a los tres tipos de estudios.

Sin embargo, una situación tal no describe con rigor lo que está ocurriendo en la actualidad, ya que la difusión de los ordenadores personales y, en general, las facilidades de cálculo a que ha dado lugar la popularización de los paquetes estadísticos estandarizados tales como el SPSS, han conducido a que el interés por la estadística haya crecido considerablemente entre los estudiantes de Sociología y entre los propios sociólogos profesionales y académicos.

A este cambio ha contribuido poderosamente un nutrido grupo de profesores de Sociología, que con una buena preparación en Estadística y en Informática, han iniciado la enseñanza de la Estadística desde comienzos de la presente década de los 80 en las Facultades de Ciencias Políticas y Sociología españolas —en los momentos presentes ya existen cinco de tales Facultades en las Universidades españolas y en los próximos años es muy probable que abran sus puertas otras tantas—, teniendo en cuenta dos principios que entiendo son básicos para integrar felizmente la Estadística en el plan de estudios de estas Facultades. En primer lugar, los sociólogos implicados en la enseñanza de la Estadística, estamos teniendo en cuenta que los estudiantes de Sociología necesitan dicha disciplina para poder aprender a realizar investigación social. En consecuencia, el acento de lo que se enseña se pone más en el campo de la “estadística real (aplicada)” que en el de la “estadística matemática”. Más adelante volveremos sobre este punto.

En segundo lugar, los sociólogos estamos tratando de enseñar Estadística a los estudiantes de Sociología teniendo en cuenta que la informática existe. Lo que equivale a tener en cuenta que los ordenadores, incluso los llamados personales, realizan en breves segundos operaciones complejas que antaño requerían un enorme esfuerzo personal de cálculo, incluso utilizando una “calculadora”. Consiguientemente, tratamos de que los futuros sociólogos desarrollen un conocimiento amplio de los conceptos básicos de la Estadística y de sus técnicas fundamentales, de tal forma que conozcan la técnica estadística más adecuada a utilizar según el tipo de problema que les plantee los datos de que disponen, en el bien entendido de que es más que probable que el desarrollo de las operaciones concretas de cálculo lo vayan a realizar con la ayuda de un programa informático estandarizado. Por eso, más que tratar de que desarrollen una buena habilidad para resolver problemas de tipo estadístico matemático, tratamos de que los estudiantes de Sociología adquieran una buena comprensión de las ideas fundamentales que subyacen en los principios de la estadística inferencial, y no un conocimiento de “recetas estadísticas” desvinculadas de los problemas que plantea la investigación social.

## 2. LA ASIGNATURA DE ESTADÍSTICA EN LOS PROGRAMAS DE ESTUDIO DE SOCIOLOGÍA

La "Estadística aplicada a las Ciencias Sociales" es el nombre de la asignatura que se incluye en el primer curso de los actuales programas de estudio de las licenciaturas de Sociología y de Ciencia Política. Con este nombre ya no aparece más en el resto de ambos programas de licenciatura, aunque se continúa estudiando Estadística en la asignatura de Técnicas de Investigación Social, de segundo curso y, más adelante, en cuarto curso, en una asignatura optativa denominada Técnicas Avanzadas de Investigación Social.

En el programa de doctorado de la Facultad de C.C. Políticas y Sociología de la Universidad Complutense titulado "Técnicas y Métodos de Investigación", se integran dos cursos más de estadística, "Metodología multivariante de datos nominales" y "Los orígenes de la Estadística Social como Ciencia". Así pues, se observa que en el actual plan de estudios de Sociología y Ciencia Política, la Estadística aparece como asignatura troncal en primer curso, y como parte integrante de sendas materias de técnicas de investigación social en segundo y en cuarto curso. En los estudios de tercer ciclo, la Estadística es también protagonista en otras dos asignaturas optativas. (2)

Parece pertinente, pues, que tratemos con detenimiento el contenido del curso básico de Estadística aplicada a las Ciencias Sociales, ya que en el resto de las materias que se ocupan de la Estadística, se tratan en profundidad temas que están incluidos en el curso básico.

La enseñanza de la Estadística en muchos departamentos de sociología americanos y europeos, es deudora de las aportaciones del sociólogo norteamericano Hubert M. Blalock, que con la publicación en 1960 del texto **Social Statistics**, estableció un modelo de "programa de la asignatura", que con ligeras variaciones, hemos continuado explicando muchos sociólogos desde la década de los 60.

En realidad, y como he señalado con anterioridad, la innovación más importante que se ha introducido en la década de los 80 en la enseñanza de la Estadística a los estudiantes de Sociología, es el manejo de los paquetes estadísticos para ordenador, pero en el diseño de la asignatura, continúa siendo actual el modelo establecido hace casi tres décadas por Blalock.

---

(2) En las propuestas de contenido de las nuevas licenciaturas de Sociología y Ciencia Política, la Estadística aplicada a las Ciencias Sociales continúa considerándose una asignatura troncal.

A pesar de que muchos estadísticos matemáticos no son partidarios de distinguir dos ramas en el árbol común de la Estadística, en Sociología se distingue con toda nitidez la **Estadística Descriptiva** de la **Estadística Inferencial o Inductiva**. En esto, no se hace otra cosa que seguir las funciones generales que según Kruskal (1978) realizan los métodos estadísticos, a saber: 1) resumen de los datos y extracción de información relevante de los mismos, esto es, de las mediciones observadas, sean éstas numéricas, clasificatorias, ordinales, o de otro tipo; 2) y evaluación de los modelos y pautas que ofrecen los datos, pero que se encuentran ocultos por la inherente variabilidad aleatoria de los mismos; y 3) contribuir al diseño eficiente de experimentos y encuestas. La cuarta función que señala Kruskal, la de facilitar la comunicación entre los científicos, no tiene una influencia directa en el diseño de los programas de Estadística aplicada a las Ciencias Sociales, en la medida en que es inherente a la propia enseñanza universitaria el establecimiento de canales de comunicación científica.

Pues bien, la función de "resumen de los datos", se trata de satisfacer mediante la enseñanza de los conceptos y técnicas básicas de la Estadística Descriptiva, esto es, de las técnicas de organización de datos y distribución de frecuencias. En relación a la distribución de frecuencias se estudian los aspectos que son habituales en su tratamiento, esto es, representación gráfica, medidas de tendencia central, variación o dispersión y forma de la distribución.

Dado que en Sociología abundan más los datos medidos al nivel nominal y ordinal, que los que alcanzan el nivel de intervalo o de ratio, se pone especial atención en la enseñanza de la Estadística en los problemas que plantea la medición de variables, y en la necesidad de tener en cuenta el nivel de medición a la hora de utilizar un tipo u otro de estadísticos. Quizá en este aspecto debemos ser los sociólogos muy cuidadosos, al igual que les ocurre a los psicólogos, ya que a diferencia de los economistas, los datos empíricos que solemos disponer son con frecuencia bastante débiles desde el punto de vista de su nivel de medición, fundamentalmente los datos obtenidos a partir de encuestas.

Precisamente el uso que hacen los sociólogos de las encuestas y la necesidad creciente de utilizar modelos evaluadores, justifica la importancia que se concede en el programa de estudios de Estadística al estudio de la lógica y procedimientos para la inferencia o inducción. Habida cuenta de la escasez de conocimientos matemáticos y estadísticos con que llegan muchos estudiantes a las Facultades de Ciencias Políticas y Sociología, se hace necesario introducir en el programa de estudios de Estadística nociones básicas de probabilidad y muestreo.

En el estudio de las distribuciones de probabilidad se pone especial atención en la distribución normal por el papel central que juega esta distribución en la estadística aplicada. Y por lo que se refiere al muestreo, se introduce el estudio de los procedimientos aleatorios y no aleatorios, teniendo en cuenta que en el curso de Técnicas de Investigación Social se trata el muestreo aplicado a la investigación social con mayor detenimiento. Por eso, sólo se incluyen en el programa introductorio de Estadística las nociones fundamentales de muestreo.

Un capítulo muy importante en el programa que venimos comentando es el referente al estudio de las pruebas de decisión estadística y estimación de parámetros en la investigación social. Especial atención se dedica en este apartado a la formulación de las hipótesis estadísticas, a la elección de una prueba estadística, al nivel de significación y al tamaño de la muestra, y de manera especial a la distribución muestral, con la consiguiente inclusión del teorema del límite central y de la ley de los grandes números.

Dentro de este apartado dedicado a la Estadística inferencial, se estudian las pruebas de hipótesis en el caso de una muestra, siempre en función de los distintos niveles de medición. Así se estudia la distribución binomial y la correspondiente prueba binomial, la prueba de chi-cuadrado para una sola muestra y las distribuciones muestrales de las medias, incluyendo la distribución *t* de Student.

Nótese que en repetidas ocasiones hemos señalado la diferencia entre Estadística matemática y Estadística aplicada, al tiempo que mostrábamos que en el plan de estudios de Sociología y Ciencia Política, la Estadística que se enseña es claramente del último tipo. Existe ya una cierta tradición académica en el campo de la Estadística, que insiste en destacar la naturaleza dual de esta disciplina en las Ciencias Sociales. Y es que el frío y abstracto universo euclidiano que estudia la Estadística Matemática, no se adapta siempre muy bien a los dominios de las Ciencias Sociales, que no es otro que el mundo de las relaciones sociales y de la conducta humana.

Pienso al escribir estas notas en las reflexiones que a este respecto nos han brindado estadísticos tales como Woytinsky (1954) y Mosiman (1957), cuando señalan que la Estadística, en tanto que rama de las Matemáticas que por definición trabaja con conceptos abstractos, tiene que adaptarse a los datos reales del mundo social, con todas sus ambigüedades e imprecisiones. Por eso, según estos autores, la Estadística real (aplicada) es una mezcla de arte y de ciencia, ya que en lugar de la ausencia de error que es propia de la Estadística matemática, la Estadística real requiere tomar medidas sujetas a error, a la vez que realiza formulaciones en términos cuantitativos sobre entidades particulares y ya se sabe que

al medir cosas reales nunca se puede alcanzar la exactitud perfecta, pues entre otras cosas, sólo por el hecho de medirlas, las cosas reales cambian.

Consecuentemente con todo ello, los sociólogos tratamos de que los estudiantes de Sociología, al enfrentarse con la asignatura de Estadística, tomen rápidamente conciencia de la diferencia entre Estadística matemática y Estadística aplicada, siendo esta última una rama de la primera. Una toma tal de conciencia implica que se reconozca claramente lo que se está haciendo, por ejemplo, que ambas disciplinas tienen terminologías diferentes.

Siguiendo las observaciones que al respecto hace Mosiman, tratamos de que los estudiantes de Sociología, cuando estudian muestreo, perciban que mientras la Estadística Matemática puede considerar una muestra de tamaño " $n$ " extraída de una población de " $N$ " elementos con la limitación  $0 \leq n \leq N$ , en donde  $n$  puede ser efectivamente 0 o  $N$ , en la Estadística aplicada una muestra de cero elementos significa que no existe muestra y que la expresión una muestra del 100 por ciento no significa, propiamente, una muestra.

