



ÉRASE UNA VEZ UN NÚMERO

John Allen Paulos



METATEMAS

LIBROS PARA PENSAR LA CIENCIA

¿Hay algo más opuesto que los números y las historias de ficción? Los números son abstractos, seguros y eternos, pero a la mayoría de las personas les parecen fríos e inertes. Las buenas historias están llenas de vida: nos involucran emocionalmente, aportan matices..., pero carecen de rigor y sus posibles verdades siempre son escurridizas y objeto de polémica. Se diría que, para comprender el mundo que nos rodea, los números y las historias son métodos casi incompatibles.

«Érase una vez un número» nos descubre que las historias y los números no son tan diferentes como cabría imaginar y que en realidad están relacionados de un modo sorprendente y fascinante. Los conceptos de lógica y probabilidad surgieron de intuiciones sobre el funcionamiento de determinadas historias, y los lógicos de hoy en día se dedican a conjeturar posibles maneras de afrontar situaciones reales a partir de métodos matemáticos. Incluso la teoría de la complejidad enfoca las cadenas numéricas y las cadenas narrativas de manera parecida.

En las páginas de «Érase una vez un número», John Allen Paulos tiende un puente entre las dos culturas con su peculiar estilo ingenioso y desenfadado. Además de lúcidos resúmenes sobre la más actual teoría de la información, el lector encontrará por ejemplo, entre chistes y anécdotas divertidas, instrucciones para organizar el mastodóntico timo de la pirámide; una disparatada conversación entre Groucho Marx y Bertrand Russell o explicaciones de por qué es mucho más probable sentirse ofendido que ofensor.

Al más que astuto filósofo David Hume, quien escribió:

«Sólo puedo considerarme una especie de residente o embajador del reino del saber en el de la conversación, y creer que tengo la obligación continua de fomentar las buenas relaciones entre estos dos estados, que tanto dependen uno de otro».

Índice de contenido

Introducción

1. Entre las historias y las estadísticas
 - Atisbos primitivos
 - Las historias como contexto estadístico
 - Apunte para un cuento matemático
 - Ámbitos narrativos y ámbitos estadísticos
 - Demasiados rasgos, pocas personas
 - Estereotipos, fantasías y conservadurismo estadístico
2. Entre la subjetividad y la probabilidad impersonal
 - Puntos de vista minoritarios, individuos y estadísticas
 - La ley de Murphy y la sensación de que nos tratan mal
 - Psicología, perspectivas y paranoia
 - Claves bíblicas y escándalos sexuales
 - Cuento con truco y apareamiento probabilístico
 - Teorema de Bayes y revisión de las propias historias
 - Historias jurídicas complejas y redes de inferencia
 - O. J. Simpson y el estadisticidio
3. Entre el discurso informal y la lógica
 - La lógica informal y nosotros
 - Comodines en el póquer y en la vida
 - Reglas, sustitución y probabilidad
 - Lógica intensional y explosiones combinatorias
 - Los huesos pelados de la lógica estándar
 - Situaciones, semántica y estadísticas
 - El acervo narrativo
 - Parábola de las feministas despiadadas y el mercado de valores
 - Autobiografías, modelos y novelas
 - Apéndice: humor y cálculo
 - Groucho conoce a Russell
4. Entre el significado y la información

Información: perspicacia de un conferenciante, distracciones de una niña
Criptografía y narrativa
De Occam y las historias de dos dígitos
Nuestros horizontes de complejidad
Puede prepararse un menú gratuito si el vertedero (de Ramsey) es lo bastante grande
Historias, analogías y orden gratuito
Complejidad, caos y calderilla

5. Salvar el abismo

Loterías y deseos que se hacen realidad
Promesas, preguntas e intenciones implícitas
Dos culturas, un solo localismo
La encrucijada del médico
El medio ambiente y otras tierras de nadie
Unas palabras sobre (proto)religión

Bibliografía

Autor

Introducción

Sería exagerado afirmar que nuestra relación es hostil; yo vivo, yo me dejo vivir, para que Borges pueda tejer su literatura y esa literatura me justifica [...] No sé cuál de los dos escribe esta página.

