Autor: Darrell Huff

Ilustrado por: Irving Geis



DARRELL HUFF



Ilustración: IRVING GEIS

ISBN 0-393-31072-8 1954



Norton & Company Inc

El texto de este libro ha sido bajado de: http://www.fce.unl.edu.ar/catedras/backend/materiales/731.pdf

ÍNDICE

Capítulo	TEMA	Página
	Introducción	1
1	La muestra que presenta un factor de influencia en sí misma	3
2	El promedio bien escogido	13
3	Las pequeñas cifras que no aparecen	19
4	Mucho ruido y pocas nueces	29
5	El gráfico exclamativo	33
6	El personaje de la gráfica	37
7	La cifra indirectamente relacionada	43
8	El "post hoc" aparece de nuevo	51
9	Cómo estadistiquear	59
10	Cómo enfrentarse con las estadísticas	73
	Glosario	85

Darrell Huff

INTRODUCCIÓN

-Hay mucho crimen por allí – dijo mi suegro al poco tiempo de haberse trasladado de Iowa a California. Y así era, en el periódico que leía. Es un periódico que no pasa por alto ningún crimen en su propia área y que tiene fama de prestar más atención a un asesinato en Iowa que el mismo periódico principal de aquella localidad.

Desde el punto de vista estadístico, la conclusión de mi suegro no poseía fundamento técnico. Se basaba en una muestra marcadamente subjetiva. Al igual que muchas otras estadísticas aún más falseadas, la estadística de mi suegro era tendenciosa. Daba por supuesto que el espacio del periódico dedicado a informar sobre crímenes era la medida del porcentaje de criminalidad. Hace algunos inviernos, doce investigadores informaron independientemente y sometieron cifras sobre las píldoras antihistamínicas. Todos ellos demostraron que un porcentaje considerable de resfriados desaparecía después del tratamiento. Siguió un gran alboroto, al menos en los anuncios, y se consiguió el lanzamiento de un producto farmacéutico de los que hacen época en el campo comercial. Este éxito se basaba en la eterna esperanza, siempre a punto de renacer, así como en el curioso desdén por examinar más allá de las estadísticas un hecho que se conoce desde hace largo tiempo.

Como señalaba, no hace mucho tiempo, Henry G. Felsen, humorista y no precisamente una autoridad en la medicina, un tratamiento adecuado cura un resfriado en siete días, pero si se deja que se resuelva por sus propios medios, se arrastra durante una semana.

Lo mismo ocurre con muchas de las cosas que usted lee y oye. Los promedios y las relaciones, las tendencias y los gráficos no son siempre lo que parecen. Puede haber más de lo que ven los ojos y puede haber mucho menos.

El lenguaje secreto de las estadísticas, tan atrayente a una cultura que se apoya en los hechos, se emplea para causar sensación, deformar, confundir y simplificar en demasía. Los métodos y los términos estadísticos son necesarios para informar sobre los datos masivos de las tendencias sociales y económicas, las situaciones de los negocios, las encuestas de "opinión" y los censos; pero sin escritores que utilicen las palabras con honradez y precisión y sin lectores que sepan lo que significan, el resultado no es más que pura semántica sin sentido alguno.

En la literatura popular sobre temas científicos, la estadística falseada ocupa un lugar destacado, a cargo de héroes de bata blanca que trabajan a deshoras en un laboratorio pobremente iluminado. Al igual que "un toque de brocha o un poco de carmín", las estadísticas desvirtúan los hechos importantes. Una estadística bien presentada es mejor que la "gran mentira" de Hitler; desorienta, pero no debe convencerle.

Este libro es una especie de "rudimentos" sobre la manera de utilizar las estadísticas para engañar. En conjunto puede parecer un manual para desaprensivos. Quizá pueda justificarlo como aquel ladrón retirado que con la publicación de sus recuerdos ofrecía un curso para graduarse en el arte de tirar la piedra y esconder la mano: los desaprensivos ya conocen estos trucos; los hombres honrados deben aprenderlos en defensa propia.

LA MUESTRA QUE PRESENTA UN FACTOR DE INFLUENCIA EN SÍ MISMA



"El promedio de los componentes de la promoción 1924 de la Universidad de Yale", informaba la revista Time, en un comentario de un artículo aparecido en el Sun de Nueva York, "gana 25.111 dólares al año". ¡Magnífico!

Pero, espere un momento. ¿Qué significa esta cifra impresionante? ¿Según parece, es prueba de que si envía a su hijo a Yale, no tendrá usted que trabajar en su vejez ni él tampoco?

Dos cosas referentes a esta cifra saltan a la vista a la primera ojeada recelosa. Es de una sorprendente precisión y tiene pocas probabilidades de ser exacta.

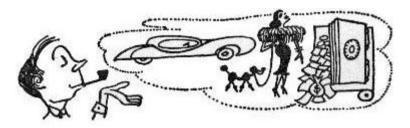
Es extremadamente difícil llegar a conocer hasta el último dólar los ingresos medios de cualquier grupo que se formó hace tiempo. Tampoco es muy probable que usted recuerde con tal precisión sus propios ingresos del año pasado, a menos de que éstos provengan enteramente de un salario; pocas veces, los ingresos del orden de 25.000 dólares provienen totalmente de un salario; la gente que se encuentra en este nivel disfruta, probablemente, de inversiones bien distribuidas.

Además, este atrayente promedio está calculado, sin duda, partiendo de las cantidades que DIJERON los graduados de Yale.

Incluso si en el año 24 figuraban todos en el cuadro de honor por su integridad, no podemos estar seguros de que todos ellos continúen siendo honestos al cabo de un cuarto de siglo y que estos informes sean todos fidedignos. Algunas personas, cuando se les pregunta acerca de sus ingresos, exageran por vanidad o por optimismo. Otros los minimizan, especialmente cuando se trata de las declaraciones a Hacienda, y después de hacerlo, temen contradecirse en cualquier otro documento. ¿Quién sabe hasta dónde pueden llegar los inspectores? Es posible que estas dos tendencias, la de sobre-

Darrell Huff

estimar y la de subestimar, se neutralicen, pero no es probable.