O por poner otro ejemplo característico, cuando explicamos la teoría de la probabilidad a los estudiantes de Sociología, procuramos que diferencien con toda claridad que en Matemáticas, la probabilidad es una precisa fracción matemática en la que no existe nada de "probable", mientras que en la Estadística real o aplicada, con el término probabilidad se denota algo que no es necesariamente cierto.

Si el sociólogo que enseña Estadística a los estudiantes de Sociología va teniendo éxito en la explicación del programa de la asignatura, cuando ha terminado de explicar los capítulos correspondientes a la Estadística univariante y a los primeros capítulos de la Estadística inferencial, es de esperar que los alumnos estén en condiciones de asimilar los capítulos correspondientes a la Estadística bivariable y multivariable, y que hayan adquirido los suficientes conocimientos matemáticos como para distinguir sin dificultad la naturaleza diferenciada de las pruebas paramétricas y de las no paramétricas.

De este modo podrán asimilar mejor los estudiantes las nociones fundamentales de la Estadística bivariable, y la estrecha relación que guardan las distribuciones bivariables con las etapas iniciales de la investigación social. Aunque desde un punto de vista estrictamente estadístico-matemático, el cálculo de porcentajes en una tabla bivariable no presenta mayores dificultades, su correcta interpretación desde el punto de vista sociológico no es tan sencilla, por lo que se presta bastante atención en el programa de la

asignatura a la presentación y análisis de las tablas bivariantes, y al cálculo e interpretación de los porcentajes en una tabla.

El estudio de las características de una asociación de dos variables sirve de marco en el que ocuparse de las medidas de asociación, de los conceptos de independencia estadística y de asociación perfecta. La diferenciación entre medidas simétricas y asimétricas de asociación se hace conjuntamente con la distinción entre medidas de asociación para variables nominales (el coeficiente  $\lambda$ , el coeficiente  $Tau\text{-}y$  de Goodman y Kruskal), y medidas de asociación para variables ordinales (coeficiente  $Tau\text{-}a$  de Kendall,  $\Gamma$  de Goodman y Kruskal,  $d$  de Sommers,  $Tau\text{-}b$  de Kendall,  $\rho$  de Spearman).

Nótese que el estudio de estas medidas de asociación se hace primeramente desde un punto de vista descriptivo y como paso previo al estudio inferencial de las pruebas de decisión para el caso de dos muestras, en donde se utilizarán algunos de tales coeficientes para probar la hipótesis de asociación. En este apartado dedicado a las pruebas de decisión para el caso de dos muestras, se estudian también las pruebas de la diferencia entre dos medias y entre dos proporciones, la prueba de  $\chi^2$  y el coeficiente de contingencia  $C$ .

El tratamiento de las medidas de asociación para variables de intervalo, conduce al estudio de la regresión y correlación. La escasa preparación matemática de muchos de los estudiantes que estudian el primer curso de Sociología, aconseja una vez más introducir la noción de regresión estudiando en primer lugar las propiedades de las ecuaciones de una recta, como modelización matemática de la relación entre dos variables estadísticas. De este modo, se facilita la comprensión por parte de los estudiantes de la ecuación de regresión y el ajuste por mínimos cuadrados.

También en el caso de la regresión y de la correlación, se diferencia con claridad la parte descriptiva de la parte inferencial, de tal modo que el estudiante adquiera una buena comprensión, de cara a la futura realización de investigación social, de la aplicación del coeficiente de regresión y del coeficiente  $r$  de correlación de Pearson para la prueba de hipótesis o test de asociación de dos variables de una población medidas en la correspondiente muestra.

Con el análisis de la varianza tanto para variables paramétricas como no paramétricas, y de la prueba  $F$ , se completa la parte del programa de la asignatura de Estadística aplicada a las Ciencias Sociales, dedicada a la Estadística univariante y bivalente.

La enseñanza de los anteriores temas "teóricos" se complementa en clase con la realización de ejercicios prácticos, de tal manera que el estu-

diante de Sociología va relacionando en todo momento a lo largo del curso la adquisición de los conceptos estadísticos fundamentales, con la solución de problemas estadísticos en base a variables sociológicas con distintos niveles de medición. En la Facultad de Ciencias Políticas y Sociología de la Universidad Complutense de Madrid, el libro de Manuel García Ferrando titulado **Socioestadística** (1988) y el de M. José Mateo titulado **Estadística en Investigación social** (1989) cubren de forma aproximada la parte de programa de la asignatura que se ha señalado con anterioridad, en su dimensión teórica y práctica, respectivamente, aunque naturalmente se utilizan otros textos como el referido de Blalock (1979), para ampliar la enseñanza de la asignatura, así como los apuntes de clase que cada profesor considera oportuno explicar.

Llegados a este punto en la explicación del programa, el curso suele estar bastante avanzado y entonces surge la duda de cómo explicar la última parte del programa, que está dedicada al análisis estadístico multivariable, en el escaso tiempo que queda hasta la celebración de los exámenes finales. La forma más usual de proceder consiste en explicar de forma resumida los problemas estadísticos que plantea la introducción de una o más variables en la relación básica entre dos variables, y el papel del ordenador en la solución matemática de los complejos cálculos que es preciso realizar en el análisis estadístico multivariable, dejando para cursos posteriores el estudio detallado de cada una de las técnicas estadísticas multivariadas más utilizadas en la investigación sociológica. Las Unidades Didácticas de Estadística Aplicada a las Ciencias Sociales correspondientes al primer curso de la licenciatura de Sociología y de Ciencias Políticas de la UNED, responde también a este modelo de asignatura (M. José Mateo y Manuel García Ferrando, 1989).

De hecho, tal estudio lo realiza en parte el estudiante de Sociología en la asignatura de Técnicas de Investigación Social que se da en segundo curso, y de forma optativa en cuarto curso en la asignatura de Técnicas Avanzadas de Investigación Social. El texto de J. J. Sánchez Carrión titulado **Introducción a las técnicas de análisis multivariable aplicadas a las ciencias sociales** (1984), se suele utilizar a este respecto sobre todo en el curso más avanzado.

Ahora bien, con independencia de que el mayor o menor número efectivo de días lectivos permita explicar con diverso grado de detalle a los estudiantes de primer curso algunas técnicas de análisis multivariable, se procura que al menos quede bien claro para los futuros sociólogos que el carácter multidimensional de buena parte de los fenómenos sociales exige análisis multivariados, y que tales análisis, por su complejidad matemática, eran hasta hace pocos años tan sólo asequibles a los estudiantes bien

preparados en Matemáticas. Pero en la actualidad, gracias a los ordenadores de alta velocidad y a los paquetes estadísticos estandarizados, se han convertido en herramientas analíticas de fácil manejo por todo aquel estudiante capaz de comprender las nociones fundamentales de un curso introductorio de análisis estadístico y con unas nociones poco más que elementales de informática. Con tan corto bagaje, son muchos los estudiantes que con independencia de la solidez de su formación matemática previa, pueden enfrentarse al manejo de un programa estadístico estándar con un ordenador personal.

Los sociólogos actuales, inmersos en un mar de información cuantitativa, encuentran dificultades muy grandes en ordenar conceptual y teóricamente la disciplina. Sin embargo, es mi opinión que los nuevos desarrollos de la Estadística, como es el caso del análisis multivariable, una vez se generalice su uso entre los sociólogos, van a facilitar enormemente la necesaria tarea de síntesis que en la actualidad requiere la Sociología. Veamos esto con mayor detalle.

La comprensión de una realidad social compleja y multidimensional obliga al investigador social a introducir en sus análisis, un número creciente de variables que reflejen la multitud de factores que intervienen en la producción de los fenómenos sociales. La elaboración de la relación entre dos variables fue la respuesta dada por Lazarsfeld (1955) a los problemas que surgen al introducir un factor de prueba en una relación bivariable. Pero el método de análisis lazarsfeldiano se hace inviable cuando se consideran simultáneamente los efectos de tres a cuatro variables de control (\*).

Si los científicos sociales no han hecho con anterioridad un uso mayor del tratamiento analítico con cuatro o cinco variables simultáneamente se ha debido, precisamente, a las dificultades de cálculo que se originan con ello. Por eso, el tipo de técnicas estadísticas que eran de uso corriente para el sociólogo no pasaban del tratamiento bi o trivariable de los fenómenos sociales. Sin embargo, la difusión del uso del ordenador en los últimos años, ha permitido de pronto la desaparición en la práctica de las dificultades de cálculo, ya que la gran capacidad operacional de los modernos ordenadores permite resolver, en cuestión de breves minutos, problemas estadísticos cuyo cálculo manual exigiría el esfuerzo de largas horas a equipos completos de investigadores.

---

(\*) En el periodo transcurrido entre la redacción del presente artículo y la corrección de las pruebas de imprenta, ha sido publicado el libro de J. J. Sánchez Carrión titulado *Análisis de tablas de Contingencia* (Madrid, CIS, 1989), que se ocupa monográficamente del análisis de datos sociológicos siguiendo "el paradigma de Evaluación de Lazarsfeld".

En tal sentido, la situación con la que se encontrará el científico en la actualidad, desde el punto de vista del tratamiento estadístico de la información empírica, es auténticamente revolucionaria en relación a la situación que disfrutaban los sociólogos de generaciones anteriores, que sólo disponían de una tecnología elemental para almacenar datos y resolver sus problemas de cálculo.

El uso y difusión del ordenador ha revolucionado, pues, el panorama informativo, estadístico y de cálculo de las ciencias en general, y de las ciencias sociales en particular, haciendo posible y necesaria la utilización de métodos de análisis multivariantes y la elaboración y contraste de modelos matemáticos para tratar de explicar la realidad social.

No existe apenas dificultad entre los sociólogos interesados por cuestiones metodológicas, en aceptar que el factor más importante en la rápida difusión de las técnicas multivariantes en la investigación sociológica, es la amplia disponibilidad de programas de ordenador. De hecho, la difusión de tales programas ha permitido superar, en cierta medida, el desequilibrio que existía por un lado entre los nutridos bancos de datos sociológicos que es posible encontrar en muchas universidades e institutos oficiales y privados de investigación social, y por otro lado, el pobre análisis estadístico que habitualmente se ha venido haciendo a partir de tales datos.

En España, antes de que se difundiesen los programas de ordenador para el análisis multivariante, se habían realizado ya numerosas y amplias encuestas sociales que posteriormente, eran tratadas simplemente a niveles bivariable o todo lo más trivariante. La difusión en la década de los setenta de tales programas —como, por ejemplo, el Statistical Package for the Social Sciences (denominado abreviadamente SPSS) o el Biomedical Computer Programs (BMD)— ha permitido una paulatina incorporación de los investigadores sociales españoles a la utilización de tales técnicas, que todavía son, con frecuencia, poco comprendidas en sus fundamentos lógicos.