Jorge Luis Borges

Al oír la risa ahogada, la joven se volvió para mirar a su tío, al que tanto había querido antaño. Le arrebató sin miramientos los papeles que llevaba en el bolsillo de la camisa hawaiana mientras el hombre retrocedía con nerviosismo hacia la puerta de la habitación del hotel. Ella, sin ocultar el asco que le producían los lloriqueos de su tío y su hipocresía, exclamó con voz silbante:

—El veintidós coma ocho por ciento de todas las suspensiones de pagos declaradas entre julio de 1995 y junio de 1997 se atribuyó a la mala asesoría jurídica, hasta el nueve coma dos por ciento por encima del bienio anterior.

—He hecho lo que he podido —balbuceó el hombre de ciento treinta kilos. Lo que menos quería en este mundo era aumentar el enfado de su irritada sobrina, que a pesar de su esbelta figura, sus cincuenta kilos y su cara de ángel era muy capaz de causar daños de consideración. Una vez a salvo en el pasillo, sin embargo, el tío se armó de valor y propuso:

—Un metaanálisis hecho sobre un abanico de estudios sugiere que menos del cuarenta por ciento de los casos de

impericia jurídica se debe a la mala intención, la contrapartida de la simple incompetencia.

Al oír aquello, la joven se lanzó sobre él y le dio un zarpazo en el grueso cuello con sus dedos fuertes y afilados, arrancándole la camisa de la espalda cubierta ya de sangre.

Como ilustra la situación descrita, las historias que contamos todos los días conviven en incómodo equilibrio con estadísticas de presunta pertinencia, incluso cuando no se contradicen. Nuestras historias están llenas de personas que hacen cosas por deseo, por miedo, y posiblemente por una pasión inmoderada por los macarrones. Las circunstancias y situaciones concretas ocupan un lugar destacado en todas las descripciones. Sin embargo, en estadística no suele haber agentes; sólo demografía, leyes generales, procesos. Los detalles y concreciones suelen desestimarse por irrelevantes.

El desencuentro entre historias y números abarca desde lo tópico (tomar una yuxtaposición por una relación causal) hasta lo abstruso. Un reciente e insólito caso de choque frontal entre nuestra sed de historias comprensibles y la atracción por las estadísticas impersonales es el fenómeno de las claves bíblicas. La manía empezó cuando Eliyahu Rips y otros dos matemáticos israelíes publicaron un artículo en una revista de estadística en el que al parecer sugerían que la Torá (el Pentateuco, los cinco primeros libros de la Biblia) contiene muchas «series de letras equidistantes» (en adelante SLE) que indican relaciones significativas entre personas, acontecimientos y fechas.

Una SLE es una sucesión de letras (hebreas en este caso) separadas entre sí por distancias iguales. Las palabras del texto se juntan y los espacios que las separan se pasan por alto. Así, la palabra *generalización* contiene una SLE que da «nazi» (geNerAliZaclón) con una separación fija de dos letras. Por lo general, las separaciones de las SLE son mayores y tras alguna inicial afectan, por ejemplo, a las le-

tras vigésimo tercera, cuadragésimo sexta, sexagésimo nona, nonagésimo segunda, etcétera.

El artículo decía que en la Torá solían aparecer juntas las SLE de los nombres de ciertos rabinos famosos que vivieron siglos después de los tiempos bíblicos y las SLE de sus fechas de nacimiento, y que la probabilidad de que esto ocurriera era mínima. El comité editorial de la revista enfocó la publicación del artículo como una especie de acertijo matemático: entre las muchas cosas susceptibles de tener escasa probabilidad, ¿cuáles la tienen realmente? No fue ésta, sin embargo, la lectura que tuvo. Algunos grupos se lanzaron sobre este «indicio», como ya habían hecho con anteriores hallazgos numerológicos cristianos e islámicos, y declararon que era una prueba de la inspiración divina de la Torá. *The Bible code*, el best-seller internacional de Michael Drosnin, fue más allá y alegó haber encontrado en la Torá un anuncio profético del asesinato de Itzhak Rabin y otros acontecimientos contemporáneos. Como tampoco aquí falta la sempiterna *Kennedy connection*, no es de extrañar que haya una SLE de «Kennedy» cerca de otra de «Dallas». Aunque de la solución de este enigma y de las matemáticas elementales que hay detrás de las claves bíblicas hablaré más adelante, lo que aquí interesa destacar es que nuestra sed de historias, agentes y motivos es tan intensa que muchos ven significado en series de letras fuera de contexto.