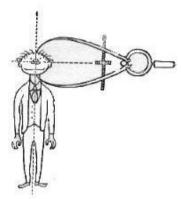
Quizá una tendencia sea más fuerte que la otra, pero ¿cuál de las dos? Empezamos a explicarnos una cifra de la cual nos dice el sentido común que difícilmente puede representar la verdad. Ahora pongamos el dedo en el probable origen del error más patente, origen que puede hacer figurar la cifra de 25.111 dólares como "ingresos medios" de algunos hombres cuya media real podría muy bien hallarse hacia la mitad de esta cantidad. Se trata de la técnica del muestreo, que es la espina dorsal de la mayor parte de las estadísticas que usted haya de encontrar en toda clase de temas. La base de esta técnica es simple, aunque en la práctica sus sutilezas han conducido a toda clase de caminos equívocos, algunos de los cuales distan de ser respetables. Si se tiene un saco de alubias, rojas y blancas, existe sólo una manera de averiguar exactamente cuántas hay de cada color: ¡contarlas! Sin embargo, puede averiguarse con aproximación la cantidad de alubias rojas de una forma más fácil; se coge un puñado de alubias, se cuentan las que lo componen y se da por sentado que la proporción será la misma en todo el caso. Si la muestra es lo bastante extensa y bien seleccionada, representará al conjunto con bastante aproximación en la mayoría de los casos. Si no es así, puede ser menos precisa que una hipótesis sensata, y no tiene nada que la apoye, a no ser un aire aparente de precisión científica. Es una triste verdad el hecho de que detrás de lo que leemos o creemos saber están las conclusiones derivadas de muestras subjetivas y demasiado pequeñas, o ambas cosas a la vez.



El informe sobre los universitarios de Yale deriva de una muestra. Podemos estar seguros de ello, porque la razón nos dice que nadie puede ponernos en contacto con todos los miembros vivientes de la promoción de 1924.

Al cabo de veinticinco años, debe haber muchos cuyas direcciones se desconozcan. Y, entre los de direcciones conocidas, muchos no contestarán al cuestionario, sobre todo por tratarse de un cuestionario de carácter personal. ... continúa...





Espero que usted no sea un *snob*, y yo no cuido de la compra y venta de propiedades. Pero, pongamos por caso, que usted lo es y que éste es un trabajo, y que usted está buscando una finca para comprar, en una carretera no lejana al valle de California donde yo vivo.

Descubierto su punto flaco, procuro hacerle saber que los ingresos medios del vecindario son de 15.000 dólares anuales. Puede que esto determine su decisión de vivir allí; sea como sea, usted compra la finca y esta hermosa cifra se graba en su cerebro. Es muy posible, ya que según hemos acordado para la finalidad de este relato es usted un poco snob, que mencione esta cifra cuando explique a sus amigos dónde vive.

Al cabo de un año o poco más, nos encontramos de nuevo. Como miembro de un comité de contribuyentes, redacto y distribuyo una instancia pidiendo que se rebajen los impuestos, o las contribuciones, o la tarifa de autobuses. Alego que no podemos hacer frente al aumento; después de todo, los ingresos medios del vecindario son solamente de 3.500 dólares al año.

Quizás usted se una a mi comité en esto — usted no solamente es snob, también es tacaño —, pero no deja de sorprenderse al oír la menguada suma de 3.500 dólares. ¿Miento ahora, o mentía el año pasado?

No puede culparme ninguna de las dos veces. Esta es la ventaja esencial de mentir con estadísticas. Ambas cifras son promedios legítimos, que se han obtenido de forma lícita. Ambas representan los mismos datos, los mismos individuos, los mismos ingresos. A pesar de todo, se hace patente que al menos una de ellas es tan engañosa que se aproxima a una mentira manifiesta. Mi truco consistió en utilizar un tipo de promedio distinto en las dos ocasiones, ya que la palabra "promedio" tiene un significado muy vago.

Es un truco utilizado con mucha frecuencia, a veces de forma inocente, otras con intención culpable, por individuos que desean influenciar la opinión pública o vender espacio publicitario. Cuando le dicen que algo es un promedio, usted no sabe todavía

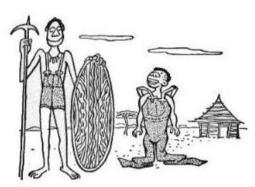
Darrell Huff

gran cosa del mismo, a menos de que pueda averiguar de qué promedio se trata: media, mediana o moda.

La cifra de 15.000 dólares que utilicé cuando deseaba manejar una cifra elevada, es una media aritmética de los ingresos de todas las familias del vecindario. Se obtiene sumando todos los ingresos y dividiendo la suma por el número de éstos. La otra cifra, más baja, es la mediana, la cual nos dice que la mitad de las familias en cuestión tienen unos ingresos superiores a los 3.500 dólares al año y la otra mitad disfruta de unos ingresos inferiores a esta cifra. Podría haber utilizado también la moda, o cifra que aparece mayor número de veces en una distribución. Si en este vecindario existe mayor número de familias con ingresos del orden de 5.000 dólares al año, en comparación con cualquier otra cantidad, el ingreso modal es de 5.000 dólares al año.

En este caso, como ocurre generalmente cuando se trata de cifras de ingresos, el "promedio" no especificado no representa nada. Otro factor de confusión es el hecho de que, en algunos tipos de información, todos los promedios son tan semejantes entre sí, que para la información no resulta de vital importancia hacer distinciones entre ellos.

Si usted lee que la estatura media de los hombres de alguna tribu primitiva es solamente de metro y medio, dispone de una información bastante buena acerca



de la estatura de aquella gente. No necesitará preguntar si se trata de una media, una mediana o una moda; vendrá a ser lo mismo (naturalmente, si posee una fábrica de prendas de trabajo para los africanos, necesitará una información más amplia que la obtenida con cualquier promedio. Esto guarda relación con las clases y las desviaciones, punto que trataremos en el próximo capítulo).

Los distintos promedios coinciden muy aproximadamente en un punto cuando se manejan datos relacionados con muchas de las características humanas; estos datos tienen la propiedad de estar dentro de lo que se llama la distribución normal. Si se traza una curva para representarla, se obtiene algo parecido a una campana, y la media, la mediana y la moda coinciden en el mismo punto.

Por consiguiente, cualquier promedio es tan bueno como otro para describir la estatura de los hombres; pero no lo es para describir sus carteras. Si se hiciera una lista de los ingresos anuales de todas las familias de una ciudad determinada podría resultar que fluctuasen desde cantidades muy pequeñas hasta 50.000.... continúa...

CÓMO MENTIR CON ESTADÍSTICAS Darrell Huff

CAPÍTULO 3

LAS PEQUEÑAS CIFRAS QUE NO APARECEN

Los titulares dicen que entre los usuarios de la pasta de dientes Doakes la frecuencia de caries dentales disminuye en un 23 %, y continúan informando de que usted puede vivir con un 23 % menos de dolores de muelas. Estos resultados, según podrá averiguar, proceden de un laboratorio "independiente" y prestigioso, y van acompañados de certificación expedida por un perito diplomado. ¿Qué más quiere?