Pero, ¿por qué la revolución de los ordenadores se ha traducido, en el campo de la Sociología, en la difusión de las técnicas de análisis multivariante?. Sencillamente, porque la tecnología del ordenador aplicada a las ciencias sociales apenas permite otro uso, a diferencia de lo que ocurre en el campo de las ciencias naturales. La difusión de los ordenadores se ha producido en una época en que la mayor parte de las ciencias naturales poseían ya un número sustancioso de leyes bien desarrolladas, basadas en un razonamiento deductivo. En tal situación, el ordenador ofrece unas excelentes oportunidades para la creación de modelos y la programación

de una compleja red de operaciones que permiten la comprobación a gran escala de las leyes físicas. El éxito obtenido en la exploración del espacio exterior, por medio de sofisticados cohetes y satélites artificiales, es la mejor prueba de la aplicación con éxito de la tecnología del ordenador a la física aplicada y de la resolución de los complejos problemas teóricos que ello comporta.

Sin embargo, la situación en las ciencias sociales, incluida la Sociología, está muy lejos de asemejarse siquiera a la que ofrecen las ciencias naturales. La ausencia de leyes sociales invariables es notoria, y buena parte de la investigación social continúa siendo de naturaleza empírica, exploratoria o inductiva. En consecuencia, el ordenador se ha utilizado primordialmente, en el campo de la Sociología, para analizar, seleccionar, almacenar, clasificar y procesar, en general, datos sobre actitudes, opiniones, valores y comportamientos que provienen, fundamentalmente, de las encuestas sociológicas. Dada la gran capacidad del ordenador para procesar tal tipo de datos, y partiendo de la reconocida complejidad multidimensional de la realidad social, lo más natural ha sido utilizarlo para investigar apropiadamente, es decir, multidimensionalmente, dicha realidad.

La rápida difusión y aceptación del análisis multivariable en el campo de la sociología se refleja en los siguientes hechos:

1.- Se ha desarrollado y se ha puesto a la disposición del gran público un número creciente de programas de ordenador para tales técnicas, tales como los ya citados, SPSS o el BMD. 2.- Prácticamente todos los artículos de naturaleza empírica que se publican en las revistas de sociología profesionales más prestigiosas, utilizan casi exclusivamente técnicas estadísticas multivariadas, dejando ya como obsoletos los tratamientos meramente porcentuales de la información. 3.- Los libros y artículos metodológicos recientes, de naturaleza empírica, que se ocupan de problemas sociológicos contienen casi exclusivamente tratamientos multivariados.

La reciente publicación del libro de J. J. Sánchez Carrión titulado **Introducción al análisis de datos con SPSS/PC +** (1988), es un buen ejemplo de la preocupación pedagógica de los sociólogos por hacer accesible a los estudiantes de ciencias humanas y sociales el análisis multivariable de datos, que tal como he señalado anteriormente, es el que en realidad exige el análisis de una realidad social casi siempre multidimensional. El carácter novedoso de la "revolución multivariable" en el análisis de datos sociales, bien merece que nos detengamos en su reflexión, y en los mismos términos en los que tratamos de explicar su contenido a los estudiantes de Sociología.

### 3. El análisis multivariable en la investigación sociológica

Aceptada la importancia del análisis multivariable en Sociología, surge el problema de cómo transmitir su conocimiento a unos estudiantes que no han tenido anteriormente oportunidad de estudiar cálculo matricial, y que tampoco van a poder hacerlo con el programa oficial de estudios de las Licenciaturas de Ciencias Políticas y Sociología.

Llegados a este punto, caben dos opciones para los estudiantes. La primera de ellas es la de ampliar los conocimientos estadísticos "fuera del programa", con la ayuda de algún profesor del departamento de Métodos y Técnicas de Investigación Social, y elegir posteriormente en cuarto curso la asignatura de Técnicas Avanzadas de Investigación Social. Se trata, evidentemente, de una opción minoritaria.

La segunda opción es la de adquirir unas nociones algebraicas elementales sobre cálculo matricial, familiarizándose el estudiante con nociones tales como matriz general de datos, matriz de varianza y covarianza, matriz de correlaciones, inversión de matrices, vectores y valores propios, componentes principales y unos pocos más. Con este bagaje, se puede estar en condiciones de adentrarse en el estudio de las técnicas multivariadas, que en mi libro anteriormente citado **Sociestadística (1988)**, en el que se dedica un capítulo a la explicación de las anteriores nociones elementales, se presentan diferenciando entre técnicas basadas en la dependencia entre las variables —regresión múltiple, análisis de camino (path analysis), análisis canónico, análisis discriminante múltiple y análisis multivariable de la varianza—, y técnicas basadas en la interdependencia de las variables —análisis factorial, análisis de conglomerados (cluster analysis), análisis de segmentación (tree analysis) y técnicas de escalamiento multidimensional—.

Naturalmente, el nivel de comprensión que puede adquirir el estudiante medio de Sociología de las técnicas de análisis multivariable anteriormente señaladas, no puede ser más que introductorio, justo el necesario para poder seguir la lectura de los trabajos de sociología empíricos que empleen tales técnicas, o incluso para poder aprender a realizar por sí mismos investigación social empírica, aunque en caso de tratar problemas complejos contando siempre con la elaboración de otro investigador más especializado en técnicas estadísticas en informática.

En tanto los programas oficiales de estudios de las licenciaturas de Sociología y de Ciencia Política, sólo incluyan un curso de Estadística aplicada a las Ciencias Sociales, la enseñanza del análisis multivariable presentará serios problemas, tal como se acaba de señalar. Confiemos en que en el nuevo programa de estudios que actualmente se está preparando

do, se incluyan dos cursos de Estadística, de tal manera que en el primero de ellos se estudie la Estadística univariable y bivariable, mientras que en el segundo se estudien con mayor profundidad las técnicas de análisis multivariable.

Pero si sólo se cuenta oficialmente con un curso de Estadística, estaremos obligados a "comprimir" la asignatura, confiando en que el estudiante vaya asimilando realmente los conceptos estadísticos "aprendidos", a medida que va ampliando a lo largo de la carrera sus conocimientos de investigación social y sociología empírica. En cualquier caso, el estudio de las técnicas de análisis multivariable en conjunción con el aprendizaje del manejo de al menos un paquete estadístico estándar para ordenador, va a ser cada vez más exigible a los futuros sociólogos que aspiren a ejercer su profesión en un contexto ocupacional que ha hecho del tratamiento multidimensional de la información social, su rasgo más definitorio y característico.

## BIBLIOGRAFÍA

En la presente bibliografía sólo se reseñan los autores citados en el texto de este artículo. En las bibliografías que incluyen los libros citados de Sánchez Carrión (1984) y García Ferrando (1988), se pueden encontrar buena parte de los trabajos más relevantes de Estadística y Sociología que se suelen utilizar en nuestros contextos académicos y profesionales.

BLALOCK, HUBERT M., *Social Statistics*, Nueva York, McGraw-Hill, 1960.

CATTELL, R. B. (ed.), *Handbook of Multivariate Experimental Psychology*, Chicago, Rand McNally & Co., 1966.

GARCÍA FERRANDO, MANUEL, *Socioestadística. Introducción a la Estadística en Sociología*, Madrid, Alianza Editorial, 1988.

KENDALL, MAURICE G., *Multivariate Analysis*, Londres, Charles Griffin and Co. 1975.

KRUSKAL, WILLIAM H., "The field of Statistics", en W. H. Kruskal y J. M. Tanur (eds.), *International Encyclopedia of Statistics*, Nueva York, Free Press, 1978, págs. 1701-1091.

LAZARSELD, PAUL F., "Interpretation of Statistical Relations as a Research Operation", en Paul F. Lazarsfeld, y Morris Rosenberg, *The Language of Social Research*, Nueva York, Free Press, 1955, págs. 115-125.

MATEO, M. JOSE, *Estadística en Investigación Social. Ejercicios Resueltos*, Madrid, Paraninfo, 1989.

- MATEO, M. JOSE y M. GARCIA FERRANDO, Estadística aplicada a las Ciencias Sociales, Madrid, Universidad Nacional de Educación a Distancia, 1989.*
- SANCHEZ CARRION, J. J. (compilador). Introducción a las técnicas de análisis multivariable aplicadas a las ciencias sociales, Madrid, Centro de Investigaciones Sociológicas, 1984.*
- SANCHEZ CARRION, J. J., Introducción al análisis de datos con SPSS7PC+, Madrid, Alianza Editorial, 1988.*
- SANCHEZ CARRION, J. J., Análisis de Tablas de Contingencia, Madrid, Centro de Investigaciones Sociológicas, 1989.*

## Comentarios a la enseñanza de la Estadística en Sociología

### COMENTARIO

FRANCISCO ALVIRA MARTIN

Universidad Complutense de Madrid

En líneas generales, el artículo de Manuel García Ferrando, describe adecuadamente la realidad existente como no podría ser de otra manera, no en vano el autor ha sido responsable de la disciplina "Estadística aplicada a las Ciencias Sociales" durante muchos años en la Facultad de Ciencias Políticas y Sociología de la Universidad Complutense. No obstante, me gustaría matizar aquí algunos de los temas recurrentes de dicho artículo, para completar así la perspectiva argumental seguida por su autor.

Es cierto que los estudiantes de Sociología —y los sociólogos— muestran —¿mostramos?— escasa preparación matemática. Pero no hay que achacar exclusivamente a esta falta de preparación y a la ausencia antaño de paquetes estadísticos y ordenadores personales, el escaso interés por la estadística dentro de los estudios de Sociología.

Cuando ya tardíamente se creó la primera Facultad de Ciencias Políticas y Sociología, estaba de moda la mal llamada perspectiva "cualitativa", contraria a la estadística, y a las también mal llamadas técnicas cuantitativas. Esta perspectiva cualitativa se equiparaba a términos positivos como progresismo, crítica y modernidad. Del mismo modo, la llamada perspectiva cuantitativa se igualaba a funcionalismo, americanismo y algo "déjà vu", términos con connotaciones negativas.

A duras penas se logró introducir en el plan de estudios la asignatura de "Estadística aplicada" y algunos más de técnicas de investigación en número claramente precario y mínimo. Durante muchos años, el Departamento de Métodos y Técnicas de Investigación que tiene a su cargo la docencia de estas asignaturas de Estadística y Análisis de Datos, quedó relegado a un plano secundario dentro de dicha Facultad.

La vena a-empírica y a-cuantitativa de la Sociología española de los años 70 y parte de los 80, tiene su reflejo no sólo en el "desprecio" por el análisis de datos, la estadística, la medición, etc..., sino también en:

- la escasa dotación de plazas de numerarios durante muchos años y,
- el "apropiamiento" de las disciplinas de análisis de datos por parte de departamentos sustantivos.