Los dos ejemplos que acabamos de ver ilustran sendas formas desaconsejables de tender un puente entre las historias y las estadísticas. El presente libro trata de formas más inteligentes de abordar y explorar el abismo que las separa. También comenta el acientífico tema de cómo compaginamos las historias y las estadísticas en nuestra vida; nuestro modo de compaginarlas contribuye a definir quiénes somos.

Casi todo el mundo ha visto esos carteles turísticos donde los puntos de interés locales aparecen en primer tér-

mino mientras el resto del mundo retrocede hasta perderse en el lejano horizonte. Nuestro mundo psicológico es igual de egocéntrico: los demás son el telón de fondo de nuestra vida y, aunque nos fastidie, nosotros somos el de los demás. ¿Es posible conciliar los carteles turísticos y las fantasías sobre nosotros mismos con los mapas exactos, las complejidades externas y un enfoque incorpóreo que no venga de ninguna parte?

La cuestión es, una vez más, hasta qué punto puede cerrarse o por lo menos aclararse el abismo lógico y psicológico que hay entre las historias y las estadísticas, y los abismos paralelos que median entre la subjetividad y la probabilidad impersonal, entre el discurso informal y la lógica, entre el significado y la información. Entre la literatura y la ciencia hay una complementariedad igual de incómoda. Los análisis literarios de perspectivas individuales, desarrollos posibles, arquetipos útiles y singularidades extrañas se avienen mal con la cháchara científica sobre objetividad, resultados definitivos, verdades universales y casos generales. Con pocas líneas de diferencia, palabras como «bendita suerte» y «milagro» conviven de mala manera con «azar» y «coincidencia».

¿Se puede salvar la brecha que hay entre las personas para quienes el mundo es o blanco o negro y las que lo ven desde el punto de vista del azar y los números, entre las humanidades y las ciencias, y, llevando las cosas a los extremos, entre los teóricos de la conspiración y los «seres anónimos»? En un mundo cada vez más interrelacionado, ¿merecen respeto las actitudes y perspectivas personales que no desdeñan la objetividad científica? En ese caso, ¿cómo explicamos el hecho de que casi todo el mundo se sienta personalmente ofendido y que casi nadie se tenga por ofensor?

¿Cómo se traza, por ejemplo, una imagen coherente que abarque a la vez los significados humanos y los fragmentos aislados de información? ¿De qué manera las histo-

rias (por ejemplo, la de la joven y su corpulento tío) y las estadísticas (por ejemplo, la incompetencia jurídica) llegan a resultar pertinentes y vinculantes? ¿Son los conceptos no estadísticos refinamientos y destilaciones de ideas sugeridas por historias y acontecimientos repetitivos? ¿Cuáles son las consecuencias narrativas de las ideas matemáticas de complejidad y «orden gratuito»? ¿Qué tienen en común las interpretaciones de la literatura y las aplicaciones de las estadísticas? ¿Qué tiene que ver la crítica literaria con la criptografía?

En los ensayos que siguen espero arrojar un indirecto pero penetrante rayo de luz sobre estos interrogantes. Algunas rarezas y ciertos problemas relacionados con la sima abierta entre las historias y las estadísticas (como tomar las anécdotas por muestras estadísticas o, al contrario, tomar las medias por descripciones de casos concretos) son resultado de aplicar una lógica apta para un dominio, en otro completamente distinto. A diferencia de la lógica que rige las matemáticas y las ciencias físicas, las verdades de la lógica informal y cotidiana dependen fuertemente del contexto y del individuo, que son aspectos insustituibles de toda situación. Proyectar condiciones específicas, por ejemplo, de un juego o una actividad, o de determinadas creencias religiosas, sobre el universo físico (o, por el contrario, derivar de las leyes de la física estrategias para los juegos y actividades, o para las creencias religiosas de una persona) es sólo un ejemplo de esta confusión de dominios. Lo mismo ocurre, con alguna pequeña diferencia, cuando se atribuye significado a las claves de la Biblia.