A pesar de ello, si usted no es excesivamente crédulo u optimista, recordará por experiencia que una pasta de dientes raras veces es mucho mejor que cualquier otra. Entonces, ¿cómo consigue Doakes informar de tales resultados? ¿Puede contar mentiras de tal magnitud? No, ni hace falta que lo haga. Existen maneras más fáciles y efectivas de salirse con la suya.

El comodín principal en este caso es la muestra inadecuada – es decir, estadísticamente inadecuada –; aunque para la finalidad de la Doakes sea del todo satisfactoria. El grupo "test" de usuarios, según descubre usted posteriormente al leer la letra menuda, era solamente de una docena de personas. (Hay que dar las gracias a la Doakes en este caso, por la oportunidad que representa esta información.)

Algunos anunciantes la omitirían y publicarían la engañosa estadística sin indicar de qué clase de trapacería se trataba. Su muestra de una docena no está mal del todo, tal como se hacen generalmente estas cosas. Un producto que se llama Dentífrico del doctor Cornish apareció hace unos cuantos años en el mercado con la pretensión de haber conseguido "un éxito considerable en la corrección de... caries dentales".

La novedad del dentífrico era que contenía urea. Unas investigaciones de laboratorio parecían haber demostrado su valor para tal aplicación. El supuesto era desatinado porque el trabajo experimental había sido preliminar y se llevó a cabo sólo con seis casos.

Pero volvamos a examinar cómo la Doakes puede conseguir fácilmente un titular sin que contenga mentira alguna, y además una certificación de lo afirmado. Hagamos que cualquier pequeño grupo de personas controle sus caries durante seis meses, después sometámoslo al dentífrico de la Doakes. Pueden ocurrir tres cosas: que presente un número mayor de caries, que presente un número inferior o que presente el mismo número aproximadamente. Si se da la primera o la última de las posibilidades, la Doakes and Company archiva las cifras (fuera de todo alcance) y vuelve a probar.

Tarde o temprano, por efecto del azar, un grupo experimental presentará una gran mejora, digna de un titular y quizá de toda una campaña publicitaria. Esto ocurrirá tanto si usan el dentífrico Doakes como si emplean perborato sódico o el mismo dentífrico que venían utilizando.

La razón de utilizar un grupo experimental pequeño es la siguiente: en un grupo extenso, las diferencias que se producen por azar serán pequeñas e indignas de ser publicadas en grandes titulares.

Darrell Huff



Dar a conocer una mejora de un 2 por 100 no ayudará a vender mucha pasta dentífrica. Usted puede comprobar por sí mismo, y con un gasto muy pequeño, cómo pueden obtenerse por azar resultados que no indican nada — partiendo siempre de un número reducido de casos —. Eche una moneda al aire. ¿Cuántas veces saldrá cara? La mitad de las veces, naturalmente. Todo el mundo lo sabe.



Muy bien, comprobémoslo y veamos... He probado diez veces y he conseguido ocho veces cara: demuestro que con las monedas sale cara el 80 por 100 de las veces. Bien; para las estadísticas del dentífrico esto va sobre ruedas. Ahora, pruébelo. Puede que consiga un resultado del 50 por 100 pero probablemente no será así; su resultado, igual que el mío, corre el riesgo de diferir bastante del 50 por 100. Pero si su paciencia aguanta hasta las mil tiradas podrá estar casi seguro (aunque no completamente seguro) de alcanzar un resultado de aproximadamente la mitad de "caras", resultado que representará la probabilidad real. Sólo cuando se realiza un número importante de ensayos, el cálculo de probabilidades proporciona una predicción útil. ¿Cuántas tentativas son necesarias? También esto tiene su intríngulis. Depende, entre otras cosas, de la extensión y la variedad de la población que se estudie por muestreo. Y a veces, la magnitud de la muestra no es lo que parece.

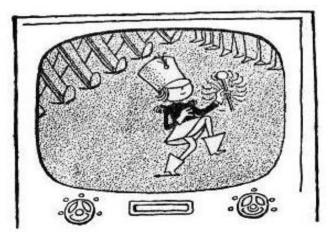
Relacionado con lo que acabamos de decir, encontramos un... continúa...

MUCHO RUIDO Y POCAS NUECES

Si no le importa, empezaré atribuyéndole dos hijos. A Pedro y a Linda (podríamos haberles puesto nombres que estuviesen más de moda) les han pasado un test de inteligencia, como se hace con otros muchos niños en el transcurso de sus años escolares. Los *tests* mentales de cualquier tipo son uno de los principales fetiches mágicos de nuestro tiempo; por consiguiente, tendrá que insistir un poco para que le informen de los resultados de los *tests*. Esta información es tan esotérica, que a menudo se cree que solamente está segura en manos de psicólogos y educadores, y quizá sea verdad. Sea como sea, usted se entera de que el C. I. (Cociente Intelectual) de Pedro es de 98 y el de Linda de 101. Usted sabe, naturalmente, que el C. I. se basa sobre 100 como promedio o "normal".

¡Ah! Linda es su vástago más listo. Además, está por encima de la media. Pedro está por debajo de la media, pero vale más que no hablemos mucho de ello. Tales conclusiones son simples tonterías.

Para aclarar la situación, observemos el hecho de que lo medido por un test de inteligencia no es exactamente lo mismo que lo designado generalmente con la palabra inteligencia. Olvida factores importantes como la capacidad de mando y la imaginación creadora. No tiene en consideración el factor social o la aptitud musical, artística, etc., sin mencionar factores de la personalidad tales como la voluntad y la integración emocional. Es más, los *tests* aplicados con mayor frecuencia en la escuela son los de tipo económico, de grupo, que dependen en gran parte de la facilidad de lectura; tanto si es listo como si no lo es, el que no lea con gran rapidez pierde toda probabilidad de éxito.



Pongamos por caso que reconocemos todo esto y estamos de acuerdo en que el C. I. es simplemente la medición de una aptitud, vagamente definida, para el manejo de abs-

Darrell Huff

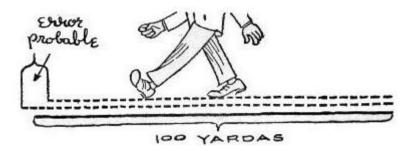
tracciones en "conserva", y que a Pedro y a Linda les han pasado el que generalmente se considera como el mejor de los *tests*, el Stanford-Binet-Revisado, que se suministra individualmente y no requiere habilidad especial de lectura.

Los *tests* de inteligencia pretenden ser una muestra del intelecto. Al igual que cualquier otro producto que tenga su origen en la técnica del muestreo, el C. I. es una cifra que presenta un error estadístico, el cual expresa la precisión o intervalo de confianza de la cifra.