La actitud anti-empiricista, anti-cuantitativista y anti-estadística de la Sociología española es persistente y ha persistido, aunque esté en retroceso. Así, la mayoría de los artículos que aparecen en las revistas de la profesión en España, no utilizan la Estadística y, de hacerlo, lo hacen de un modo muy rudimentario.

Estoy afirmando justo lo contrario que señala Manuel García Ferrando en su artículo, pero claro está, él no se refiere a las publicaciones de la Sociología española, sino de la Sociología internacional donde, efectivamente, es cierto que reflejan cada vez más un mayor conocimiento y utilización de técnicas estadísticas avanzadas. Esto no sucede apenas en España (véase F. Alvira, 1989).

Afortunadamente, en los últimos años, el péndulo se ha inclinado, quizá en exceso, por lo cuantitativo, la estadística, el análisis de datos, etc..., pero no sólo porque el desarrollo de la informática lo permite y lo potencia. La razón esencial estriba en que las técnicas de investigación, en general, y la estadística, en particular, se presentan como herramienta imprescindible para el desempeño profesional del sociólogo. La estadística es necesaria no sólo para investigar, sino también —y de un modo creciente— para el desempeño del papel del sociólogo en la Administración Pública y en las empresas.

Es el mercado quien ha hecho imprescindible esta vuelta a lo cuantitativo. El sociólogo no puede, sin las técnicas estadísticas, competir con otros profesionales —psicólogos, economistas, etc...— y, no puede tampoco ejercer adecuadamente su profesión.

La consecuencia última de este proceso ha sido el cambio de actitud de los sociólogos y la potenciación creciente de todo lo "cuantitativo", incluyendo la Estadística.

Conviene también matizar la influencia que García Ferrando atribuye a H. M. Blalock en el modo como la Estadística se estructura y se ofrece en Sociología. No todos los sociólogos han "bebido" de Blalock reflejando luego esta fuente en sus enseñanzas. Yo mismo me introduje en la Estadística de la mano de dos bioestadísticos —W. J. Dixon y F. J. Massey— para continuar con dos economistas —Wannacott y J. Kmenta—, un psicólogo

—A. Trersky— y L. Kish. Otros sociólogos han bebido de otras fuentes, existiendo de hecho múltiples autores que han influenciado el tipo de Estadística aprendida por estos profesionales de las ciencias sociales.

Estos autores provienen tanto de la propia Sociología y de la Estadística como de disciplinas afines: la Economía y la Psicología.

La disparidad de influencias y formación tiene como resultado perspectivas también diferentes, aunque comunes en lo básico, sobre el papel que tiene que jugar la Estadística en la Sociología y la formación de los Sociólogos. En la actual estructura del plan de estudios apenas si tienen relevancia temas tan vitales como el muestreo y la medición, temas a los que debería darse un mayor papel.

Por ello, creo que el futuro de la Estadística en la Sociología debería de enfocarse de otra manera. Es ridícula la actual situación en la que la estadística enseñada en Psicología, Economía y Sociología, etc..., forma compartimentos estancos sin que exista fertilización mutua. Esta situación carece de sentido cuando, por ejemplo, el análisis de series temporales o de ecuaciones estructurales ya no es algo exclusivo de la Economía, ni el análisis factorial algo exclusivo de la Psicología. Las diferentes ciencias sociales y/o de la conducta están utilizando las mismas técnicas estadísticas, aunque haciendo hincapié en diferentes aspectos.

Gran parte de la Estadística o aplicaciones de la Estadística que se enseña en Psicología tiene cabida en Sociología y en Economía y viceversa. Por ello, creo que debería imponerse una economía de esfuerzo que se orientaría en una doble dirección:

— desarrollo de cursos comunes a las ciencias sociales de Matemáticas y Estadística teórica.

— cursos de técnicas estadísticas aplicadas desde la diferentes ciencias sociales, pero abiertas a todas las demás ciencias sociales.

Obviamente, este desarrollo es más fácil llevarlo a cabo en segundo y tercer ciclo, pero creo que, en cualquier caso, resultaría muy fructífero.

## REFERENCIA

- F. ALVIRA MARTIN. *"El desfase entre teoría e investigación en la Sociología"*. Ponencia presentada en el III Congreso de Sociología, S. Sebastián. Septiembre 1989.

## Sobre la utilización de la Estadística por los sociólogos

MIGUEL BELTRAN

Universidad Autónoma de Madrid

Nada tengo que comentar, y menos que añadir, al trabajo de mi colega y amigo, profesor García Ferrando, acerca de la enseñanza de la Estadística a los estudiantes de Sociología. Aprovecharé, sin embargo, la hospitalidad que esta prestigiosa Revista me ofrece para decir alguna cosa en relación con el uso que los sociólogos hacemos de la Estadística, uso que a veces me parece en exceso expeditivo y poco cuidadoso. En caso de que tengan algún valor mis deshilvanadas reflexiones, no será difícil deducir de ellas la conveniencia de introducir determinados énfasis en la enseñanza de la Estadística a quienes se forman como sociólogos, con lo que no me habré alejado en exceso de la preocupación didáctica que preside estas páginas.

Los sociólogos (o, más exactamente, muchos de ellos) utilizan habitualmente una amplia serie de técnicas estadísticas de análisis de datos, adecuadas no sólo para el método cuantitativo, sino también para el histórico y el comparativo (en la medida, claro está, en que se historifican o comparan magnitudes variables): es obvio que las ciencias sociales no han de manejarse necesariamente con métodos "puros", sino que con frecuencia utilizan conjuntamente las herramientas heurísticas que proporcionan dos o más de ellos. Resultaría ocioso recordar aquí la extrema importancia que para la historia económica y social ha tenido un movimiento como el de los **Annales**, al incorporar los datos y las técnicas cuantitativas a la descripción y explicación históricas (habiéndose incluso llegado al pintoresquismo de hablar de **cliometría**).

Convendrá recordar, de todas formas, que las ciencias sociales han de manejar el análisis estadístico con la necesaria sensibilidad crítica, ya que no se trata en modo alguno de "ciencias exactas": al no ser su modelo las ciencias de la naturaleza, el hecho de que puedan y deban cuantificar cuando así lo exige su objeto no implica en modo alguno un mimetismo naturalista que carecería de justificación. Desde luego, fenómenos como los demográficos, o las infinitas pautas recurrentes en toda suerte de ámbitos sociales, exigen el método cuantitativo y, consecuentemente, las técnicas de análisis estadístico. Pero recuérdese que con ello se aborda sólo una dimensión del objeto, que en su totalidad concreta presenta una mucho mayor complejidad.

En todo caso, la estadística es sin duda la más común y la más canónica de las técnicas de análisis. Desde luego, no es la única de naturaleza empírica: todas lo son, en la medida en que todas se dirigen a la **cosa** para añañarla, frotarla, desconstruirla y reconstruirla, para escudriñarla y hacerle

decir lo que **quiere decir** y por sí misma no puede. Es claro que el tipo de "frotamiento" (la técnica) ha de ser muy diferente según el método de acceso a la "cosa" (la dimensión del objeto) que se esté manejando, manejo que nunca será resultado de una opción caprichosa por parte del investigador, sino que vendrá exigido por la dimensión del objeto que se haya preferido (ahora sí) considerar. Pero —hay que repetirlo— tanto el método cuantitativo como el cualitativo, el histórico como el comparativo, e incluso el crítico-racional, son todos ellos métodos empíricos que han de habérselas con evidencias externas, con aspectos o dimensiones del objeto de conocimiento de las ciencias sociales. Pues bien, no creo que sea muy exagerado afirmar que el análisis estadístico es hoy en día la herramienta analítica más común en las ciencias sociales; puede incluso decirse que hay una cierta tendencia a su abuso: la tan denostada **cuantofrenia** descansa en el mismo mecanismo que el martillo en manos de un niño, que descubre enseguida que todo es martillable (aunque los adultos no puedan compartir tan destructiva convicción). Es claro que los martillos son imprescindibles, pero sólo para ciertas cosas: un clavo no se clava con un paraguas, pero uno no evita mojarse con la lluvia dándose martillazos en la cabeza.

He dicho más arriba que el análisis estadístico es la forma más canónica de análisis. En efecto, está cuidadosamente codificado, y sus reglas son claras y rigurosas; puede estudiarse en manuales bien establecidos, e incluso utilizarse con calculadoras y ordenadores que rutinizan las operaciones intermedias hasta llegar al resultado deseado. Todo el mundo puede valorar sin ambigüedad la utilización hecha por el investigador de las técnicas de que se trate, y replicarlas cuando sea necesario. Y, en caso de duda, siempre podrá recurrirse a una fuente que indique **cómo hacerlo**. Nada de esto se parece, por supuesto, al mundo de las técnicas de análisis cualitativo, en las que el análisis del discurso puede ser plausible y convincente o puede no serlo: si no se me toma al pie de la letra, me atrevería a decir que el análisis cualitativo tiene más de arte que de técnica (aunque estoy seguro de que tanto los psicoanalistas como los hermenéutas rechazarían semejante calificación).

No se crea, sin embargo, que ese carácter canónico del análisis estadístico permite una utilización aproblemática: no obstante su **manualidad** (y a veces precisamente por ella), el manejo de las técnicas estadísticas por muchos investigadores resulta inquietante. La cuantofrenia no es sino pérdida del sentido teórico o, si se prefiere, la primacía del método sobre el objeto. La generalizada disponibilidad de ordenadores personales y la extraordinaria riqueza y potencia de los programas estadísticos que pueden utilizarse en ellos para el procesamiento y análisis de datos, está llevando hoy día a muchos investigadores a un nivel de refinamiento estadístico que excede con mucho de las conclusiones a que son capaces de llegar, e

incluso a su comprensión del análisis que han llevado a cabo. Todo el mundo puede rotar los ejes de esta o de la otra manera en un programa de análisis factorial, pero son pocos los que saben qué sentido tiene lo que están haciendo; o, por seguir con el ejemplo, es bien fácil culminar el proceso llegando a la identificación de dos o tres factores que tienen el misterioso poder de sintetizar el **deus ex machina** del síndrome de variables manejadas: pero con frecuencia en eso se quedan la mayoría de los investigadores, que como mucho se esfuerzan en etiquetar dichos factores, concluyendo apresuradamente el trabajo cuando sería en ese punto donde debería propiamente comenzar.

Puede parecer paradójico que simultáneamente se afirme la necesidad de llevar a cabo el análisis estadístico en ordenador, y se ponga en guardia contra las miserias del proceso de datos: pero el riesgo de sacralizar los datos más deficientes y mediocres gracias al refinamiento analítico y al prestigio formal del ordenador es muy grande, y ninguno estamos inmunizados definitivamente al respecto. Por lo que se impone la inevitable receta de siempre: más teoría y más atención a la evidencia empírica (que en la práctica vienen a ser las dos caras de una misma moneda). Y, naturalmente, saber un poco de estadística. ¿Sólo un poco?

