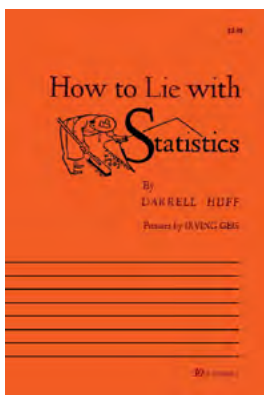


¿Estadísticas engañosas?

José Antonio **González Oreja**



HOW TO LIE WITH STATISTICS
DARREL HUFF
W. W. Norton & Company
New York, 1954

“Hay tres clases de mentiras: las mentiras, las malditas mentiras y las estadísticas.” En esta frase, atribuida generalmente al escritor, político y aristócrata británico del siglo XIX Benjamín Disraeli, podemos encontrar el origen del título que voy a reseñar: *How to Lie with Statistics* de Durrell Huff (1954). A pesar de que fue publicado hace ya más de medio siglo, y de que el contexto socioeconómico y político en el que se presentan los ejemplos ha cambiado notablemente desde entonces, la aplicación de las observaciones metodológicas realizadas por el autor es igual de válida en la actualidad. Y es que el libro es una especie de manual de cómo pueden usarse las estadísticas por los medios de comunicación (y por otros agentes interesados) para distorsionar la realidad, torturar a los datos (según la expresión de Olden y Jackson, 2000) y engañar así al lector general, quien normalmente no es experto en esta materia. En tan solo diez capítulos, escritos desde un

punto de vista crítico con la práctica de la estadística, y con un fino sentido del humor, Durrel Huff analiza en *How to Lie with Statistics* cómo se pueden manipular, de modo consciente o no, los análisis a los que se someten los datos numéricos para revelar un supuesto patrón oculto en ellos, una “verdad” que no está ahí; una mentira al fin y al cabo.

Los cuatro primeros capítulos cubren temas que se tratan también en otros libros sobre estadística. Pero muchos de quienes en algún momento nos hemos ocupado de la docencia de cursos de introducción a la estadística, o de análisis de datos, nos sentiríamos satisfechos si nuestros estudiantes dominasen de verdad el contenido de estos sencillos capítulos. El capítulo 1 arranca pasando revista a los métodos de muestreo incorrectos, que sesgan la representatividad de la muestra con respecto a la población de la que proceden y que hacen que las conclusiones obtenidas sean difícilmente generalizables. Es el caso de los conjuntos formados por individuos autoseleccionados, propuestos a sí mismos como candidatos a tener en cuenta. El capítulo 2 nos recuerda que las diferentes expresiones que permiten calcular una media (es decir, una medida de la tendencia central de una serie de datos, ya sea la media aritmética, la media geométrica, la mediana o la moda) tienen aplicaciones distintas. El utilizar una u otra como si fueran equivalentes, o el no saber cuándo debe utilizarse cada una de ellas, puede engañar al lector general. El capítulo 3 presta atención a los números que deberían estar presentes para comprender en su totalidad las estadísticas... pero que no siempre lo están. Uno de estos números es el tamaño de la muestra (n), pues su representatividad está relacionada de modo inverso con n : podemos tener una mayor certeza en que las características de la muestra se aproximarán más a las de la población general de la que se ha extraído cuando la muestra está formada por un elevado número de elementos (n grande) que cuando n sea pequeño, aunque el método de muestreo empleado sea el mismo. Podemos añadir que la falta de representatividad de la muestra con respecto a la población general será máxima cuando su tamaño sea el mínimo ($n=1$), como cuando un resultado estadístico o producto comercial (por ejemplo, una



© Nin Solis. Nuevo Valle de Moreno, Guanajuato, 2013.