La relación entre lo personal y lo objetivo suele ser sutil. Nuestro modo de abordar la definición de problemas y temas afecta a su solución; es lo que sucede, por ejemplo, cuando se prefiere los números de lotería elegidos personalmente a los dispensados por máquinas (aunque todas las series numéricas tienen la misma probabilidad de ser elegidas por la administración). En un sentido más general,

nuestra unión inextricable por un saber y unas interpretaciones implícitas comunes señala interesantes extensiones de la práctica matemática estándar.

En los descansos entre episodios y parábolas hablaré de lógica alternativa, de probabilidad y estadística, de claves, códigos y teoría de la información, de filosofía de la ciencia y de teoría literaria, y lo utilizaré para describir las intrincadas conexiones entre esas dos formas básicas de relacionarnos con el mundo que son las historias y los números. Tender un puente sobre este abismo ha sido, de un modo u otro, una preocupación latente en mis libros anteriores. Creo que es un asunto que preocupa al 63,21 por ciento de la población.

1

Entre las historias y las estadísticas

—*¡Está a punto de surgir un isótopo de bismuto!*
—*exclamé atropelladamente, viendo cómo salían despedidos del crisol de una supernova los elementos recién nacidos.*

—*¡Y que lo digas!*

Italo Calvino

¿Historias y estadísticas? ¿A qué podría referirse esta yuxtaposición de términos? ¿A la cantidad de libros históricos que se publican al año? ¿A las anécdotas que cuentan los encuestadores? ¿A las biografías de Harris, Field, Gallup y Yankelovich? Si se insistiera, la mayoría diría seguramente algo desdeñoso, por ejemplo que las historias y las estadísticas se parecen tanto como un guisante a un elefante o, por no desperdiciar la rima, como una tortuga a una lechuga; el presente libro, sin embargo, la considera seriamente.

Uno de los supuestos de que parte es que las historias y el discurso informal han dado origen, con el tiempo, a dos modos complementarios de pensamiento que se emplean por lo general en estadística, lógica y matemáticas. Aunque estas disciplinas son quizá más difíciles de abordar y pueden incluso ir contra nuestras intuiciones, puede afirmarse que primero contamos historias y luego, en el parpadeo de un eón, hablamos de estadísticas.

Hay otros binomios «obstétricos» que se parecen un poco a éste: lo particular y lo general, lo subjetivo y lo universal, la intuición y la prueba, la emoción dramática y lo intemporal, la primera persona y la tercera, lo extraordinario y lo canónico. El primer elemento de cada binomio, aunque pudiera considerarse inferior, da origen al segundo o abona su terreno. Un sentimiento subjetivo es un preámbulo necesario para la valoración de la universalidad, y la inmersión emocional en el momento tiende poco a poco a la conciencia de lo intemporal.

Meditar estas oposiciones con naturalidad sugiere que el abismo que las separa es más cuestión de tradición, grado y terminología que de profundidades insondables. Creo que es así; y dado que la sima entre historias y estadísticas es una sinécdoque del más conocido abismo entre las dos culturas que distinguió C. P. Snow, la literaria y la científica, algunos temas que trato tienen un alcance mayor de lo que pueda pensarse cuando surjan. (En ocasiones emplearé el término estadística en un sentido muy general). *Sinécdoque* es un término literario que designa una figura de dicción en que se toma la parte por el todo, y viceversa, y esto es lo que ocurre cuando se toma una muestra por toda la población. Con este detalle de pedantería tenemos ya tendido el primer cordel sobre el abismo.