Un investigador social no tiene por qué ser un estadístico, del mismo modo que quien usa el ordenador para procesar datos o textos no tiene que ser programador: el investigador es **usuario final** de la estadística y, si se le permite la comparación, su actitud es análoga a la del conductor de un automóvil, que no tiene que ser un mecánico, ni siquiera **amateur** (aunque no le vendrá nada mal tener la suficiente idea de mecánica para saber con cierta precisión cómo funciona su coche). Ciertamente, el análisis estadístico es imprescindible para el investigador social cuantitativo, y tiene que manejarse en él con absoluta soltura; pero dicha soltura se concreta en una solvente comprensión de las técnicas que utiliza y de los **estadísticos** que calcula, no en el detallado dominio de la fundamentación matemática de unas y otros. Claro está que ha de haber "estadística para economistas" o "para sociólogos", que es cosa muy diferente de la "estadística para estadísticos": espero que a éstos no les disguste mi insistencia en el carácter instrumental o ancilar que su disciplina tiene para otras ciencias, en su papel de conjunto de técnicas cuantitativas de análisis de datos.

Una última observación, tan trivial —o tan importante— como las anteriores: no hay que confundir la medición de las variables que uno tiene entre manos con su análisis estadístico. Este pretende sintetizar una masa de información inmanejable, determinar la relación que existe entre dos o más variables, establecer el grado de probabilidad de que se produzca en la realidad una determinada situación, poner de manifiesto cosas que estaban

latentes en los datos, etc. La medición, pues, ofrece datos, que son trabajados en el análisis para llegar a las conclusiones pertinentes. Conclusiones que no se obtienen mecánicamente, sino que implican en la mayor parte de los casos decisiones por parte del investigador, que asume siempre su propia cuota de responsabilidad en el análisis.

Pero dejemos a un lado tales cuestiones, y vayamos a una algo más concreta: mucha gente identifica sociología con encuestas y, correlativamente, con tratamiento estadístico de datos de opinión. No voy a hacer aquí ningún esfuerzo por rebatir tan abusiva y vulgar identificación, sobre todo porque hay que reconocer que la técnica de encuesta es una de las más utilizadas, y desde luego es la que con más frecuencia salta al escapate de los medios de comunicación. Quizá convenga por ello hacer aquí alguna referencia al respecto, que sin duda debería incluirse en el aprendizaje de la estadística por parte de los estudiantes de sociología.

Lo primero que creo que hay que destacar es que las muestras con las que trabajamos los sociólogos rara vez son probabilísticas: es raro que utilicemos una muestra estrictamente aleatoria, pues el problema de la identificación y localización de las unidades muestrales lleva a que, las más de las veces, nuestras muestras sean —dejando aparte su diseño y pasos previos— “de rutas y cuotas”, con lo que, en términos rigurosos y estrictos, no puede decirse que permitan la inferencia estadística. No obstante, lo cierto es que tales muestras funcionan a la perfección, como pone de manifiesto a diario la enorme masa de estudios de opinión y de mercado que se lleva a cabo en todo el mundo. Careciendo en rigor del requisito de aleatoriedad (por más que se introduzcan en la selección de las unidades muestrales ciertas dosis convencionales de la misma en la manera de diseñar las rutas y de seleccionar a los entrevistados), podríamos decir que **eppur si muove**.

Otra cuestión que me parece importante, y a la que creo que no se presta la debida atención, es la del nivel de medición de la mayor parte de las variables con las que trabajamos habitualmente los sociólogos. Pocas veces (salvo en el análisis secundario) contamos con escalas de intervalo, y muchas con escalas nominales y ordinales. Consecuentemente, con frecuencia nos está vedada la utilización de la parte más potente de la panoplia estadística. Bien es verdad que se han elaborado abundantes e ingeniosas fórmulas para minimizar tal limitación, y que la propia estadística está prestando últimamente mucha atención al análisis nominal y ordinal. Pero me parece que no hacemos suficientemente conscientes a nuestros estudiantes de esta especificidad (o, si se prefiere, limitación) de muchos de nuestros datos y, por consiguiente, no hacemos el esfuerzo que tendríamos que hacer en prepararlos con la necesaria intensidad para el análisis de datos nominales y ordinales.

Creo que ha pasado ya la época en la que era necesario prevenirse contra una ingenua "sociología de dos decimales" (por su abuso de presentar los porcentajes de esa forma): felizmente tal pretensión de precisión, derivada de una suerte de complejo de inferioridad frente a las ciencias físico-naturales, ha dado paso a una actitud mucho más seria, en la que el rigor científico no se busca tanto en los decimales cuanto en la adecuada intelección de lo que se tiene entre manos. Es cierto que quedan algunos **tics** de aquel cientifismo ya obsoleto, y podría ser un ejemplo la reverencia con que manejamos el margen de error en las encuestas (olvidando con frecuencia que la famosa situación más desfavorable de  $p=q=50\%$  vale sólo para preguntas dicotómicas, con lo que el error será mayor en las preguntas con más de dos **items** de respuesta).

Pero nada más lejos de mi intención que tomar partido por lo que en otro lugar he llamado "análisis de pan y chocolate" a causa de su simplicidad, frente al refinamiento analítico que nos permite la extraordinaria potencia que hoy tienen las herramientas estadísticas, y particularmente el análisis multivariable. La realidad social es tan oscura, difícil y elusiva que sería inaceptable limitar deliberadamente nuestra capacidad para su indagación, para describirla y explicarla adecuadamente y, por tanto, para analizarla de la manera más penetrante y completa posible. Pero justamente porque ese poder analítico de las herramientas estadísticas actuales está disponible sin necesidad de grandes esfuerzos, es por lo que hay que esforzarse cada vez más en producir datos de la mejor calidad posible, y en manejarse con teorías e hipótesis muy meditadas. De otra forma, el riesgo, como he dicho, es el de sacralizar con el prestigio del más refinado análisis estadístico lo que puede no ser más que basura.

Desde tal punto de vista creo que ha de plantearse la enseñanza de la estadística para sociólogos: no creo equivocarme al afirmar que el objetivo ha de ser formar buenos sociólogos que sean buenos **usuarios finales** de las herramientas de análisis estadístico. Pero siempre, y ante todo, buenos sociólogos. Se dirá, y con razón, que para llegar a tan banal conclusión no se necesitaban muchos argumentos. Cierto.

## COMENTARIO

JULIO CARABAÑA

Universidad de Valencia

El artículo del profesor García Ferrando me parece una buena descripción de las relaciones actuales entre Estadística y Sociología y de los problemas que plantea la enseñanza de la Estadística en las Facultades de Sociología. El profesor García Ferrando tiene una experiencia docente en la materia de

la que yo carezco, es autor de varios textos de Estadística para sociólogos y creo que no me equivoco al pensar que ha alentado la publicación de otros, como los de Mateo Rivas y Sánchez Carrión. El Departamento al que perteneció durante años en la Universidad Complutense ha realizado significativos avances en la enseñanza de la materia, y ha conseguido el funcionamiento de un aula de informática cuyos quizá cuarenta ordenadores utilizan asiduamente los alumnos. Creo que es completamente acertado por su parte destacar las diferencias entre Estadística Matemática y Estadística Aplicada, el impacto revolucionario de los paquetes de programas standard utilizables en ordenadores personales y la necesidad de ampliar la enseñanza de la Estadística y las técnicas de investigación en el Plan de Estudios de las Facultades de Sociología.

Lo que mi comentario puede aportar a su artículo es el punto de vista de un estudiante de Letras con cierta experiencia en la evaluación de proyectos de investigación empírica. Estudié Bachillerato de Letras en la década de los sesenta, cuando las Matemáticas se abandonaban en cuarto curso. Estudié luego Filosofía en la Universidad. Cuando, acabada la carrera, me dediqué a la Sociología, seguí en la Universidad de Colonia dos cursos de Estadística que me resultaron abstractos y carentes de interés porque era incapaz de ponerlos en relación con los problemas sustantivos que me interesaban. Felizmente para mí, hacia 1974 todavía se podía integrar en un equipo de investigación empírica a gente que se confundía ante una matriz de correlaciones. Tuve que aprender entonces, en un contexto ya dominado por los ordenadores, a escribir programas en SPSS y en BMDP y a interpretar sus resultados. Tuve que repasar primero mis viejos textos de Estadística, luego engolfarme en manuales avanzados y en monografías sobre análisis factorial, regresión, análisis de tablas cruzadas, etc. Para entender un poco todo esto tuve que estudiar las matemáticas —derivadas, matrices— que no había estudiado en Bachillerato. Pese a ello los fundamentos matemáticos de la Estadística me siguen resultando opacos y soy incapaz de discutir una demostración o de intentar la más trivial innovación en la materia. Mi formación estadística es la de un autodidacta sin formación matemática con cierta facilidad para aprender directamente de los libros. Aún con este pobre bagaje, publiqué en 1983 un libro sobre movilidad social basado en regresiones múltiples que creo que resulta ilegible para el 98 % de mis colegas.

Desde 1984 hasta 1986 trabajé en la Unidad de Investigación del Ministerio de Educación (parte primero del INCIE y luego del CIDE). Entre mis obligaciones se contaba la evaluación de proyectos e informes de investigación, en su mayor parte hechos por profesores universitarios relacionados con los ICEs. También participé en la selección de becarios de investigación para el organismo, provenientes de las Facultades de Psicología,

Pedagogía y Sociología. Al mismo tiempo, doy clase de Sociología General en el Primer Curso de Sociología y Ciencias Políticas, y hago cursos de doctorado que suponen que los alumnos pueden leer artículos de revista en inglés. Estas actividades me han permitido apreciar un importante progreso en la extensión y dominio de las técnicas estadísticas tanto por los investigadores como por los licenciados, a la vez que detectar problemas importantes. Los breves comentarios que siguen se refieren a la dificultad y a la necesidad de superarlos.

1. Los estudiantes de Sociología, dice el profesor García Ferrando, necesitan de la Estadística para poder aprender a realizar investigación social (p. 2). He de reconocer que esta afirmación toca uno de los puntos que más me confunden. Por un lado, yo no fui capaz de enfrentarme seriamente a la Estadística mientras no tuve que hacer investigación social. Cuando sólo tenía que leerla, me saltaba sin más lo que consideraba detalles "técnicos", los daba por buenos y me limitaba a la introducción y a las conclusiones. He visto después que es la tendencia de la mayor parte de la gente, por lo que mi conclusión es que la **didáctica** de la Estadística necesita imperiosamente de la práctica para interesar a alumnos que en su mayor parte creen haber desertado definitivamente de las Matemáticas y tienden a tomarse la Estadística como un engorro. Pero, por otro lado, la mayor parte de los alumnos no se van a dedicar nunca a la investigación social, sino, a lo sumo, a interpretar y usar investigaciones y fuentes estadísticas elaboradas por otros. Los alumnos sólo aprenden Estadística si esta se les enseña como si **estuvieran haciendo** investigación, pero sospecho que es sumamente difícil crear esta impresión en primero, cuando desconocen casi totalmente lo que es la Sociología y su curiosidad va por otros caminos. Mi única contribución a este problema consiste en incluir estudios de tipo estadístico en mi curso de Sociología General, pero creo que es evidente que en primero faltan tanto la motivación como el conocimiento de base necesarios para que los alumnos puedan apreciar lo que la Estadística significa para la Sociología.

