

PROBABLEMENTE, EL MEJOR LIBRO SOBRE ESTADÍSTICA JAMÁS ESCRITO



**Cómo vencer a las probabilidades
y tomar mejores decisiones**

HAIM SHAPIRA

PROBABLEMENTE, EL MEJOR LIBRO SOBRE ESTADÍSTICA JAMÁS ESCRITO

**Cómo vencer a las probabilidades
y tomar mejores decisiones**

HAIM SHAPIRA

PROBABLEMENTE, EL MEJOR LIBRO SOBRE ESTADÍSTICA JAMÁS ESCRITO

DE LA EDICIÓN ORIGINAL

© Haim Shapira, 2024

Primera edición publicada en el Reino Unido y en los Estados Unidos en 2024 por Watkins, sello editorial de Watkins Media Limited, Unidad 11, Shepperton House, 83-89 Shepperton Road, Londres N1 3DF. Con la intermediación de Vicki Satlow of The Agency srl.

DE LA EDICIÓN ESPAÑOLA

© Editorial Planeta, S.A., 2025

De la traducción: Pedro Pacheco, 2025

Realización: Planeta

1.^a edición

geoPlaneta

Diagonal 662-664. 08034 Barcelona

geoplaneta@planeta.es - www.geoplaneta.com

ISBN: 978-84-08-30833-1

Depósito legal: B. 11.680-2025

Impresión y encuadernación: Liberdúplex

Printed in Spain - Impreso en España

La lectura abre horizontes, iguala oportunidades y construye una sociedad mejor. La propiedad intelectual es clave en la creación de contenidos culturales porque sostiene el ecosistema de quienes escriben y de nuestras librerías.

Al comprar este libro estarás contribuyendo a mantener dicho ecosistema vivo y en crecimiento.

En **Grupo Planeta** agradecemos que nos ayudes a apoyar así la autonomía creativa de autoras y autores para que puedan seguir desempeñando su labor. Dirígete a CEDRO (Centro Español de Derechos Reprográficos) si necesitas fotocopiar o escanear algún fragmento de esta obra. Puedes contactar con CEDRO a través de la web www.conlicencia.com o por teléfono en el 91 702 19 70 / 93 272 04 47.

Queda expresamente prohibida la utilización o reproducción de este libro o de cualquiera de sus partes con el propósito de entrenar o alimentar sistemas o tecnologías de inteligencia artificial.



SUMARIO

PRIMERA PARTE: ESTADÍSTICA

1. La estadística brilla con luz propia	11
2. Superados por los números	17
3. El atraco con jugo de limón y el efecto Dunning-Kruger	31
4. Personas y porcentajes	43
5. Percentiles altos y bajos	53
6. ¿Qué es la pobreza?	59
7. Una lección sobre los promedios	65
8. Medias móviles	81
9. Algunos dígitos principales	95
10. Las historias nunca están completas	107
11. Distribuciones y desviaciones	121
12. ¿Vale más un gráfico que mil palabras?	135
13. Escalando hasta la cima	141
14. Correlación no es lo mismo que causalidad	149

SEGUNDA PARTE: PROBABILIDAD

15. El mundo de las probabilidades	175
16. El lado oscuro de la moneda	179
17. La historia de los tres prisioneros	183
18. El lado (ligeramente) más brillante de la moneda	189
19. El caso del jugador equivocado	193
20. Cuando lo improbable pasa a ser probable	199
21. Contra todo pronóstico	205
22. Manual sobre probabilidades	211

23. Un juego de niños solo para adultos	223
24. La (supuesta) paradoja del Caballero de Méré	235
25. ¿Quién quiere ser millonario?	241
26. Canastas y cumpleaños	247
27. Pánico, crímenes y fechorías	257
28. La triste historia de Sally Clark	267
29. El juicio del siglo	273
30. Amaneceres, duelos y democracia	279
31. El cuervo blanco	297
32. Grandes esperanzas	301
33. Una cuestión de sobres	315
Una última oportunidad	323
Agradecimientos	333
Notas	335
Índice	339