Atisbos primitivos

Las ideas de probabilidad y estadística no aparecieron de repente, con todo el aparato con que las vemos en los cursos de matemáticas. En la antigüedad ya se atisbaron los conceptos de media y variabilidad. Los huesos y las piedras se empleaban para jugar a los dados. En la literatura antigua hay referencias a la probabilidad. La importancia del

azar en la vida cotidiana se entendía claramente, por lo menos la entendían algunos. No es difícil imaginar pensamientos probabilísticos revoloteando en la mente de nuestros antepasados. (Con un poco de suerte habré vuelto antes de que se lo coman todo; es poco creíble que dejaran intacto el ganado y se llevaran su cosecha de bellotas; él suele exagerar su número de presas).

Las ideas de azar y probabilidad se formalizaron milenios después, cuando Pascal y Fermat las perfeccionaron para resolver ciertos problemas de juego allá por el siglo XVII. Laplace y Gauss, ciento cincuenta años más tarde, las desarrollaron y aplicaron a cuestiones científicas; y Quetelet y Durkheim las emplearon en el siglo XIX para comprender las regularidades de los fenómenos sociales. (Hay más posibilidades de sacar un seis tirando un solo dado cuatro veces que de sacar dos seises tirando dos dados veinticuatro veces; la probabilidad de que una partícula se desintegre en el próximo minuto es de 0,927; las encuestas realizadas en los colegios electorales revelan que cuatro de cada cinco ciudadanos a favor de la ley de control de armas dieron su voto a Gore).

Después de este rápido viaje por la historia de la estadística, permítaseme reducir la velocidad para señalar algunos de los muchos antepasados corrientes de las ideas más sobresalientes sobre probabilidad y estadística. Pensemos en primer lugar en las nociones de tendencia centralizadora: media, mediana, clase modal, etcétera. Lo más seguro es que surgieran de palabras cotidianas como *habitual*, *acostumbrado*, *típico*, *mismo*, *regular*, *mayoría*, *clásico*, *estereotipo*, *esperado*, *vulgar*, *normal*, *corriente*, *medio*, *convencional*, *tópico*, *mediano*. Cuesta imaginar a los prehistóricos, incluso a los que carecían del vocabulario descrito, sin algún barrunto de lo típico. Es de creer que fenómenos o seres como las tormentas, los animales y las piedras que se presentasen una y otra vez condujeran de manera natural a la idea de recurrencia típica o media.

Veamos también las precursoras de las ideas de variación estadística: desviación estándar, varianza, etcétera. Se trata de *insólito, peculiar, extraño, singular, original, extremo, especial, diferente, único, anormal, distinto, dispar, raro, demasiado*, etcétera. Una expresión como *fuera de lo común*, que indica algo extraordinario, viene muy al caso, porque una observación que está en la «cola» de la gráfica de una distribución estadística está fuera de lo habitual y señala una gran desviación de la magnitud en cuestión. Con el tiempo, cualquier situación o entidad que se repita sugerirá la idea de excepción. Si unos acontecimientos son corrientes, otros son raros.

La probabilidad está implícita en palabras como *casualidad, acaso, posibilidad, destino, dioses, hado, fortuna, suerte, coincidencia, azar* y muchas otras. Adviértase que la sola admisión de la idea de posibilidades alternativas, esencial para la narrativa, casi supone la idea de probabilidad; unos argumentos se considerarán más probables que otros. La necesidad de destacar aspectos de situaciones y entidades repetitivas conduce igualmente al concepto de muestreo, clave en estadística, y que se refleja en palabras y expresiones como *ejemplo, caso, representativo, observación, espécimen y muestra*. Del mismo modo, el proceso mental natural de asociar dos animales o cosas sugiere la importante idea de correlación, que tiene los siguientes correlativos (por así decirlo): *asociación, conexión, relación, vinculación, conjunción, conformidad, dependencia, proporción* y la siempre predispuesta *causa*.

Como R. P. Cuzzort y James Vrettos expusieron en *The elementary forms of statistical reason*, incluso nociones estadísticas menos conocidas como control, estandarización, comprobación de hipótesis, análisis bayesiano (la revisión de las estimaciones de probabilidad a la luz de nuevos datos) y categorización, se corresponden con ideas y locuciones de sentido común que son parte inseparable del saber humano y de la literatura. A semejanza de aquel personaje