2. Los ordenadores y los paquetes de programas son en cierto sentido una bendición, pues evitan los cálculos y alivian su enseñanza. Por desgracia pueden ser también una peste en manos de usuarios que desconocen sus fundamentos y sus limitaciones. Cuando su uso está basado en la fé, no en el conocimiento, fomentan peligrosamente la tendencia a no pensar. Además, sociólogos y economistas tienden en general a ignorar algunas de las condiciones estrictas sobre medida y distribución de las variables de la Estadística Matemática. Esta ignorancia está muchas veces justificada, por ejemplo, cuando se usan medidas ordinales en vez de cardinales en las regresiones, o cuando se desprecian pequeñas desviaciones de la normalidad en las distribuciones. Pero otras veces lo están menos, como cuando

se usan variables dependientes dicotómicas en regresiones o se desprecian grandes desviaciones de la normalidad. El simple usuario tiende a ignorar **siempre** las restricciones, a meter sin más en el ordenador lo que tiene y a esperar la salida. Incluso si es consciente del fundamental principio "garbage in, garbage out", suele ser incapaz de identificar correctamente la basura.

Esta sustitución del pensamiento por el cálculo, del investigador por el ordenador y de la formulación de hipótesis por el cruce de todo por todo da lugar a una variada teratología. Puedo recordar algunos notables monstruos —serían muchos más con una mínima propensión mía al coleccionismo: un profesor titular de psicología que obtuvo correlaciones mayores que uno, una catedrática de pedagogía con dificultades para comprender que el cuadrado de una correlación era menor que ésta, etc. Más interesantes que los casos extremos son, sin embargo, los usos sistemáticamente aberrantes: por ejemplo, análisis estadísticos de muestras casuales de 15 o 20 países, la sustitución sistemática de hipótesis sustantivas por la hipótesis nula o la constante confusión entre diferencias estadísticamente significativas y diferencias materialmente importantes.

3. He señalado las dificultades de enseñar Estadística, incluso al nivel más elemental, y la necesidad de una buena comprensión de sus fundamentos para su manejo correcto. Es claro que soy un poco más pesimista que el Profesor García Ferrando cuando afirma que ordenador y paquetes estadísticos son "de fácil manejo por todo aquel estudiante capaz de comprender las nociones fundamentales de un curso introductorio de análisis estadístico y con unas nociones poco más que elementales de informática". Creo que habiendo comprendido **muy bien** bastante más que lo que se aprende en un curso introductorio el estudiante está listo para comenzar un largo proceso de aprendizaje por ensayo y error. Este proceso está necesitado de continua vigilancia, en especial para el planteamiento y la interpretación correcta de los análisis multivariados. La capacidad de cálculo de los paquetes supera claramente la capacidad de interpretación de los estudiantes con esta formación elemental; conozco mucha gente con dificultades para entender, por ejemplo, la relación entre el análisis de varianza y el de regresión múltiple y, por supuesto, para la lectura de la mayor parte de los artículos de las revistas sociológicas. Naturalmente, sólo una pequeña parte de los estudiantes aprovecha el curso elemental de Estadística. En realidad, los estudiantes de estadística se pueden dividir en dos grandes grupos, el poco numeroso de los que entran en la materia y consiguen avanzar en su estudio y el muy numeroso que, refugiándose en la teoría, nunca llegan a tener un atisbo de sus fundamentos.

Vistas así las cosas, se impone de modo aún más apremiante la conclusión del Profesor García Ferrando sobre la conveniencia de reforzar la

Estadística en los planes de estudio, de tal modo que no hubiera licenciado ni diplomado en Sociología incapaz de realizar con soltura análisis multivariados mediante un ordenador. Para lo cual habría que introducir al menos otro curso de Estadística, y además reforzar la enseñanza con numerosas prácticas, dejando claro desde el principio a los alumnos que se trata de materia absolutamente imprescindible o troncal. Las nuevas Facultades en particular, en vez de copiar miméticamente el plan de estudios existente en la Facultad actual, no deberían dejar escapar la ocasión que les brinda la inauguración de los suyos, pues es sabido que es casi imposible reformar un plan de estudios una vez establecido. En mi opinión, el prestigio y la consideración profesional de los titulados en Sociología dependen crucialmente de que la mayoría alcance una formación estadística a por lo menos este nivel.

## COMENTARIO

R. MODESTO ESCOBAR

Universidad Complutense de Madrid  
Centro de Estudios Sociales. Instituto Juan March

Encontramos en el artículo de Manuel García Ferrando la idea central de que en Sociología el análisis idóneo es el multivariante y que gracias a la existencia del ordenador un alumno de esta rama, en principio con escaso bagaje matemático, puede llegar a comprender y utilizar esta técnica a lo largo de los pocos cursos de técnicas de análisis por los que atraviesa en la carrera. Junto con esta idea desarrolla el autor una brillante exposición de cuáles deben ser los conocimientos estadísticos que han de aparecer en un programa de formación de sociólogos.

Al hilo de estos argumentos, me propongo realizar en este espacio una reflexión sobre para qué, dónde y cómo se desarrolla o debería desarrollarse la enseñanza de herramientas de análisis cuantitativo utilizadas en la ciencia sociológica.

La Sociología, una ciencia social, cuyo origen se remonta al inicio del siglo XIX no constituye un cuerpo de conocimientos homogéneos y objetivos tal como soñaron los positivistas al pensar que la aplicación de los principios y métodos de las Ciencias Naturales al campo de lo social daría lugar a una especie de ingeniería humana capaz de mejorar las condiciones de habitabilidad del hombre. Sin embargo, se vislumbra en la actividad de aquéllos que se enorgullecen o tienen complejo de inferioridad de ser sociólogos un conjunto de aproximaciones contrapuestas al mismo objeto. Cabe citar en este contexto dos oposiciones, cuyos extremos se desarrollan indefectiblemente con desigual habilidad o dedicación de tiempo por los profesionales que desempeñan cotidianamente su trabajo.

Tenemos, por un lado, la oposición entre el trabajo teórico y el empírico. El primero pone énfasis en el desarrollo del pensamiento acumulado sobre lo que es la sociedad y cuáles son sus mecanismos de integración y/o cambio. El segundo se basa en la explicación y/o descripción del mundo social observado. Ambos han sido caracterizados y criticados por un clásico de nuestro siglo: Wrigth Mills, quien etiqueta sus consecuencias llevadas al extremo como "gran teoría" y "empirismo abstracto" respectivamente. "Intelectualmente esas escuelas representan abdicaciones de la ciencia social clásica. El vehículo de su abdicación es la pretenciosa superelaboración de 'método' y 'teoría'; y la razón de ésta es la falta de conexión firme con problemas sustantivos" (Mills, 1985; 91).

De esta oposición se deduce claramente que el analista teórico no se interesa, ni hace uso de las técnicas estadísticas de análisis. Pero no todo empirista concibe la estadística como su principal herramienta de trabajo. La segunda polémica que delimita, en el marco del saber sociológico, el campo de interés por el conocimiento y aplicación de los métodos numéricos de análisis es la que se produce por la controversia sobre la naturaleza de los datos sociales y su posible análisis. Quienes conciben la especificidad de las Ciencias Sociales y abogan por una perspectiva comprensiva (M. Weber, 1983; 5) defienden la utilización (casi) exclusiva de métodos cualitativos de análisis. Por el contrario, los partidarios de métodos numéricos reconocen el valor heurístico, o al menos contrastador, de la medición cuantitativas y el análisis estadístico de los datos.

Desde un punto de vista profesional, el titulado como sociólogo puede desarrollar su actividad en tres diferentes campos: la producción, la transmisión o la aplicación de su saber especializado. A estos procesos corresponden las actividades de investigación, enseñanza y ejercicio profesional, **strictu sensu**.

La primera de ellas juega un papel predominante en las disciplinas sociales. "El estudiante de medicina no se plantea que una vez acabada la carrera tenga que hacer investigación médica. Se plantea, ante todo, la aplicación de los conocimientos adquiridos durante sus estudios, conocimientos que son el resultado de horas y horas de investigación hechas por otros... Este no es el caso de la Sociología, que como tal Ciencia, no dispone de leyes o generalizaciones empíricas que hagan posible su aplicación inmediata a la realidad social a través de una práctica profesional. Es ésta una de las razones por las que el sociólogo debe investigar y crear conocimientos en lugar de aplicarlos" (Francisco Alvira, 1984; 57). Es decir, el sociólogo debe producir nuevo conocimiento; pero no sólo porque el desarrollo de nuestras generalizaciones es muy bajo —¿son posibles en nuestros dominios cognitivos?—, sino también porque la realidad que se intenta comprender, explicar o meramente describir es cambiante.

Consecuencia importante de ello es que en la formación de sociólogos profesionales es imprescindible proporcionarles el conocimiento de las herramientas para la investigación y, en general, para el desarrollo del trabajo intelectual. Ahora bien, no hay que caer en la falacia de identificar al potencial investigador con un experto en la Estadística. Hoy en día, en nuestro país y también allende nuestras fronteras, muchos sociólogos (especialmente los académicos) sólo tienen conocimientos muy someros de Estadística e incluso algunos investigadores —especialmente los de más avanzada edad— no pasan más allá del análisis de porcentajes que proporcionan los cuestionarios (instrumento de obtención de datos que para bien o para mal se identifica como el más específico de esta rama del conocimiento). Y no por ello, se han dejado de realizar excelentes estudios sobre la sociedad española, ni acertadas precisiones electorales o fructíferas investigaciones de mercado.

Quizá sean éstas últimas las que en nuestro contexto cultural están impulsando en mayor medida el desarrollo de las técnicas cuantitativas más complejas no sólo en el terreno cuantitativo, sino también en el cualitativo. En el primer aspecto —el más relevante en el contexto de esta reflexión sobre la relación entre Estadística y Sociología— resulta que un análisis multivariante de datos (segmentación, factorial, regresión, cluster o cualquier otro semejante) está más solicitado por la investigación comercial que por la académica. Si en otros contextos culturales distintos al español existen publicaciones sociológicas en la que es necesario incluir una técnica o una fórmula sofisticada de análisis; aquí más bien se fomenta lo contrario. Compruébese, por ejemplo, la serie de artículos aparecidos en la **Revista Española de Investigaciones Sociológicas** o en la **Revista Internacional de Sociología**.