LA ESTADÍSTICA BRILLA CON LUZ PROPIA

Cada noche, antes de irme a dormir, pongo dos despertadores, uno al alcance de la mano y el otro un poco más lejos. Lo hago porque sé que las probabilidades de que me despierte sin ellos son casi nulas, y si solo tuviera uno al alcance de la mano, podría apagarlo sin necesidad de levantarme y la probabilidad de que me levantara seguiría siendo bastante escasa. De ahí que necesite dos.

Cuando salgo de la cama y me preparo el desayuno, me doy cuenta de que mi queso favorito contiene (lamentablemente) un 38 % de grasa y 450 calorías por cada 100 gramos. De camino al trabajo, pongo en marcha la aplicación de navegación por satélite Waze, que envía datos como la ubicación y la velocidad a todos los demás usuarios. Gracias a estos datos, Waze proporciona información sobre los atascos, los obstáculos y los radares de tráfico. Una luz de advertencia parpadea en el salpicadero, avisándome de que el aceite del motor está un 4 % por debajo de la cantidad óptima.

Llego al trabajo y abro el periódico de la mañana, y en la portada se habla de una nueva encuesta de opinión pública con los mismos resultados de siempre. Por no hablar de las páginas de deportes, repletas de todo tipo de tablas, datos y estadísticas. O de la sección de economía.

Dejo el periódico a un lado y me conecto a internet. El boletín mensual al que estoy suscrito me sorprende con una estadística fantástica: una de cada 400 personas que viven en la actualidad es descendiente de Gengis Kan, el conquistador mongol de Asia y Europa del Este. Como es evidente que

prefiero las bicicletas a los caballos, probablemente soy una de las 399 restantes.

De regreso en casa, veo las noticias de la noche. La pantalla de mi televisor está repleta de los datos y las frases de carácter estadístico que aportan los presentadores y reporteros. Con cierta asiduidad, cada noche vemos reportajes sobre temas como el salario medio en Estados Unidos, el aumento de los costes de la construcción, los índices de desigualdad económica en el Reino Unido o la proporción de sexos entre los estudiantes de enseñanza superior en Israel (hay mayoría femenina tanto en los estudios de grado como en los de máster, e incluso en los de doctorado).

También utilizan expresiones como «índice de pobreza» o «índice del coste de la vida»; me pregunto cómo los calculan.

Cansado de las noticias, me conecto a mi canal de YouTube para leer algunas reacciones a un vídeo que publiqué la noche anterior. Veo que tiene 1003 visitas, y puedo agruparlas por país de origen, grupo de edad o plataforma de acceso a YouTube. Y, entre otros datos, me muestra la valoración de ese vídeo en comparación con los diez últimos que he subido.

Los números, los datos, el pensamiento estadístico y las probabilidades son casi omnipresentes en nuestro mundo moderno y resultan fundamentales en nuestras vidas, seamos o no conscientes de ello.

LAS ESTADÍSTICAS ESTÁN EN TODAS PARTES

A principios del 2020, cuando empezó la pandemia de la covid-19, el público no solo era consciente de cómo se propagaba la enfermedad a escala mundial, sino que también conoció términos estadísticos como «coeficiente de infección», «falso positivo» y «falso negativo». Puede que oyeras hablar de la diferencia entre escalas lineales y logarítmicas (véase el capítulo 13). Oímos hablar de porcentajes de personas vacunadas, no vacunadas y hospitalizadas; de la edad media de las personas que murieron a causa de la enfermedad en comparación con la longevidad de la población en general; y del ritmo al que la pandemia se estaba extendiendo a nivel local

y mundial. Nos informaron sobre el número de víctimas mortales por millón de personas; sobre la tasa de ocupación de camas en las unidades covid de los hospitales; y sobre el número de personas aterrorizadas por la posibilidad de enfermar. Nos hablaron de la inmunidad de rebaño y reflexionamos sobre la correlación (otra palabra de moda en estadística, de la que hablaremos más adelante) entre el estatus socioeconómico y las posibilidades de contraer el virus. En resumen, todos esos análisis podrían explicarse en un curso académico decente de introducción a la estadística, junto con un breve taller sobre probabilidades.