dieta, un nuevo medicamento o lo que sea) viene “avalado” por determinada entidad o individuo (Paulos, 1988). Otro de los números que no siempre están presentes, y que permite entender la importancia de las conclusiones emanadas de ciertos análisis estadísticos, es el valor de la P ; o sea, la significación estadística del análisis realizado a partir de los datos. En realidad no siempre es fácil comprender el significado técnico del valor de la P (y se han publicado algunos libros cuyos títulos apuntan en este sentido; Moyé, 2006; Vickers, 2010), pero el lector instruido en estadística puede valorar por sí mismo si un resultado es relevante como para merecer su atención o se trata tan solo de un resultado más que entra dentro de lo que cabe esperar bajo una hipótesis nula concreta. Como dice el autor, lo que resulta engañoso del “numerito” que debería estar presente pero no lo está es que, en general, su ausencia o eliminación pasa desapercibida, y en ello radica el secreto del éxito de este engaño. La importancia que tiene el tamaño del error asociado a una medida cualquiera es analizada en el capítulo 4. Es el caso de los estimadores puntuales *versus* estimadores por intervalo: para muestras de un tamaño n , cuanto más estrecho es el intervalo (es decir, cuanto más precisa es la estimación), menos fiable resulta, y viceversa (Paulos, 1988).

Los siguientes tres capítulos abordan, entre otros, puntos relacionados con el mal uso de las representaciones gráficas de la información estadística. El capítulo 5 presta atención al uso falaz de figuras y gráficas. Alguno de los “trucos” utilizados por quienes así

tergiversan los datos consiste, por ejemplo, en manipular los ejes que indican los valores de las variables que se muestran en la gráfica, de modo que las diferencias entre las variables, o las tendencias a lo largo de un eje, se distorsionen respecto a la realidad. El capítulo 6 considera los gráficos pictóricos, o pictogramas, como herramientas que transmiten fácilmente información errónea; una cuestión, básicamente, de pérdida de proporciones, ya que la variable que se representa es unidimensional pero la imagen que la simboliza en el pictograma no lo es. Peor aún, los pictogramas se usan muchas veces de modo incorrecto porque se quiere engañar de modo consciente a un lector inconsciente del engaño. A este respecto considero que sigue siendo útil el trabajo de Wainer (1984), escrito con el mismo espíritu crítico y burlón que el libro de Huff (1954), pues ambos nos muestran algunas de las manipulaciones que podemos encontrar en este campo. Un par de libros más modernos que cubren estos puntos son los de Willis (2004), escrito como una introducción, y Tufte (2001), quizás la guía definitiva a seguir para representar visualmente información cuantitativa. El capítulo 7 analiza las relaciones aparentes que se dan entre algunas variables que solo están relacionadas de modo marginal (y eso, con suerte), y que hacen posible una vez más que el lector incauto resulte burlado. Es el caso de las correlaciones entre pares de variables (por ejemplo, cuando la variable A aumenta o disminuye, la variable B también aumenta o disminuye), que pueden hacerse pasar como relaciones causales (por ejemplo,

cuando la variable A aumenta o disminuye, esto hace que la variable B también aumente o disminuya; la pequeña diferencia con la frase anterior es de importancia capital). Y es que, en realidad, *correlation is not causation* (Havens, 1999). Yo mismo me entretuve hace años en recopilar algunos ejemplos de relaciones aparentes y correlaciones sin sentido (González Oreja, 1999), relaciones que han documentado bien muy diversos autores. La posible dependencia de las dos variables analizadas (A y B) de una tercera que no se tiene en cuenta (C) invita a prestar una mayor atención a los resultados obtenidos con base en un estudio observacional. Por ejemplo: (a) existe una elevada correlación positiva y significativa entre (A) las ventas anuales de chicle y (B) la incidencia del crimen en los Estados Unidos de América; ahora bien, no es lícito concluir que prohibiendo la venta de chicle podría reducirse la incidencia del crimen, pues ambas variables dependen de una tercera (C): el tamaño de la población analizada. (b) Existe también una elevada correlación positiva y significativa entre (A) los salarios de los profesores y (B) las ventas de licores alcohólicos; en realidad, carece de sentido concluir que, dado que la correlación estadística existe, debe existir también una relación del tipo causa-efecto entre las variables analizadas. Y (c) se ha documentado una correlación positiva y significativa entre el índice de divorcios de un país y sus importaciones de plátanos.

