

Revisión de libros

Principios de Estadística Aplicada (Jorge Ortiz Pinilla)

José Fernando Zea^a
jfzeac@gmail.com

El libro *Principios de Estadística Aplicada* introduce al lector al razonamiento estadístico. A diferencia de la mayoría de textos, enfatiza en las fuentes de datos, los mecanismos de selección de estos, temas para los cuales dispone de diversos ejercicios, tanto prácticos como teóricos. Cada fórmula que aparece en el texto está acompañada de su razonamiento y la explicación de su construcción, y se complementa con ejercicios teóricos y prácticos que permiten al estudiante tener una comprensión detallada de cada uno de los temas, sin necesidad de recordar fórmulas que en muchos casos son memorizadas por estudiantes y aprendices sin comprender realmente qué conceptos están detrás de cada una de las fórmulas que se utilizan en la práctica cotidiana de la estadística.

El libro tiene, además de sus fortalezas conceptuales, ejemplos en el paquete estadístico R en los capítulos referentes a inferencia estadística y en los comienzos del libro tiene una propuesta de ejercicios para adquirir las nociones básicas del paquete.

El autor comienza por establecer una distinción entre los estudios observacionales y los estudios experimentales, y desarrolla los conceptos básicos estadísticos de población, el método de observación y la estrategia de muestreo, todo enmarcado dentro de un contexto y una delimitación conceptual que es inherente a la disciplina en la cual la estadística está siendo utilizada para resolver un problema. A diferencia de la gran mayoría de los textos introductorios que inmediatamente después de exponer los conceptos básicos proceden a desarrollar el análisis con datos agrupados, el autor aborda desde sus comienzos las reglas fundamentales del conteo, las diferentes formas de extracción de objetos de una urna, analogía que permite comprender las diferentes estrategias de muestreo utilizadas en la práctica, en el capítulo se desafía al estudiante a comprender una amplia gama de ejemplos y desarrollar ejercicios que permiten al lector entender que antes de emprender cualquier estudio hay diferentes mecanismos con los cuales se puede acceder a los elementos de una población y la manera de cuantificar las posibilidades que se tienen para cada caso particular.

En el tercer y cuarto capítulo del libro, el autor desarrolla los temas con los que tradicionalmente los libros introductorios de estadística comienzan: los conceptos

^aDocente de la Facultad de Estadística, Universidad Santo Tomás.

de población, características, las variables, sus clasificaciones y la estadística descriptiva. Las variables son funciones de los elementos del espacio muestral y, basado en el rango que toman las variables (que pueden ser números reales, subconjuntos de estos o etiquetas), se pueden clasificar en variables cualitativas o cuantitativas y específicamente pueden ser numéricas, intervalo, de razón, continuas, discretas, de clases numéricas, de categorías ordenadas, rangos y nominales, los diferentes ejemplos y ejercicios que desarrolla el autor logran que a esas alturas el lector comprenda así las etapas iniciales de una investigación estadística.

En el capítulo cuatro, el concepto de colectividad es abordado como cualquier subconjunto de la población y parte de un conjunto más grande: la σ -álgebra. Enmarcados los elementos en este conjunto y definidas las variables de medición, el autor procede con una exposición clásica de la construcción de tablas de frecuencias, la elaboración de gráficos básicos, las técnicas de agrupamiento de variables continuas y la elaboración de estadísticas resúmenes agregadas, cada uno de los temas desarrollados con un formalismo matemático que raramente tiene un texto introductorio en estadística, pero sin abrumar al lector neófito gracias a sus diferentes visualizaciones, ejemplos y la pertinencia de los ejercicios propuestos. En este capítulo el autor define el promedio aritmético con tres enfoques diferentes: como el valor más cercano de los datos, un primer modelo estadístico, y desde el punto de vista físico, como el centro de gravedad de los datos. Adicionalmente, otro aporte interesante es la construcción de las medidas de dispersión a partir de las diferencias que se presentan entre todos los puntos, el autor desarrolla esta medida en muestreo con reemplazo y en muestreo sin reemplazo, derivando las dos expresiones ampliamente conocidas de la varianza; este aporte es muy interesante y permite al lector decidir cuándo utilizar una de las fórmulas de la varianza, decisión que depende del esquema de selección para el caso de una población finita. A través de los ejercicios propuestas se desarrolla la descomposición de la varianza, tema raramente tratado en estadística descriptiva y que es un primer acercamiento al análisis de varianza de un modelo lineal.

El capítulo cinco, Asociación, desarrolla el análisis estadístico bivariado, las medidas de covarianza y correlación, sus diferentes propiedades son derivadas de una manera muy intuitiva, asimismo desarrolla el análisis estadístico de combinaciones lineales y el método de regresión de mínimos cuadrados.

Un aporte muy original del libro es el desarrollo estadístico que se realiza sobre poblaciones extraídas con un esquema de urna, el autor analiza así datos dicotómicos y numéricos, derivando para cada uno de estos casos los valores esperados y varianzas de los diferentes esquemas lo cual permite al lector comprender el uso de las dos diferentes varianzas, con denominador n y con denominador $n - 1$ y tener una mejor comprensión del análisis estadístico de poblaciones finitas. El autor enfoca el análisis de urnas, utilizando únicamente las técnicas de conteo sin el uso formal de conocimientos de probabilidad. Otra temática desarrollada en el capítulo es la del teorema del límite central, teorema que se introduce para una muestra finita obtenida con reemplazo a través del uso de simulación, si bien el autor desarrolla su forma funcional, lo cual puede ser prematuro sin tener los

conceptos de probabilidad que el autor deja hasta el siguiente capítulo. Las distribuciones muestrales asociadas a las distribuciones χ^2 , *t*-student y *F* se ilustran al final del capítulo, las lecturas de estos temas sin conocimientos previos de probabilidad pueden dificultarse, por lo tanto, sería deseable abordar el capítulo siete sobre probabilidad.

En el capítulo siete, el autor trata de una manera intuitiva los conceptos básicos, las funciones de probabilidad y las funciones de densidad y presenta una buena gama de ejercicios para el lector, le proporciona los conocimientos necesarios para comprender satisfactoriamente los principios de inferencia estadística, temas que se desarrollan en los dos últimos capítulos del libro (capítulos ocho y nueve), en los que se desarrollan de manera muy esquemática y ordenada los procedimientos para realizar juzgamientos de hipótesis y estimación puntual y por intervalo de una media.

Finalmente, en el último capítulo se explica la inferencia sobre distribuciones de probabilidad, conocida como la binomial (para proporciones), para conteos de éxitos en un esquema sin reemplazo (hipergeométrica) y para medias, diferencias de medias, datos pareados, pruebas de varianzas y cocientes de varianzas, cada una de estas pruebas desarrollada con datos reales en el paquete estadístico R.

El libro es ideal para lectores que hayan tenido una exposición previa a la estadística y quieran madurar su pensamiento estadístico y comprender la racionalidad que hay detrás de cada método, fórmula y gráfico que se desarrolla en cualquier curso introductorio. Como primer acercamiento a la estadística, es un libro muy interesante para matemáticos, ingenieros y en general profesionales que tengan una buena fundamentación matemática, ya que con el lenguaje de la matemática se puede comprender toda la riqueza conceptual que tiene el libro. El libro es una referencia rápida para estadísticos y usuarios de la estadística que utilicen R en lo concerniente a pruebas de hipótesis gracias a sus buenos ejemplos en este paquete.

Recibido: 6 de marzo de 2013

Aceptado: 20 de marzo de 2013