Las estadísticas de este tipo también son muy importantes para todo Gobierno que se precie. Gracias a ellas, pueden tomar decisiones calculadas sobre salud pública, seguridad, transporte, medio ambiente, delincuencia, cuidado de los ancianos, educación y muchas otras cuestiones relevantes. Los modelos estadísticos desempeñan un papel crucial en casi todas las predicciones relativas a terremotos, tarifas de seguros, cambios políticos, etcétera. Las estadísticas ayudan a gestionar las crisis y a planificar con suficiente antelación catástrofes naturales o daños colaterales de las guerras.

El personal sanitario utiliza las estadísticas para analizar las reacciones de los pacientes a los distintos tratamientos. Los expertos médicos las usan para extraer conclusiones de los datos recopilados y tomar decisiones. Las estadísticas muestran si un determinado fármaco ayuda a curar una determinada enfermedad, y los médicos recurren a ellas para, por ejemplo, decidir si deben recomendar a los pacientes con cáncer de próstata someterse a una prostatectomía radical o limitarse a vigilar cuidadosamente su estado. Familiarizarse con la terminología y las nociones de estadística y probabilidad podría evitar numerosos errores y sesgos en la investigación médica.

ESTADÍSTICAS PELIGROSAS

Una vez mantuve con mi buen y sabio amigo el profesor Igor Elman, de la Facultad de Medicina de la Universidad de Harvard,

una conversación sobre el uso de las estadísticas en medicina. Para exponer un ejemplo de estudios irresponsables, Igor me mostró una breve carta al director, escrita por Jane Porter y Herschel Jick, de la Universidad de Boston, que el prestigioso *New England Journal of Medicine* publicó en 1980. En ella se afirmaba que, entre los casi doce mil pacientes hospitalizados a los que se había administrado opiáceos, «únicamente se produjeron cuatro casos de adicción razonablemente bien documentada en pacientes sin antecedentes de adicción».¹ Esa carta y el estudio que mencionaba se han citado desde entonces en numerosas ocasiones, lo que acabó provocando que se adoptara un enfoque indulgente de los opiáceos y que tales fármacos se prescribieran a discreción para tratar el dolor, partiendo de la base de que era poco probable volverse adicto a ellos.

Muchos médicos empezaron a recetarlos a la ligera y sin restricción alguna para consumo doméstico, ignorando tanto los defectos metodológicos del estudio (aunque se realizó con una impresionante muestra de 39 946 pacientes) como el hecho de que su conclusión solo era válida para individuos tratados con dosis bajas de opioides y estrechamente vigilados por el personal médico.

«En la actualidad, unos cuarenta años después de que se publicara esa carta —me contó Igor—, nos enfrentamos a un terrible problema de salud: solo en Estados Unidos hay más de dos millones de personas adictas a los opiáceos, muchas de las cuales son susceptibles de contraer diversas enfermedades y, además, presentan elevadas tasas de mortalidad. Tenemos unos 130 casos diarios de muerte por sobredosis, lo que supone cerca de 50 000 muertes anuales.» (Algunos expertos creen que las cifras son muy superiores.)

Herschel Jick afirmó en el 2017, en declaraciones a la Radio Pública Nacional de Estados Unidos, que, si hubiera sabido en 1980 lo que sabe hoy, no habría enviado esa carta. Las estadísticas tienen un enorme poder, sobre todo cuando los encargados de tomar decisiones importantes se apoyan en ellas. Por eso pueden ser peligrosas cuando se sacan de contexto.