El problema de la conexión causal (*post hoc ergo propter hoc*) es tratado también en el capítulo 8: el hecho de que el acontecimiento B tenga lugar después de A no implica –necesariamente– que A haya causado B. De hecho, la ausencia de relaciones lógicas o evidencia empírica que ligue al suceso A con el B puede degenerar en una especie de pensamiento mágico o supersticioso, una falacia más del razonamiento (Ennett, 2012). El capítulo 9 recoge algunas otras formas en las que es posible utilizar las estadísticas para transmitir prácticamente la información “que uno quiera”. Y los ejemplos en este capítulo me resultan especialmente ilustrativos. Si el producto 1 ha pasado de costar 20 unidades (ya sean pesos, dólares, euros o cualquier otra moneda) el año anterior a solo 10 este año, pero el producto 2 ha pasado de costar 5 el año pasado a 10 este año,

© Nin Solis. Santo Domingo, Guerrero, 2013.



¿cabe concluir que la cesta de la compra formada por ambos productos, o el coste de la vida, ha subido?; ¿o que ha bajado; o que se ha mantenido estable? En realidad, podemos *demostrar* cualquiera de estas tres posibilidades (pp. 118-119) cambiando tan solo el período que consideremos como referencia, o utilizando una u otra medida de tendencia central.

El último capítulo del libro abandona el tono irónico con el que se escribieron los demás y se enfoca en cómo distinguir *el trigo de la paja*, si se me permite la expresión; es decir, en cómo reconocer los datos valiosos y útiles de entre toda la cantidad y variedad de falsedades tras las que puede ocultarse la información estadística. En este capítulo en especial, el autor avanza a mediados del siglo XX en lo que se llamaría tiempo después pensamiento crítico, tan de moda a inicios del siglo XXI (Halpern, 2003; Jenkins, 2015). Así, presenta cinco preguntas con las que podemos rebatir a un estadístico y evitar prestar atención a algo (o a alguien) que no merece nuestro tiempo.

1. ¿Quién dice que las cosas sean así? Pues es más que probable que haya sesgos ocultos, conscientes o no, con respecto al autor de la afirmación estadística que nos ocupa. Y, aunque no haya tales sesgos, el simple hecho de que provengan de determinada fuente no garantiza su veracidad (el argumento de autoridad, otra falacia; Ennett, 2012).

2. ¿Cómo se ha llegado a tal conclusión? O sea, un análisis crítico de los métodos empleados, desde la etapa inicial del muestreo, pasando por el tamaño (n) de la propia muestra, hasta la pertinencia de los análisis estadísticos de los datos.

3. ¿Qué información falta? Como los pequeños números que deberían acompañar a las estadísticas. Y es que en la información que falta¹ puede estar la clave para entender si las conclusiones estadísticas emanadas de los datos son relevantes o no.

4. ¿Qué ha cambiado entre los datos crudos y las conclusiones extraídas de los análisis realizados? Pues no siempre las conclusiones están soportadas por los datos, aunque es más que frecuente que así se nos presente.

5. ¿Tiene sentido? Como el uso de números que se expresan con una precisión exquisita en cuanto a las cifras significativas, pero que son innecesarias en muchos casos, cuando no completamente increíbles.² O, también, como la extrapolación de relaciones entre variables, que pueden estar bien fundamentadas dentro de cierto rango, pero cuyo comportamiento fuera de los valores extremos desconocemos; extrapolar por debajo del mínimo o por encima del máximo en los que la relación es válida puede llevar a obtener conclusiones sin sentido, especialmente cuando la extrapolación se hace hacia el tiempo futuro. En las extrapolaciones hacia el futuro, una práctica habitual que se sitúa en la base de los ejercicios de predicción, la premisa implícita es que las tendencias encontradas hasta el presente serán igualmente válidas en el futuro; ahora bien, afortunadamente o no, las cosas en el futuro cambian y a menudo ya no son como cabría esperar. De lo contrario, termina su libro Huff (1954: 140), la vida sería aburrida.