La información exigida a estos *tests* se obtiene en forma parecida a lo que usted haría para estimar la calidad del maíz de un campo, yendo al mismo y arrancando unas mazorcas aquí y allá al azar. Cuando usted hubiera arrancado y examinado un centenar de mazorcas, pongamos por caso, se habría formado una idea bastante buena de cómo era el campo. Su información tendría exactitud suficiente para ser utilizada al comparar este campo con otro, supuesto que los dos campos no fuesen muy similares. Si lo fuesen, tendría que examinar muchas más mazorcas, asignándolas a clases diferentes por medio de un estándar preciso de calidad.

La precisión atribuible a su muestra como representación de todo el campo es un dato que puede representarse en cifras: el error probable y la desviación estándar.

Supongamos que tiene que medir el perímetro de varios campos, dando pasos a lo largo de las vallas que lo circundan. La primera cosa que debe hacer es comprobar la exactitud de su sistema de medición, recorriendo un número determinado de veces una extensión de la cual sepa que tiene una longitud de cien yardas. Supongamos que la variación media sea de tres yardas. Es decir, en la mitad de las veces que probó hizo hasta tres yardas menos de las cien y en la otra mitad sobrepasó la longitud exacta hasta en tres yardas. Su error probable sería el de tres yardas en cada cien, o el 3 %.



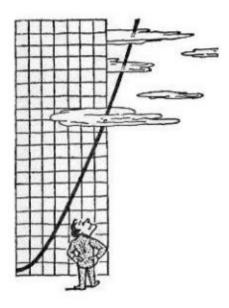
A partir de aquí, cada tendido de valía que midiese cien yardas según sus pasos, podría registrarse como 100 ± 3 yardas.

(La mayoría de los expertos en estadísticas prefieren utilizar una medición semejante a ésta y que se conoce con el nombre de desviación típica.)

Esta toma en consideración los dos tercios de los casos en lugar de tomar exactamente la mitad, y es mucho más manejable desde el punto de vista estadístico. Para nuestra finalidad seguiremos con el error probable, que es el utilizado todavía en relación con el Stanford-Binet.

Al igual que con nuestro hipotético recorrido, se ha hallado que el... continúa...

EL GRÁFICO EXCLAMATIVO

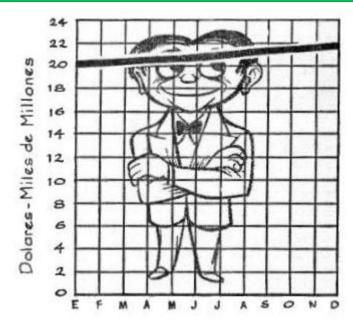


Existe cierto temor ante los números. No mucha gente emularía, tratándose de números, la confianza de Humpty Dumpty, personaje de "Alicia en el País de las Maravillas", el cual afirmaba "que era dueño de las palabras que utilizaba". Quizá sufrimos un trauma originado por las matemáticas del bachillerato.

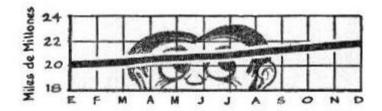
Sea cual fuere la causa, este temor crea un problema real al escritor que desea ser leído, al publicitario que espera con su texto vender productos, al editor que desea popularizar sus libros o revistas. Cuando los números en forma tabular son tabú y las palabras no van bien, como ocurre a menudo, solamente queda una solución: Dibujar un cuadro estadístico.

Entre los cuadros o gráficos más simples figura el lineal. Es de gran utilidad para representar tendencias; cosa que prácticamente interesa a todo el mundo para presentar, aprender, descubrir, lamentar o prever algo. Vamos a hacer que nuestro gráfico muestre cómo aumentó la renta nacional un 10 % en un año. Empecemos con un papel cuadriculado. Ponga el nombre de los meses en la línea marginal inferior. Indique los millares de millones en la línea marginal izquierda. Sitúe los puntos y trace una línea roja; su gráfica presentará este aspecto.

Esto está claro. Da a conocer lo que ocurrió durante el año y lo da a conocer mes por mes. El que va de prisa puede ver y comprender, porque toda la gráfica guarda proporción y existe la línea cero como base para su comprobación. El 10 % parece un 10 %, una tendencia ascendente que es sustancial, pero quizá no sea deslumbrante.



Todo va muy bien si su deseo es facilitar información. Pero suponga que usted desea llevarse la palma en una discusión, sorprender al lector, moverle a la acción, o vender-le algo. Para tales finalidades, a este gráfico le falta impacto. Corte la parte inferior.



Así va mejor. Además, ha ahorrado papel (a señalar si algún individuo quisquilloso pusiera objeciones a sus gráficos trucados). Las cifras son las mismas y también la curva. El gráfico es el mismo. No se ha falseado nada -excepto la impresión que produce – porque ahora el lector apresurado ve que la renta nacional ha subido en doce meses la mitad del espacio de papel que ocupa; todo gracias al hecho de eliminar la mayor parte de la gráfica. Al igual que las partes de la oración omitidas en las frases que se utilizan en la clase de gramática, se da por "sabida". Naturalmente, el ojo no "sabe" lo que no figura, y una pequeña tendencia ascendente se ha convertido, en el aspecto óptico, en una subida importante.

Ahora que usted ya se ha practicado en el engaño, ¿por qué limitarse a esa mutilación? Existe todavía un truco que vale lo que doce como éste. Con él podrá conseguir que el modesto aumento del 10 % parezca mucho más deslumbrante de... continúa...

Darrell Huff

CAPÍTULO 6



EL PERSONAJE DE LA GRÁFICA

Hace una década más o menos se habló mucho de la "gente pequeña", expresión que abarcaba prácticamente a todos nosotros. Cuando esto empezó a parecer demasiado condescendiente, nos convertimos en el "hombre común". Poco después esto también pasó al olvido, lo que tampoco me parece mal; pero el hombrecito se quedó con nosotros. Es el personaje de la gráfica.

En un gráfico descriptivo, el hombrecito representa un millón de hombres; una bolsa de dinero o un montón de monedas representa mil o un millón de dólares; el perfil de un ternero, el suministro de carne para el año siguiente. Es un sistema muy útil. Posee lo que se llama atractivo visual. Y puede convertirse en un mentiroso elocuente, tortuoso y con enormes posibilidades de éxito.

El padre de todos los gráficos descriptivos es el gráfico normal de barras, método simple y bien conocido para representar cantidades cuando hay que comparar dos o más. El gráfico de barras puede ser un engaño también. Mire con recelo cualquier versión en donde las barras cambien de anchura, o de longitud, mientras continúan representando el mismo factor, así como cualquier gráfico que represente objetos de tres dimensiones cuyos volúmenes no puedan compararse con facilidad.