Sin embargo, seamos cautos en lo que concierne a las posibilidades del análisis multivariable. Existe una amplia polémica sobre su uso. Además del carácter cuantitativo que los caracteriza, es probable que las estructuras matemáticas que los sostienen no sean las pertinentes para el análisis de lo social. Incluso cabría también ser cautos en cuanto a los datos que se introducen en el mágico ordenador, siendo así que el desarrollo de las técnicas de medición de los hechos sociales aún están en un estadio de gran imperfección. Un ejemplo claro que tendría que darnos que pensar a los sociólogos cuantitativos son los escasos avances que han otorgado a nuestra disciplina la aplicación de modelos causales matemáticos como el análisis de senderos o el desarrollo de los modelos de ecuaciones estructurales (LISREL). Como Manuel García Ferrando (1985; 164) escribió en otro lugar, "el desarrollo teórico y la construcción de indicadores válidos y fiables de los conceptos claves en la sociología, se encuentran si no estancados, sí al menos muy retrasados con respecto a las técnicas estadísticas".

Surge, pues, de todo este razonamiento la pregunta sobre la utilidad de la Estadística para el sociólogo. La respuesta que se puede deducir de lo anterior es que es importante; pero no fundamental. Además, partiendo de la realidad, comentada por el artículo en esta revista de M. García Ferrando, sobre el escaso conocimiento matemático con el que acceden los estudiantes de Sociología —muchos de los cuales la consideran como una rama de las letras, en lugar de la de ciencias— se podrían considerar distintos niveles de aprendizaje. Bechoffer (1981; 501-502) afirma en un contexto británico de la Sociología que la "premisa de que los estudiantes sólo aprenden los métodos de investigación si son muy elementales o si sienten necesidad de ello, conduce a tres conclusiones: Primero, que debería ser obligatoria la enseñanza de los métodos de carácter sencillo que proporciona a los estudiantes el acercamiento y las destrezas básicas. Segundo, que debería existir la posibilidad de ir más allá de lo simple para aquéllos que lo deseen... En tercer lugar, que debería promoverse cualquier forma de enseñanza que incremente el sentimiento de la necesidad de aprendizaje sobre los métodos".

Esta propuesta de diferenciación no está reflejada en el presente plan de estudios de las Facultades de Sociología españolas. M. García Ferrando expone las tres asignaturas metodológicas (dos obligatorias y una opcional) que componen el conjunto de veintidós asignaturas obligatorias y otra veintena de optativas entre las que hay que seleccionar seis. El mismo autor expone la necesidad y la futura implantación en el nuevo plan de estudios de otra asignatura de Estadística. En la Universidad Autónoma de Barcelona, el plan es bastante diferente como consecuencia de que sólo ofrece enseñanza de la Sociología a partir del segundo ciclo al que se accede tras tres años de estudios en otra Facultad más un curso de adaptación. En éste último se ubica la asignatura de Estadística. En cuarto se imparten las Técnicas de Investigación Social y en quinto existe una materia de Análisis de Datos. La diferencia estriba en que la formación estadística se concentra al final de los estudios, en lugar de en los comienzos. ¿Qué alumno se acuerda al final de la licenciatura de lo que es un coeficiente de correlación, si después de primero no le vuelve a aparecer en ninguna línea del **currículum**?

Bien sabido es que también en la Universidad Complutense se ofrece un programa de doctorado en Metodología con algunos seminarios de contenido estadístico o que existen cada vez más cursos de postgrado, organizados por el Colegio de Licenciados, la Asociación de Sociología, los incipientes Institutos Universitarios u otras entidades públicas o privadas. Sin embargo, habría que insistir en formar para que un licenciado pueda tener un conocimiento profesional, al menos a los interesados. Para eso, habría que ofrecer más cursos optativos en el segundo ciclo a quienes deseen

profundizar en técnicas de análisis cuantitativas o en cualitativas. Hay personas que abogarían por una especialidad de licenciatura en Sociología centrada en la metodología; otros han propuesto una diplomatura (sólo primer ciclo) para formar técnicos destinados a las empresas de investigación de mercados.

Lo penoso de todas estas alternativas es que es muy posible que no se lleven a cabo o quizá, si lo hacen, que sean de escasa calidad por la insuficiencia de profesores que pudieran hacerse cargo de esta tarea. De todos los departamentos de la Facultad de Ciencias Políticas y Sociología de la Universidad Complutense, el de Metodología es al que menos candidatos se presentan para ocupar plazas vacantes. Los encargados de las asignaturas, hoy sociólogos ayer economistas, se han formado fundamentalmente en el extranjero, si no autodidácticamente. La escuela de verano de Essex, financiada por el European Consortium for Political Research y la Universidad de Michigan han ofrecido y ofrecen oportunidades de aprendizaje a quienes han querido o quieran aprender Estadística y el uso de ordenadores. Experiencias interesantes, que se han intentado importar al Instituto de Cooperación Iberoamericana bajo la dirección de J. J. Sánchez Carrión con una escuela de verano de un mes intensivo de duración.

Desde nuestro punto de vista, la enseñanza universitaria de la Estadística con el fin de la profesionalización de la Sociología y de otros campos de las Ciencias Sociales debe incardinarse en un proyecto global de enseñanza metodológica, en la que deben tenerse en cuenta los siguientes objetivos:

- a) estimular el surgimiento de planteamientos adecuados de investigación para su posterior desarrollo correcto.
- b) presentar los instrumentos más usuales para la recogida de datos e insistir en las posibilidades y alcances de cada uno de ellos. (V. gr. la encuesta, la entrevista, la observación participante, el grupo de discusión, el estudio de casos...)
- c) desarrollar en el alumno la capacidad de seleccionar el instrumento más adecuado a las necesidades de información que tiene planteadas.
- d) articular los tres objetivos anteriores de forma tal que el alumno sea capaz de construir diseños de investigación coherentes.

La Estadística debe considerarse como una herramienta para la fase de análisis de los procesos de investigación. En concreto, los objetivos y la metodología de la formación cuantitativa de los alumnos de Ciencias Sociales tendrían que ser los siguientes:

- a) Comprensión de los conceptos básicos de la Estadística.
- b) Desarrollo en el alumno de la capacidad de abstracción que requiere

el uso de formulaciones matemáticas, así como el recuerdo de las reglas básicas del lenguaje y operaciones algebraicas.

c) Potenciación de la habilidad de la aplicabilidad de la Estadística mediante la exposición de los fundamentos y requisitos de los procedimientos de todo análisis cuantitativo de los datos. Es decir, a partir de una formulación teórica o un problema a investigar, el alumno ha de llegar a saber cuál es la técnica estadística más adecuada para su verificación o resolución.

d) La ampliación de la facultad de lectura y utilización de estadísticos provenientes de las siguientes fuentes:

- 1) Publicaciones oficiales.
- 2) Artículos y monografías que utilicen técnicas cuantitativas.
- 3) Salidas de ordenador, producidas por los paquetes estadísticos más usuales: (STAR, BARBO, SPSS, BMDP, SAS...)

e) Fomento de la incipiente capacidad del alumno para traducir el lenguaje matemático al lenguaje sociológico, desarrollándole la habilidad para elaborar informes interpretativos a partir de los datos —ya transformados o susceptibles de transformación— que se les ofrezcan.

## BIBLIOGRAFIA

- ALVIRA, F. (1984). "La investigación sociológica", en S. del Campo, *Tratado de Sociología*. Madrid: Taurus.
- BECHOFFER, F. (1981). "Substantive Dogs and Methodological Tails: A Question of Fit", en *Sociology*, n.º 15, pp. 499-507.
- GARCIA FERRANDO, M. (1985). "Análisis y modelización causal en sociología" en *Revista Española de Investigaciones Sociológicas*, n.º 29, pp. 143-164.
- MILLS, C. W. (1985). *La imaginación sociológica*. México: F.C.E.
- WEBER, M. (1983). *Economía y sociedad*. México: F.C.E.

## COMENTARIO

JUAN JAVIER SANCHEZ CARRION

Universidad Complutense de Madrid

Como complemento al artículo de García Ferrando, centrado fundamentalmente en tratar los contenidos de la enseñanza de la estadística en Sociología, en estas páginas vamos a ofrecer algunas consideraciones de

naturaleza pedagógica sobre estas mismas enseñanzas. En primer lugar nos vamos a referir a los conocimientos y actitudes de los alumnos de Sociología y Ciencias Políticas, para pasar a continuación a hacer algunas consideraciones sobre la forma que creemos más adecuada de explicar estadística en el contexto de las Facultades de Ciencias Sociales— hablando en términos más generales, las reflexiones se pueden extender a la enseñanza de las técnicas de investigación. Como Anexo incluimos nuestro programa de Estadística Aplicada las Ciencias Sociales, asignatura que impartimos en Primer Curso de la Facultad de CC.PP. y Sociología.

## LOS CONOCIMIENTOS Y ACTITUDES

Al hablar de los alumnos de Sociología y su relación con la estadística podemos hacer 2 grupos. Existe el grupo de los cursos iniciales, caracterizado por su deficiente formación matemática —los mismos alumnos suelen decir que son “de letras”—, y con escaso interés por la estadística y, más en general, las técnicas de investigación. Por otro lado está el grupo de los alumnos de los últimos cursos, que si bien por lo que se refiere a su formación matemática son similares a los alumnos de primero, sin embargo cambian en lo relativo a su interés por las técnicas.

El primer grupo lo forma un colectivo preocupado por los problemas sociales, para quienes los números resultan de poco interés —en el límite, si les gustaran los números habrían elegido una carrera de Ciencias. Su formación matemática suele ser deficiente, lo cual tampoco ayuda mucho para introducirles en temas estadísticos.

Los alumnos de los cursos superiores comienzan a preocuparse por su futura salida profesional, y se dan cuenta de que la Sociología como profesión tiene casi todo que ver con la investigación social. A la vez, no toda la investigación social tiene que ver con el análisis estadístico de los datos, pero sí que es evidente que para bien o para mal una gran parte de la legitimidad actual de la Sociología le viene de que hace encuestas en las que se cuenta el número de personas que compran, venden, comen, votan, etc. Y este recuento implica conocer unas técnicas que hacen uso de la estadística. En estos cursos superiores son los alumnos los que andan de cabeza tratando de aprender SPSS/PC+, el análisis factorial y, en general, cualquier técnica que les capacite para el ejercicio profesional.

Las características diferenciales de los alumnos de primeros y últimos cursos hace que el enfoque que se le da a la enseñanza de las técnicas de investigación sea diferente según se trate de uno u otro colectivo. Fundamentalmente, en primero se trata de lograr el interés de los alumnos, algo que se puede dar por supuesto en los alumnos de los últimos cursos.

## LA ENSEÑANZA DE LA ESTADÍSTICA

Antes de entrar a explicar nuestro programa de estadística queremos añadir un par de consideraciones a nuestras reflexiones anteriores sobre los conocimientos y actitudes de los alumnos de Sociología.

