

MANUEL GARCIA FERRANDO. Socioestadística. Alianza Editorial, Madrid. 1.ª Edición 1985, 1.ª Reimpresión 1987. 491 páginas.

MANUEL GARCIA FERRANDO. Ibáñez, Jesús y Alvira, Francisco (compiladores).– El análisis de la realidad social. Métodos y técnicas de investigación. Alianza Editorial, Madrid. 1.ª Edición 1986. 461 páginas.

por
GONZALO ARNAIZ TOVAR
Ministerio de Economía y Hacienda

El primero de los libros que se van a comentar, según afirma su autor en la introducción, responde a la necesidad de “contar en nuestra universidad con un texto introductorio de estadística para sociólogos que, sin entretenerse excesivamente en la derivación de fórmulas matemáticas, ofreciese una visión rigurosa y amplia del uso actual del análisis estadístico en la investigación social”.

El segundo, en cuya redacción han intervenido catorce autores, pretende cubrir el temario de lo que puede ser un Curso Introductorio de nivel intermedio de Técnicas de Investigación Social. Los compiladores aspiran con el libro “a ofrecer una visión de conjunto de las diferentes perspectivas de investigación social más utilizadas de un lado y de otro, con el propósito manifiesto de que el estudiante y el estudioso de las Ciencias Sociales encuentren en él una aproximación lo suficientemente rigurosa, precisa y crítica a los procedimientos más utilizados de la investigación de la realidad social. Aproximación que, por otra parte, aspiramos permita ir más allá de la simple y, con frecuencia, ingenua dicotomía cuantitativo/cualitativo”. Por consiguiente el libro presenta un temario más amplio y con distinta orientación que el primero y, aunque tiene determinados puntos de contacto, posibilita, sobre todo, situar el análisis cuantitativo en sociología y, en general, en las ciencias sociales, en una perspectiva más amplia, aceptando “el carácter problemático de todo saber social”.

Socioestadística opta, lo que parece acertado, por presentar el temario en tres partes claramente diferenciadas. La primera (capítulos, 2, 3, 4, 5 y 6) trata del análisis estadístico cuando se estudia una variable en la población; la segunda (capítulos 7, 8, 9, 10 y 11) se refiere al estudio de dos variables y la última (capítulos 12, 13, 14 y 15) a varias variables.

Cada parte se estructura de forma parecida: se empieza con el análisis descriptivo para pasar después a los aspectos inferenciales. La primera parte cuenta con una buena introducción a la Estadística Descriptiva, presentando intuitiva y sencillamente los conceptos, haciendo hincapié en los procedimientos gráficos y comparando las ventajas e inconvenientes de las distintas medidas que tratan de resumir los datos. Sin embargo, hay que hacer dos observaciones:

– La primera se refiere a las páginas 94 y 95 donde se presentan “otras propiedades interesantes de la media”. Tal y como se presentan queda la duda de por qué son interesantes ya que no se hace ninguna precisión en ese sentido ni se muestra si otras medidas de tendencia central cumplen las propiedades. Además, la “demostración” que se da de la propiedad b) es incomprensible y la propiedad d) está confusamente enunciada.

– La segunda es la forma de presentar la distribución normal. Dada la importancia de dicha distribución no parece adecuado presentarla donde se hace, sin haber introducido todavía el concepto de población ni de parámetros de la misma, lo que desdibuja la presentación. Aunque el autor ha mantenido, lo cual es de agradecer, un nivel matemático elemental, no es conveniente realizar afirmaciones del tipo “... resulta útil, desde un punto de vista matemático, pensar en términos de distribuciones basadas en un número infinito de casos. Tales distribuciones vendrán representadas por curvas cuyos extremos se van acercando suavemente al eje X, pero sin cruzarse con él”, que, además de incorrectas e incomprensibles, contribuyen a fomentar el “argumento de autoridad matemática”.

Por último, se echan de menos en esta parte algunas consideraciones sobre las transformaciones de los datos.

Los capítulos 4, 5 y 6 presentan el aspecto inferencial básico. Se comienzan (capítulo 4) con una breve introducción a las probabilidades para pasar a presentar posteriormente el muestreo. Se trata de un capítulo informativo y bien construido; como el propio autor indica en una nota de pie de página, los apartados 4.3.3. (Estimadores y errores de muestreo) y 4.3.4. (Determinación del tamaño de la muestra) rompen el esquema seguido y deben leerse después de los capítulos 5 y 6.