INMERSOS EN LA ESTADÍSTICA

Lo cierto es que la mayoría de los investigadores de todos los campos de la ciencia llevan a cabo estudios estadísticos en algún momento. La estadística ha adquirido un estatus único entre las disciplinas académicas, ya que se requiere cierto pensamiento estadístico en casi todas las fases de la investigación: planificación del estudio, selección de muestras, gestión de datos e interpretación de los resultados. Los recientes paradigmas relacionados con la recopilación y el análisis de macrodatos están cambiando las reglas de casi todos los juegos.

Además de en la ciencia, la estadística está presente en los mercados financieros, el deporte y la publicidad, es decir, en casi todas partes. En el creciente campo de la inteligencia artificial, la estadística desempeña asimismo un papel importante, y todos sus expertos conocen bien los conceptos de probabilidad y estadística. Los grandes modelos lingüísticos que hay detrás de herramientas como ChatGPT se fundamentan en el análisis de grandes cantidades de datos, lo que implica el uso de métodos estadísticos.

Una vez, un amigo que estaba planificando unas vacaciones familiares me planteó una pregunta estadístico-probabilística en absoluto trivial: «Según TripAdvisor, ¿qué debería elegir, un hotel con una calificación de 98 basada en unas 300 opiniones o un hotel con una calificación de 95 basada en más de 1000 opiniones?». Evidentemente, 98 es algo mejor que 95, pero más de 1000 opiniones tienen un peso más fiable que 300. ¿Cómo resolver este dilema?

Si leer los datos anteriores te ha resultado agotador, ahora puedes tomarte un respiro. Párate a reflexionar sobre la pregunta de mi amigo: ¿qué hotel elegirías?

PERO ¿QUÉ ES LA ESTADÍSTICA?

He tenido la suerte de enseñar estadística a estudiantes de matemáticas e informática, pero también en escuelas de ingeniería, física, agricultura, trabajo social, ciencias del comportamiento, economía, biología y gestión empresarial. Que

yo sepa, no hay ninguna otra disciplina que se enseñe en todas esas escuelas. La estadística es importante por sí misma. La razón de ello es que se trata de un lenguaje especial, un requisito para comprender nuestro mundo moderno.

Así pues, ¿qué significa «estadística»? Sorprendentemente, no existe una respuesta definitiva.

En la bibliografía encontrarás diversas definiciones cuyo denominador común es que la estadística es **una ciencia que permite analizar datos para ayudarnos a extraer conclusiones**.

La siguiente definición es tal vez más precisa:

Estadística: Conjunto de términos, reglas y métodos utilizados para (1) recopilar datos, (2) analizarlos, (3) extraer conclusiones basadas en esos datos.

Y ahora, después de haber hablado tanto de la ciencia de la estadística, debo añadir que es la más extraña de todas las ciencias. Esta opinión se fundamenta en varias observaciones, de las que citaré solo dos:

1. **Idiosincrasia:** Es posible que dos expertos trabajen con los mismos datos, examinen los mismos resultados y, aun así, lleguen a conclusiones diferentes.
2. **Arrogancia** (la razón principal): A nadie se le ocurriría presumir de ser un experto en ecuaciones diferenciales, mecánica de líquidos o genética molecular si nunca ha estudiado a fondo estos campos científicos. No ocurre lo mismo con la estadística. La gente cree tener un buen conocimiento de este campo científico, aunque nunca lo haya estudiado en serio.

Incluso los hay que dicen cosas como: «Dejemos que las cifras hablen por sí solas». ¿En serio? Nunca he oído al número 7 pronunciar un discurso de motivación ni entablar una conversación sincera con el número 3. ¿Y tú?

Sin embargo, pese a que no puedan hablar directamente por sí mismos, los números tienen mucho que decirnos: descubramos cómo lo hacen.