La primera tiene que ver con el hecho de que las clases a las que nos dirigimos están compuestas por más de 100 alumnos, de los que puede haber un 60 % que, como consecuencia de la política ministerial de obligar a elegir carrera en función de las notas de la Selectividad, están allí porque no les dejan estar en otro sitio. Este elevado número de alumnos y su heterogénea composición dificulta enormemente cualquier proyecto de enseñanza de una materia experimental como la estadística.

La segunda consideración está relacionada con el carácter "problemático" de la enseñanza de la estadística en los primeros cursos del área de las ciencias sociales —problematicidad que parece compartir la estadística con otras materias técnico-matemáticas. Dicho muy brevemente: la enseñanza de la estadística en las Facultades de Ciencias Sociales dejará de ser un problema el día en que no tenga que explicarse. Está claro que hay que saber estadística para realizar investigaciones sociales, lo que desde un punto de vista ideal no está tan claro es que la enseñanza de la estadística —por lo menos a nivel de los primeros cursos— haya que efectuarla en una asignatura independiente de otras asignaturas de tipo sustantivo (estructura social, sociología electoral, cambio social, etc.). Nuestra opinión es que sería más lógico fundir la enseñanza de materias sustantivas y técnicas.

Conocemos experiencias en las que se explica estructura social investigando sobre la estructura social. En el mismo curso se explican diferentes teorías sobre la movilidad social, las clases sociales, etc., se enseñan los conocimientos estadísticos y se facilitan datos de la estructura social del país, bien en forma bruta (matrices de  $p$  variables por  $n$  unidades) o elaborada (matrices de correlaciones, tablas de contingencia, etc.). Luego, en vez de ser el profesor el que dice que hay tantos individuos con estudios superiores (distribución de frecuencias de la variable), que además resultan ser aquellos cuyos padres también tienen una formación superior (relación entre las variables "estudios padre" y "estudios hijo"), con lo cual queda claro que en el país existe una movilidad ... son los propios alumnos los que descubren estos aspectos analizando los datos que se les facilitan. De esta manera pueden estudiar un tema de estructura social al mismo tiempo que aprenden estadística.

La experiencia a la que nos hemos referido se realiza en el curriculum de Sociología de la Universidad de Harvard. En este curriculum son los profesores los que proponen los nombres y contenidos de sus asignaturas. En

nuestra Universidad, donde el profesor se encuentra con una asignatura predeterminada, a la que debe dar respuesta, resulta difícil fundir dos asignaturas (por ejemplo, las que estamos utilizando como ilustración: Estructura Social y Estadística) en una, siquiera sea experimentalmente, salvo que 2 profesores junten sus programas. No por ello dejamos de pensar que este sistema "integrado" de enseñanza es el más adecuado para enseñar estadística y sociología.

Hechas estas precisiones, digamos que nuestra forma de enseñar Estadística en Primer Curso pasa por combinar las explicaciones sobre las distintas técnicas de análisis, con el análisis que los alumnos hacen de datos obtenidos en una investigación por encuesta, que ellos mismos desarrollan en todas sus fases. El curso se divide en grupos de 4 o 5 alumnos, y cada grupo diseña un pequeño cuestionario, con el cual luego recoge información de aproximadamente 60 unidades.

Al comienzo del curso se explica el sistema operativo D.O.S., que es el que utilizan los ordenadores del Laboratorio de Informática de nuestra Facultad, además de un programa de análisis de datos muy sencillito, basado en menús. De esta manera, simultáneamente a nuestras explicaciones estadísticas los alumnos realizan los análisis pertinentes con sus datos: calculan medias, coeficientes de regresión, hacen tablas de contingencia, etc. Previamente han grabado sus datos en el ordenador y han creado un fichero. Al final del curso cada grupo tiene que entregar el resultado de la pequeña investigación que han realizado, en la que se supone que han aplicado de forma selectiva todas las técnicas estadísticas que se han explicado en el curso.

La experiencia es muy positiva. Hemos comprobado que no es lo mismo aprender a calcular, por ejemplo, el coeficiente de regresión o la  $\chi^2$  cuadrado con unos datos facilitados por el profesor —y seleccionados expresamente para ilustrar estos estadísticos— que aprender a calcular estos estadísticos y, además, saber dónde utilizarlos entre todas las variables de un cuestionario. Además, con este sistema los alumnos empiezan a familiarizarse con la presentación de las tablas en un informe, aprenden a comentarlas y, algo que es muy importante, tienen que escribir.

A continuación incluimos nuestro programa, donde se pueden ver los contenidos que impartimos y la forma como están organizados.

## PROGRAMA DE ESTADÍSTICA

### Objetivos del Curso

Tanto las Tablas de Contingencia como la Regresión son las técnicas estadísticas más utilizadas en las Ciencias Sociales. En este curso vamos a

centrarnos en la explicación de ambas técnicas.

El procedimiento que vamos a seguir es plantear problemas sociológicos, fáciles de entender, para mostrar a continuación el papel que tiene la estadística a la hora de darles una solución.

Las dos técnicas que hemos elegido proporcionan el lenguaje y los conceptos suficientes para que, superando su dimensión puramente técnica, podamos adentrarnos en los problemas reales de la investigación social.

En este curso se pondrá el acento en los conceptos fundamentales del análisis de los datos, sin profundizar en la teoría estadística.

En las páginas que siguen vamos a presentar el contenido del programa. Lo hemos dividido en dos partes. En una primera se hace una relación de los problemas sociológicos que vamos a estudiar en el curso (**Problemas Sociológicos**). Al final de cada problema se pone un número entre paréntesis, que hace referencia a los conceptos estadísticos que se introducen en ese problema. En la segunda parte del programa se muestra cuál es la relación de conceptos estadísticos (**Conceptos Estadísticos**).

Durante el curso los alumnos, organizados en grupos reducidos, analizarán unos datos que ellos mismos han de crear. A tal fin harán uso del paquete de análisis de datos en ordenador que se explicará en clase.

Para superar el curso es necesario aprobar el examen final y presentar el análisis al que se alude en el párrafo anterior.

## Problemas sociológicos

1. ¡Ojo! No todos los datos de las Ciencias Sociales son iguales: los datos difieren según el nivel de medida o el papel que juegan en la investigación; además, no todos se refieren a unidades del mismo tipo y hasta puede que traten de cosas que ocurren en diferentes momentos del tiempo. (1)

2. Estudio de la ideología de los alumnos de primer curso de CC. PP. y Sociología.

2.1. La ideología de los alumnos y algunas otras de sus características. (2)

2.2. ¿Adivina con que tiene que ver (se relaciona) la ideología?. (3)

2.3. Diferentes explicaciones (modelos) acerca de la ideología. (4)

2.4. De padres a hijos, ¿qué ha cambiado?. (5)

- 2.5. ¿Qué decir de la ideología de todos los alumnos de primero de España?. (6)
3. ¿Por qué ganó el gobierno el Referendum de la OTAN?
  - 3.1. Evolución de la actitud de los españoles con el paso del tiempo. (7)
  - 3.2. ¿Qué influencia tuvo la ideología de los españoles en el resultado? ¿Y el voto?. (8)
4. El mundo en cifras: indicadores socio-económicos de los mayores países del mundo.
  - 4.1. De aquella situación en la que dos personas tenían dos pollos y una se murió de hambre. (9)
  - 4.2. Adivina con qué tiene que ver (se relaciona) la tasa bruta de natalidad. ¿Y el analfabetismo?. (10 y 11)
  - 4.3. ¿Seguro que cuánto mayor es el PNB menor es la tasa bruta de natalidad?. (12)

### **Conceptos Estadísticos**

1. Conceptos fundamentales de la investigación social.
2. Distribución de frecuencias con variables nominales y ordinales. Su representación gráfica.
3. Análisis de tablas de contingencia: anatomía y representación gráfica de las tablas. Estadísticos para resumir tablas
4. La lógica de la investigación causal.
5. Análisis de tablas de movilidad.
6. El muestreo: conceptos fundamentales. Distribuciones muestrales. Intervalos de confianza. Valores observados y esperados (la Gi-cuadrado y el tamaño de N). Análisis de los residuos.
7. Análisis del cambio con una sólo variable.
8. El análisis del cambio con dos o más variables.
9. Distribución de frecuencias con variables intervalos: organización, resumen, presentación gráfica y transformación de los datos.
10. Representación gráfica de la relación entre 2 variables intervalos: la nube de puntos.

- 
11. Fundamentos de la regresión: líneas, ecuaciones y datos. El coeficiente de determinación,  $R^2$ .
  12. El análisis de los residuos. La regresión múltiple.

## **Breve respuesta a cinco comentarios sobre el trabajo "Estadística y Sociología"**

por  
MANUEL GARCIA FERRANDO

Poco tengo que objetar a los comentarios críticos que mis colegas, los profesores Alvira, Beltrán, Carabaña, Escobar y Sánchez Carrión, han hecho a mi trabajo sobre el papel de la Estadística en la Ciencia Sociológica y en el panorama de estudios de la Licenciatura en Sociología, y por extensión de la Licenciatura en Ciencia Política. De la lectura de tales comentarios he deducido que sus cinco autores coinciden en lo fundamental con las tesis principales por mí desarrolladas, a saber: el carácter instrumentalmente necesario de la estadística en la investigación social y en el programa de estudios de Sociología; el uso creciente que se realiza de la Estadística en el actual trabajo sociológico, lo que reforzaría lo anterior; y dada la amplia generalización de los ordenadores en las rutinas informáticas y de cálculo no sólo del mundo científico como del resto de las actividades sociales, la necesidad de integrar la enseñanza de la Estadística con la de los paquetes estadísticos para ordenadores, especialmente de los llamados personales.

Alrededor de estas tesis, los comentarios han desarrollado su argumentación crítica, expuesto sus puntos de vista y experiencias personales en tanto que profesores de Sociología e investigadores profesionales, e incluso hay quien como el profesor Sánchez Carrión, ha relatado su particular, e interesante, forma de enseñar Estadística a los estudiantes de Sociología y Ciencias Políticas, o como ha hecho el profesor Escobar ha desarrollado, sintéticamente, un proyecto global de enseñanza metodológica. El resultado de todo ello entiendo que ha enriquecido enormemente el contenido de mis tesis originales, por lo que como he indicado al comienzo de esta breve nota de respuesta, poco tengo que añadir a los argumentos de mis comentaristas como no sea manifestarles mi agradecimiento por el tono constructivamente crítico, y en algún momento incluso elogioso, de sus observacio-

---

nes. Igualmente, deseo extender mi agradecimiento a la Revista Estadística Española y a su director, por la oportunidad que nos ha brindado a los sociólogos de exponer públicamente nuestras experiencias y opiniones sobre la enseñanza de la Estadística en nuestros programas de estudios, en el contexto de un número monográfico de la Revista dedicado a debatir la problemática de la enseñanza de la Estadística en las Ciencias Aplicadas.