El capítulo 5 entra propiamente en el terreno de la inferencia, presentando, los contrastes de hipótesis en su aspecto general. Es ésta una decisión arriesgada ya que quizá no sea la mejor forma de empezar introduciendo la inferencia por los contrastes pues, si no se procede con cuidado, pueden quedar desdibujados los conceptos básicos de parámetros que caracterizan la distribución de la variable en la población, estimador, estadístico, distribución muestral, etc. Algo de esto ocurre en el capítulo donde no se aclaran suficientemente las cosas ni se presenta un ejemplo sencillo para seguir el razonamiento. Como ejemplo de terminología confusa y errónea valga el siguiente; en la página 162 se escribe: “supongamos que p_x es la proporción de una determinada característica o propiedad que estamos estudiando, perteneciente a una población de distribución normal... Cuando es cierta la hipótesis nula ($p_x = p$), la proporción p_o , observada en la muestra estará dentro del intervalo de probabilidad $1 - \alpha$ ”. Por último, el Teorema del Límite Central que se enuncia en el capítulo se llama, erróneamente, Ley de los Grandes Números.

El capítulo 6 presenta algunos ejemplos de contrastes (proporciones, chi-cuadrado, medias) y, al final, proporciona una visión muy correcta de la estimación por intervalo.

El libro hubiera ganado en claridad si esta parte hubiera servido para introducir las nociones fundamentales de la inferencia. La definición de sesgo de la página 197 es insostenible: "Se dice que una estimación no está sesgada cuando su distribución muestral es exactamente igual al valor del parámetro que se ha estimado. Esto es, el valor esperado a la larga es el propio parámetro".

La segunda parte del libro, dedicada a la estadística descriptiva e inferencial de dos variables contiene un capítulo muy interesante e informativo sobre las medidas de asociación para variables nominales y ordinales, donde se presentan intuitivamente los distintos coeficientes y los problemas que plantean, un capítulo sobre contrastes para dos muestras, un capítulo sobre regresión y correlación y otro sobre análisis de varianza. Estos dos últimos capítulos son muy claros; quizá debería hacerse más referencia a la validación de las hipótesis sobre las que se basan los modelos y, en concreto, al análisis de los residuos. Es un acierto esperar a desarrollar el análisis de la varianza para presentar los contrastes sobre el modelo de regresión. Hay que hacer una objeción a la forma errónea de presentar las hipótesis nula y alternativa del análisis de la varianza en la página 312: "La hipótesis nula afirma la igualdad de las medias, mientras que la hipótesis alternativa afirma que cada grupo considerado tiene una media diferente".

La parte tercera, ofrece una panorámica sobre el análisis multivariable. Se trata de, sin entrar en muchas profundidades como no puede pretender el libro, presentar una visión informativa de distintas técnicas, centrándose fundamentalmente en la regresión múltiple y el método del análisis del camino (path analysis) y en el análisis factorial de los que ofrece una introducción muy buena, planteando, además, los problemas de los distintos métodos y las cautelas a la hora de interpretar los resultados.

Al final de los capítulos del libro se presentan numerosos problemas bien escogidos.

En resumen, se trata de un libro recomendable que cumple el objetivo del autor de ofrecer una visión amplia del uso del análisis estadístico en la investigación social. Sin embargo, la parte de introducción a la inferencia, que es fundamental a la hora de transmitir sentido estadístico a los lectores, como ya se ha apuntado anteriormente, adolece de deficiencias y errores que deben ser subsanados en futuras ediciones del libro.

El segundo de los libros, al que se va a dedicar un comentario más breve, es muy recomendable, no sólo para todos los estudiosos de las ciencias sociales, sino también para la generalidad de los estadísticos.

Con independencia de que los capítulos son desiguales y que algunos son de carácter bastante técnico, el libro tiene la virtud, por un lado, de plasmar la experiencia y la reflexión sobre su trabajo de un conjunto de sociólogos y, en ese sentido, contiene en sus líneas múltiples ejemplos y enseñanzas de interés tanto de diseños experimentales, como de obtención y análisis de datos. Por otro, el libro, sobre todo en la primera y segunda parte (El Diseño de la Realidad Social), posibilita la reflexión sobre algo muy

importante para todos los implicados, y muchas veces “absortos”, tanto directa como indirectamente, como estadísticos, en el análisis de la realidad social: el por qué el para qué y el cómo de la investigación.

Los capítulos de carácter más específicamente cuantitativo presentan el problema de la medición en las ciencias sociales, una introducción sintética al muestreo (aunque los aspectos inferenciales, de nuevo, dejan bastante que desear), una muy buena introducción panorámica, al análisis de datos, una interesante reflexión, aunque demasiado sintética, sobre el análisis de los datos de encuesta y, por último, un capítulo descriptivo sobre ordenadores.