Desde su publicación, el libro de Darrell Huff se ha convertido en un auténtico *bestseller* en la materia, pues ha acumulado más de medio millón de ejemplares vendidos tan solo en la edición en inglés y se ha traducido a numerosos idiomas. De hecho, no hay otro libro “de estadística” que haya vendido más copias que el pequeño volumen que aquí reseño. Lo que no está nada mal para alguien que no tenía una educación formal al respecto, pero que sabía “leer entre las líneas de la estadística” (Steele 2005). Y que, sobre todo, sabía escribir bien.

© Nin Solis. Cruztón, Chiapas, 2013.





© Nin Solis. Cruzón, Chiapas, 2013.

NOTAS

¹ Recuerdo ahora a cierta presentadora de “Un, dos, tres... responde otra vez”, aquel famoso programa-concurso de Televisión Española que solía revelar solo una parte de la información que los concursantes le requerían: Yo nunca miento, aclaraba, pero tampoco digo toda la verdad.

² En un ejercicio de evaluación, uno de mis estudiantes calculó hace poco la *lifetime dispersal distance* (LDD), i.e., la media aritmética de 12 distancias de dispersión de animales (valores enteros dados en metros), y expresó su resultado así: $LDD=317.9166667$ m. Estoy seguro de que no presentó más decimales solo porque la calculadora no se los mostró! Obviamente, era suficiente con 317.9 m; o, redondeando, 318 m. En el mismo ejercicio, otra alumna (que había llegado de modo correcto al valor de LDD) calculó la *lifetime dispersal area* (LDA), i.e., el área de dispersión vital correspondiente (teóricamente, el área de un círculo de radio LDD), y ni corta ni perezosa llegó al siguiente resultado: $LDA=1.2$ ha. ¡Poco más de 1 ha para un círculo de más de 300 m de radio!

REFERENCIAS

Ennet B. 2012. *Logically Fallacious. The Ultimate Collection of Over 300 Logical Fallacies*. eBookIt, Boston.

González Oreja JA. 1999. *Ecología de la Recuperación de la Ría de Bilbao*. Tesis Doctoral. Departamento de Zoología y Dinámica Celular Animal, Facultad de Ciencias. Universidad del País Vasco, Leioa. 287 pp.

Halpern DF. 2003. *Thought and Knowledge. An Introduction to Critical Thinking*. Lawrence Erlbaum, New Jersey. 467 pp.

Havens KE. 1999. Correlation is not causation: a case study of fisheries, trophic state and acidity in Florida (USA) lakes. *Environmental Pollution* 106: 1-4.

Huff D. 1954. *How to Lie with Statistics*. W.W. Norton & Company, New York. 142 pp.

Jenkins SH. 2015. *Tools for Critical Thinking in Biology*. Oxford University Press, Oxford. 324 pp.

Moyé LA. 2006. *Statistical Reasoning in Medicine. The Intuitive P-Value Primer*. Second Edition. Springer, New York. 301 pp.

Olden JD, Jackson DA. 2000. Torturing data for the sake of generality: How valid are our regression models? *Ecoscience* 7(4): 501-510.

Paulos JA. 1988. *Innumeracy. Mathematical Illiteracy and Its Consequences*. Vintage Books, New York. 208 pp.

Steele JM. 2005. Darrell Huff and fifty years of *How to Lie with Statistics*. *Statistical Science* 20(3): 205-209.

Tufte E. 2001. *The Visual Display of Quantitative Information*. Second Edition. Graphics Press, Cheshire. 200 pp.

Vickers A. 2010. *What is a P-Value, Anyway? 34 Stories to help you actually understand statistics*. Addison-Wesley, Boston. 212 pp.

Wainer H. 1984. How to display data badly. *The American Statistician* 38(2): 137-147.

Willis J. 2004. *Data Analysis and Presentation Skills. An Introduction for the Life and Medical Sciences*. Wiley, Chichester. 183 pp.

José Antonio González Oreja
Escuela de Biología, BUAP
jgonzorj@hotmail.com



© Nin Solis. San Cristóbal de las Casas, Chiapas, 2013.