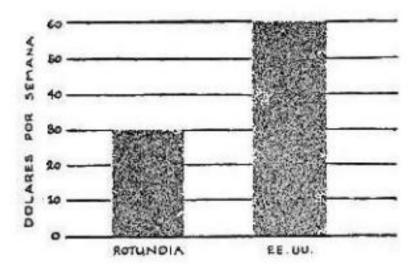
Un gráfico cuyas barras hayan sido mutiladas posee y merece la misma reputación que el gráfico recortado del cual hemos hablado con anterioridad. El gráfico de barras tiene carta de naturaleza en libros de geografía, en los estados financieros de las empresas y en las revistas de información. Lo mismo vale para su descendiente, el gráfico descriptivo.

Imaginemos que quiero representar la comparación entre dos cifras, el salario medio semanal de los carpinteros en los EE. UU. y los de Rotundia (nación imaginaría).

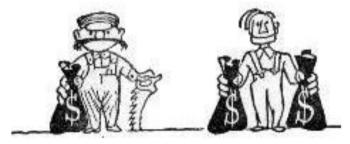
Las cantidades podrían ser 60 y 30 dólares respectivamente. Quiero atraer la atención del lector hacia este hecho y, por consiguiente, no me doy por satisfecho con imprimir las cifras solamente. Dibujo un gráfico de barras. (A propósito, si esta cifra de 60 dólares no es igual a la enorme suma que pagó usted el año pasado, cuando su porche necesitaba una barandilla nueva, recuerde que quizá a su carpintero no le haya ido siempre tan bien como cuando trabajaba para usted.

De todas formas, no he dicho de qué promedio se trataba o cómo lo conseguí; por lo tanto las lamentaciones no le llevarán a ninguna parte. Ya ve con cuánta facilidad puede parapetarse detrás de la estadística más deshonesta, si no da ninguna otra información.

Darrell Huff



Probablemente usted ha deducido que yo había inventado esta cifra como ejemplo, pero apuesto algo a que no lo habría creído así, si yo hubiera utilizado la cifra de 59,83 dólares.)



Aquí tenemos el gráfico en barras, con los dólares por semana indicados al lado izquierdo. Es un cuadro claro y honrado. El doble de dinero, duplica la barra en el gráfico y éste es también el efecto visual que produce. A pesar de ello, a la gráfica le falta atractivo visual, ¿verdad?

Puedo proporcionárselo fácilmente utilizando algo que simbolice mejor el dinero que una barra: una bolsa de dinero. Una bolsa para la menguada parte del carpintero de Rotundia y dos para el americano, o tres para el de Rotundia y seis para el americano. De cualquier forma, el gráfico continúa siendo claro y honrado, y no le inducirá a engaño incluso si lo examina con rapidez. Pero deseo más. Deseo decir que el trabajador americano está mucho mejor económicamente que el de Rotundia, y cuanto más pueda dramatizar la diferencia entre treinta y sesenta, mejor para mi argumento. A decir verdad (naturalmente, eso es lo que no deseo decir), quiero que usted infiera algo y se quede con una impresión exagerada, pero que no me coja en mis trucos. Existe un modo; es el utilizado a diario para enredarle. ... continúa...

CÓMO MENTIR CON ESTADÍSTICAS Darrell Huff

CAPÍTULO 7

LA CIFRA INDIRECTAMENTE RELACIONADA

Si no puede probar lo que desea, demuestre otra cosa y haga ver que es lo mismo. En el deslumbramiento que sigue al choque de las estadísticas con el cerebro humano, casi nadie se dará cuenta de la diferencia. La cifra que se relaciona indirectamente es un truco garantizado que le será de utilidad. Siempre lo ha sido.

Usted no puede demostrar que su preparado cura los resfriados, pero puede publicar (en grandes letras de molde) un informe, con la garantía de un laboratorio, asegurando que una onza de este producto mata 31.108 gérmenes en un tubo de ensayo en once segundos. Si lo hace, asegúrese de que el laboratorio posea una fama reconocida o un nombre impresionante. Reproduzca el informe en su totalidad. Fotografíe a un doctor en bata blanca y coloque su fotografía al margen del informe.

Pero no haga mención de los distintos trucos que ha empleado. No es asunto suyo — ¿verdad? — , hacer constar que un antiséptico que da buenos resultados en el tubo de ensayo, tal vez no cause efecto alguno al hombre después de haber sido diluido convenientemente para evitar que queme el tejido de la garganta. No comprometa el éxito diciendo qué clase de gérmenes destruyó. ¿Quién sabe cuáles son los gérmenes causantes de los resfriados, sobre todo cuando cabe la posibilidad de que no sea ningún germen? En resumen, no se conoce ninguna relación entre los gérmenes introducidos en un tubo de ensayo y la causa de los resfriados, pero la gente no va a hilar tan delgado cuando ha de sonarse.

Quizá este caso salta demasiado a la vista, y la gente empieza a darse cuenta; aunque no lo parece, a juzgar por las páginas publicitarias de la prensa. De todas formas, a continuación le ofrecemos una versión más mixtificada.

Pongamos por caso que durante un período de auge del prejuicio racial, le emplean a usted para "demostrar" lo contrario. No es un cometido difícil. Lleve a cabo una encuesta o, mejor aún, confíela a una organización de buena reputación para que la haga. Pregunte a las distintas clases de la población si creen que los negros tienen las mismas oportunidades profesionales que los blancos. Repita la encuesta periódicamente y de esta forma podrá establecer una tendencia.

La Oficina Princeton de Encuestas sobre la Opinión Pública probó una vez esta pregunta. El resultado obtenido fue una prueba patente de que las cosas, especialmente en las encuestas de opinión, no son siempre como parecen. Se formularon varias preguntas a cada una de las personas entrevistadas, acerca de la posibilidad de empleo, con intención de descubrir si sentían fuertes prejuicios contra los negros. Resultó que las personas más llenas de prejuicios contra los negros eran quienes contestaban SI a la pregunta sobre la igualdad de oportunidades. (Resultó que unos dos tercios de quienes sentían simpatía hacia los negros no creían que el negro tuviera tan buenas oportunidades de empleo como los blancos; y unos dos tercios de los afectados de prejuicios dijeron que los negros tenían tantas oportunidades como los blancos.) Es evidente que de esta encuesta podrían extraerse pocas conclusiones sobre las condiciones de empleo

Darrell Huff

de los negros, aunque podrían aprenderse cosas interesantes sobre las actitudes raciales del hombre. Como usted puede ver, si el prejuicio se extendiera durante de época de su encuesta, obtendría un número creciente de contestaciones en el sentido de que los negros tienen tan buenas oportunidades de empleo como los blancos.

Usted da a conocer los resultados: Su encuesta demuestra que los negros siempre tienen oportunidades. Ha conseguido un resultado notable utilizando la cifra que se relaciona indirectamente. Cuanto peor se presentan las cosas, mejor las hace aparecer su encuesta



O cojamos este ejemplo: "El 27 % de una extensa muestra de eminentes médicos fuma más Throaties que cualquier otra marca". La cifra puede estar falseada por muchas causas, pero esto no tiene ninguna importancia. La única contestación a una cifra que presenta tan poco fundamento es: ¿Y qué? Con todos mis respetos hacia la clase médica, ¿acaso los médicos saben más que usted acerca de las marcas de cigarrillos? ¿Poseen alguna información privada que les permite escoger el cigarrillo menos nocivo entre todos? Claro que no la poseen, y su médico sería el primero en decírselo. A pesar de ello, este "27 %" se las arregla de un modo u otro para sonar como si significara algo.

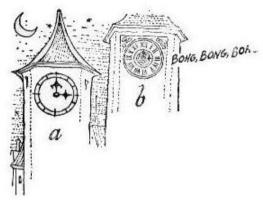
Ahora quitémosle un 1 % y veamos el caso del exprime-frutas. Se anunció como un sistema que "extrae el 26 % más de jugo", según "atestigua un test de laboratorio" y "certifica el Instituto de Economía Doméstica".

Esto suena bien. Si usted puede comprar un exprime-frutas que es un 26 00 más efectivo, ¿por qué comprar otro? Ahora bien, sin meternos en el hecho de que los tests de laboratorio (en especial de los laboratorios independientes), han demostrado las cosas más disparatadas, ¿qué significa esta cifra? ¿El 26 % de qué?

Cuando se consiguió una respuesta a esta pregunta, resultó significar solamente que con este exprime-frutas se obtenía mucho más jugo que con el antiguo exprimidor manual. No tenía relación alguna con la información que usted desearía poseer antes de efectuar la compra; aquel exprime-frutas quizá fuera el peor del mercado. Además, por ser tan precisa, esta cifra del 26 % está totalmente fuera de lugar.

Los anunciantes no son los únicos que le engañarán con cifras si les deja. Un artículo sobre la seguridad del conductor publicado en la revista This Week informaba, sin duda con las mejores intenciones y deseos para el lector, de lo que podía sucederle "si se lanzaba por la carretera a cien kilómetros por hora, haciendo eses de uno a otro lado". Usted tendría, según el artículo, cuatro probabilidades más de... continúa...

EL "POST HOC" APARECE DE NUEVO



Dados dos relojes que marchan perfectamente. Cuando "a" señala la hora, "b" toca la hora. ¿Fue "a" la causa de que "b" tocase?

En cierta ocasión, alguien se tomó un trabajo ímprobo para averiguar si los fumadores conseguían notas más bajas en sus licenciaturas que los no fumadores. Resultó que así era en efecto. Esto satisfizo a mucha gente y desde entonces, se le ha sacado mucho partido. El camino hacia la licenciatura, según parece, comienza dejando de fumar; para llevar la conclusión un paso más allá, añádase que fumar embota la mente. Supongo que el estudio que nos ocupa se hizo con toda propiedad: la muestra era suficientemente amplia y se escogió con cuidado y honradez, la correlación tenía un alto grado de significación, etc.



El engaño es viejo y, sin embargo, se introduce con frecuencia entre el material estadístico, bajo el disfraz que le proporciona un fárrago de cifras impresionantes. Es el sofisma que dice: si B sigue a A, A es la causa de B. Sin garantía alguna se saca la

Darrell Huff

conclusión de que, como sea que el fumar y las malas notas van parejas, el fumar es causa de las malas notas. ¿No podría ser al revés? Quizá las malas notas no inclinan a los estudiantes a la bebida, sino al tabaco. Examinada de cerca, esta conclusión tiene las mismas probabilidades de ser cierta que la otra, y el mismo acopio de pruebas. Pero no resulta tan satisfactoria para los propagandistas.

Sin embargo, parece mucho más probable que ninguna de estas dos cosas haya causado la otra, siendo ambas producto de un tercer factor. ¿No es posible que el individuo sociable que no toma los libros muy en serio sea también quien más fuma? ¿O encontraríamos la clave en la correlación que alguien estableció entre la extroversión y las notas bajas, correlación aparentemente más positiva que la existente entre las notas y la inteligencia? Tal vez los extrovertidos fumen más que los introvertidos. Cuando existen varias explicaciones razonables para un mismo hecho, nada le autoriza a usted a escoger la que satisfaga a su gusto, e insistir sobre la misma; pero mucha gente lo hace.

Para no caer en el error "post hoc" y creer muchas cosas que no son ciertas, debe someter a rígido examen cualquier cosa que se afirme o se relacione con otra. La correlación, esta cifra precisa y convincente que parece demostrar que una cosa ocurre a causa de otra, puede ser de distintos tipos.

Uno de ellos es la correlación producida por la casualidad; de este modo quizá consiga reunir una serie de cifras que prueben algo improbable; pero al tratar de conseguirlo otra vez, su nueva serie probablemente no lo pruebe en absoluto. Como hacía el fabricante de la pasta para los dientes que al parecer evitaba las caries, se trata simplemente de desdeñar los resultados desagradables y dar amplia publicidad a los agradables. Partiendo de una muestra reducida, existe probabilidad de hallar alguna correlación sustancial entre cualquier par de características o hechos cualesquiera.

Un tipo corriente de covariación es aquel en que hay relación real, pero no existe posibilidad de establecer cuál de las variables es la causa y cuál el efecto. En algún caso, la causa y el efecto pueden cambiar de posición a veces, o ambas variables pueden ser causa y efecto al propio tiempo. Una correlación entre ingresos y propiedad de acciones quizá sea de este tipo. Cuanto más dinero se gana, más acciones se adquieren, y cuantas más acciones se adquieren, más ingresos se obtienen; no es exacto decir simplemente que lo uno produce lo otro.

Quizá el más engañoso es el caso tan corriente en que ninguna de las variables tiene efecto alguno sobre la otra, pero existe a pesar de todo una correlación real. Utilizando este medio se ha llevado a cabo mucho trabajo sucio. El caso de las notas bajas entre fumadores figura dentro de esta categoría, igual que muchas otras estadísticas médicas citadas sin aclarar que, aun demostrándose la existencia de una relación real, el origen de la causa o efecto es sólo especulativo.

Como ejemplo de correlación absurda o falseada sobre un hecho estadístico real, un gracioso señaló que existe una estrecha correlación entre los salarios de los ministros de la iglesia presbiteriana de Massachusetts y el precio del ron en La Habana.

¿Cuál es la causa y cuál el efecto? En otras palabras, ¿se benefician los ministros de la iglesia del comercio del ron o lo manejan? La conclusión es tan... continúa...

CÓMO MENTIR CON ESTADÍSTICAS Darrell Huff

CAPÍTULO 9

CÓMO ESTADISTIQUEAR

El informar mal, utilizando material estadístico, podría llamarse manipulación estadística, y resumiéndolo en una sola palabra (aunque no sea muy buena), estadisticulación. El título de este libro y alguna de las cosas contenidas en el mismo, parecen indicar que todas estas operaciones son producto del deseo de mover a engaño. El presidente de un congreso de la Asociación Americana de Estadística me llamó la atención por tal motivo. La mayor parte de las veces no se trata de trapacerías, dijo, sino de incompetencia. Quizá haya algo de verdad en ello, pero no creo que este supuesto sea menos ofensivo que el otro para los expertos en estadísticas. Posiblemente sea de mayor importancia tener en cuenta que la tergiversación de datos estadísticos y su manipulación para conseguir una finalidad determinada, no son siempre obra de los profesionales en estadística. Lo que sale lleno de virtudes de la mesa del experto puede verse cambiado, exagerado, demasiado simplificado y tergiversado al haber tomado solamente una parte los vendedores, expertos en relaciones públicas, periodistas y redactores de textos publicitarios.

Sea quien sea el culpable en un caso determinado, se hace difícil achacarlo a su incapacidad inocente. Los gráficos falsos que aparecen en las revistas y periódicos resultan sensacionalistas por su exageración, y raras veces minimizan nada. Quienes presentan argumentos estadísticos en favor de la industria, pocas veces, según he experimentado, dan a los obreros o al público consumidor una información que mejore los hechos; generalmente los presentan aún peor. ¿Acaso emplearía un sindicato los servicios del experto en estadísticas que fuera tan incompetente como para debilitar aún más la posición de los obreros?

Mientras los errores estén todos de una parte, no es fácil atribuirlos a falta de conocimiento o a la casualidad.

Una de las formas más desleales de tergiversar los datos estadísticos es utilizando un mapa. Un mapa presenta gran cantidad de variables donde pueden esconderse los hechos y tergiversarse las relaciones. Mi trofeo favorito en este terreno es "La Sombra que se Extiende". La publicó no hace mucho el First National Bank de Boston, dándole gran difusión a través de los grupos de contribuyentes, los periódicos y la revista Newsweek.

El mapa muestra la parte de nuestra renta nacional que el Gobierno recauda y gasta. Lo hace sombreando la zona de los Estados al oeste del Misisipí (con excepción de Luisiana, Arkansas y parte de Missouri), para indicar que los gastos del Gobierno federal son iguales a los ingresos totales de la población de estos Estados.

El engaño está en escoger Estados de gran extensión, pero de ingresos relativamente reducidos, porque la densidad de población es muy baja. Con igual honradez (o fraude) el autor del mapa podría empezar sombreando Nueva York o Nueva Inglaterra, y habría resultado una sombra muchísimo más pequeña y menos impresionante. Utilizando los mismos datos, habría dado una impresión muy distinta a cualquiera que

Darrell Huff

fijase los ojos en el mapa. No obstante, nadie se preocuparía de distribuir esta versión del mapa. Al menos, no conozco ningún grupo poderoso interesado en presentar su contribución pública como menor de lo que es.

Si el objetivo del autor del mapa era simplemente dar a conocer una información, pudo hacerlo fácilmente. Pudo escoger un grupo de Estados cuya área total presentase la misma proporción respecto del área total del país, que la relación de su renta total con la renta nacional.

Este mapa se convierte en un esfuerzo flagrante por inducir al error, dado que no se trata de un truco nuevo de propaganda. Se trata de una vieja historia. Hace tiempo este Banco publicó versiones del mismo mapa para presentar los gastos federales de 1929 y 1937. Estos mapas rápidamente pasaron a figurar entre los "ejemplos horribles" del clásico libro de Willard Cope Brinton: "Ghaphic Presentation".

LA SOMBRA QUE SE OSCURECE



(Estilo Este)

(Estilo Estilo Estilo

Para demother que un enquirament hemos anad an MARYLAND, DEBWARE . RHODE ISLAND

Este método "tergiversa los hechos", decía Brinton sin ambages. Pero el First National continúa dibujando sus mapas y Newsweek y otros que deberían tener mayor conocimiento del tema y posiblemente lo tengan, continúan reproduciéndolos sin advertir al lector ni disculparse.

¿A cuánto ascienden los ingresos medios de la familia americana? Como mencionamos en un capítulo anterior, la Oficina del Censo dio para 1949 la cifra de 3.100 dólares como "ingresos medios de una familia". Pero si lee usted el reportaje publicado en un periódico por la Russell Sage Foundation, sabrá que para el mismo año los ingresos medios alcanzaron la notable cifra de 5.004 dólares. Posiblemente a usted le satisfaría que a todos nos fueran las cosas tan bien, pero quizá quede... continúa...

CÓMO ENFRENTARSE CON LAS ESTADÍSTICAS



Hasta aquí, me he dirigido a usted como si yo fuera un pirata deseando instruirle en el manejo del trabuco. En el capítulo que cierra este libro, voy a prescindir de este truco literario. Voy a referirme al propósito serio que creo se puede entrever bajo la superficie de este libro; explicando cómo debe mirarse una estadística falseada, y desenmascararla; y aún más importante, cómo reconocer los datos útiles y ciertos entre la marejada de fraudes a los cuales he dedicado los capítulos anteriores.

No toda la información estadística que caiga en sus manos puede comprobarse con la seguridad del análisis químico o como se hace en un laboratorio para valorar metales. Pero usted puede tamizar la información con 136 el cedazo representado por cinco preguntas muy sencillas. Hallando las respuestas, evitará dar por verídicas muchas cosas que no lo son.

¿QUIEN LO DICE?

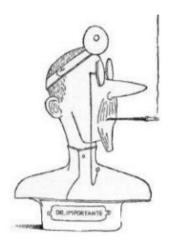
Lo primero que debe mirar es en qué sentido puede estar influida la información: el laboratorio que necesita demostrar algo para beneficio de una teoría, una reputación o unos honorarios; el periódico cuya finalidad es publicar un buen artículo; los obreros y los patronos que discuten un aumento de salarios.

Busque la influencia consciente. El método tal vez sea la desfiguración de un hecho en forma directa, o una afirmación ambigua que sirva igualmente para la finalidad que se desee y cuya culpabilidad no pueda probarse. Puede tratarse de una selección de los datos favorables con supresión de los desfavorables. Pueden ser engañosas las unidades de medida, como cuando se toma un año para una comparación y se pasa a un año más favorable para otra. Quizá se utilice una medida inadecuada; una media cuando la mediana sería mucho más informativa (quizá demasiado informativa), cubriendo el truco con la palabra "promedio" sin calificativo.

Busque con atención la influencia inconsciente: A menudo es más peligrosa. En los

Darrell Huff

gráficos y predicciones de muchos economistas y expertos en estadísticas publicadas en 1928, se infiltró de tal manera este factor de influencia que se produjeron resultados curiosos. Fueron desestimados alegremente los trastornos de la estructura económica y se adujeron toda clase de pruebas justificadas estadísticamente para demostrar que habíamos entrado en la corriente de la prosperidad.



Tal vez haga falta por lo menos una segunda inspección para enterarse de quién lo dice. El "quién" puede esconderse tras lo que Stephen Potter, el hombre del Lifemanship, llamaría probablemente un "nombre O. K.". Todo lo relacionado con la profesión médica es un "nombre O. K.". Los laboratorios científicos tienen "nombres O. K.".

Lo mismo ocurre con las instituciones, especialmente las Universidades, y sobre todo las que destacan por su labor técnica. El escritor que demostraba, como vimos unos capítulos más atrás, que una educación universitaria compromete las oportunidades de matrimonio de las chicas, hacía buen uso del "nombre O. K." de la Universidad de Cornell. Observe que mientras los datos procedían de Cornell, las conclusiones eran obra exclusiva del escritor; pero el "nombre O. K." le ayudó a crear la falsa impresión de "la Universidad de Cornell dice...".

Cuando se cite un nombre O. K., asegúrese de que la autoridad está detrás de la información, no como algo presentado al lado de la misma.

Puede que haya leído una presuntuosa información del Journal of Commerce de Chicago. Esta publicación llevó a cabo una encuesta. De 169 empresas que contestaron a un cuestionario sobre los acaparadores y el aumento abusivo de precios, dos tercios declararon que estaban reduciendo el aumento de los precios producido por la guerra de Corea. "La encuesta demuestra" (cuidado siempre y cuando tropiece con estas palabras), "que las empresas han hecho exactamente lo contrario de lo que les acusan los enemigos del sistema americano de negocios". Este es un punto clave para hacerse la pregunta "¿Quién lo dice?", ya que el Journal of Commerce podría considerarse como parte interesada. Es también un punto estupendo para formularse... continúa...

Darrell Huff

GLOSARIO

COCIENTE INTELECTUAL C.I. C. I. = (Edad mental E.M.) / (Edad cronológica E.C). en la que E.M. viene determinada por la estandarización del test mental y E.C. es la edad real del individuo.

CORRELACIÓN. Tendencia de dos series de medidas a variar en forma concomitante, según el grado de correlación que presenten. Este se mide por medio del coeficiente de correlación.

DESVIACIÓN ESTÁNDAR. Es la raíz cuadrada de la media aritmética de los cuadrados de las desviaciones de los elementos con respecto a la media aritmética.

ERROR PROBABLE. Medida de los límites dentro de los cuales, partiendo de la media de una muestra escogida al azar, habría de caer la media que se obtuviese con cualquier otra muestra.

INFLACIÓN. Reducción progresiva del valor adquisitivo del dinero debido a que demasiado dinero intenta hacerse con unos productos insuficientes. – Contrario = DE-FLACIÓN.

INTELIGENCIA GENERAL (**Factor G**). Según Stern, la Inteligencia es: "La capacidad general de un individuo de orientar conscientemente su pensamiento ante situaciones nuevas".

MEDIA ARITMÉTICA Y MEDIA GEOMÉTRICA. Dada una serie de n números, la media aritmética es igual a la suma de todos ellos dividida por n. La media geométrica es la raíz enésima del producto de esos n números. La presencia, en la serie, de un valor muy variante, descompensa más la media aritmética. Ejemplo: 2, 4, 125. media aritmética: (2 + 4 + 125)/3 = 43,666 media geométrica (más informativa)

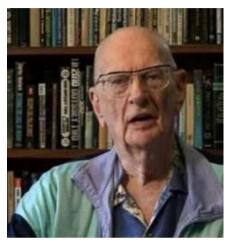
MEDIANA. En una distribución de datos sin agrupar en la cual los términos están colocados en orden creciente o decreciente, la mediana es el término que ocupa la posición central. En la serie 1, 2, 3, 5, 7, la mediana es el número 3.

MODA. La moda es el valor que se presenta con mayor frecuencia. Cuando se trata de una distribución de datos agrupados, es mejor hablar de clase modal. Para su determinación, K. Pearson ha propuesto la fórmula aproximada siguiente: (3 veces la mediana) / (2 veces la media) = MODA

MUESTREO. Selección, generalmente al azar, de un número limitado de casos de entre un grupo extenso o UNIVERSO, para someterlo a tratamiento estadístico, bajo el supuesto de que la muestra es representativa de todo el... continúa...

CÓMO MENTIR CON ESTADÍSTICAS Darrell Huff

Darrell Huff (15 de julio de 1913 - 27 de junio de 2001) fue



un escritor estadounidense, principalmente conocido por ser autor del libro *How to Lie with Statistics* (1954) "Cómo mentir con estadísticas", el libro de estadística más vendido de la segunda mitad del siglo XX. Huff nació en Gowrie, Iowa, y cursó sus estudios en la Universi-

dad de Iowa, (recibiendo los grados de BA en 1938, y MA en 1939). Antes de convertirse en escritor a tiempo completo en 1946, Huff trabajó como editor en las revistas *Better Homes and Gardens* y *Liberty* magazine. Como escritor independiente, Huff produjo centenares de artículos del tipo «Cómo hacer...» y escribió al menos dieciséis libros, la mayoría acerca de arreglos domésticos. Uno de sus mayores proyectos, que ganó un premio, fue la casa en Carmel, California donde vivió hasta su muerte.

Su mayor aportación ha sido introducir conceptos estadísticos a toda una generación de forma amena y práctica, sin evitar los conceptos matemáticos más complejos. Su libro más famoso *How to Lie* with Statistics se sigue traduciendo a nuevos idiomas y sigue siendo una referente entre los libros de texto.