Universidad de Valladolid

INCERTIDUMBRE Y CONTEXTO

Tópicos y Actividades de Estadística y Probabilidad para Enseñanza Secundaria

TRABAJO FIN DE MÁSTER DE PROFESORADO
DE EDUCACIÓN SECUNDARIA OBLIGATORIA Y BACHILLERATO

ESPECIALIDAD MATEMÁTICAS

DAVID GARROTE MORAL

TUTOR: ALFONSO GORDALIZA RAMOS

"Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write."
(H.G. Wells, 1946)

ÍNDICE

INCERTIDUMBRE Y CONTEXTO

RESUMEN

La situaciones de incertidumbre involucran fenómenos cuyo aspecto más importante es la *variablidad*; para acometer su análisis, nuestras intuiciones innatas son torpes y las estrategias que se siguen de las mismas conducen muchas veces a error. No obstante, analizar estas situaciones en orden a formar juicios, tomar decisiones y ejecutar acciones es parte del desarrollo personal insoslayable de todo ciudadano, dado que la variabilidad es omnipresente y la información, por razones diversas, nunca es completa. Es, por tanto, una responsabilidad del formador de la totalidad de la ciudadanía, es decir, el docente de enseñanza secundaria —sobre todo en su período obligatorio—, ayudar a *construir* las estrategias adecuadas a tales efectos, corrigiendo asimismo los errores primeros sobre la materia, innatos o inculcados social o culturalmente. Lo que diferencia a los *números*, objetos matemáticos abstractos, de los *datos*, descripciones cuantitativas de los objetos de las ciencias factuales y razón de ser de la estadística, es el **contexto**, cuya fuente puede ser la historia, las informaciones de los medios, los estudios sociales o económicos, los experimentos científicos, los diversos aspectos de la vida cotidiana, o incluso la misma historia de los datos (su recolección, transmisión, organización, etc.) El análisis de este contexto desde diversas metodologías es absolutamente necesario para comprender las características del fenómeno al que responden y para enunciar tesis sobre el mismo y sus propiedades, en orden a facilitar las antedichas estrategias de juicio, decisión y acción.

Por todo ello, en el presente trabajo se han escogido tópicos relacionados con situaciones de incertidumbre extractados de fuentes diversas (desde la simple cotidianeidad del alumno, pasando por informaciones procedentes de los medios de masas, conclusiones de estudios científicos, hechos registrados por la historia, tesis sobre fenómenos sociales, políticos y económicos, etc.) pero siempre en referencia a un contexto o relato que les es propio en cada caso, en vistas a dotar a las actividades correspondientes de relevancia para el alumno y sus intereses, actuales y futuros. Asimismo, en cada actividad se trata de poner de relieve la utilidad, y, en muchos casos, la necesidad de emplear los instrumentos que nos brinda la teoría de la probabilidad y la estadística, sobre todo desde sus aspectos descriptivos, pero sin obviar los matemáticos, para entender el mundo que nos rodea; sirviéndonos, además, del carácter preciso y el rigor lógico que ofrecen las anteriores disciplinas (y las matemáticas en su conjunto) para despertar en el alumno el espíritu crítico, para hacerle más reflexivo y analítico, en definitiva para desarrollarle como ciudadano consciente y activo.

INCERTIDUMBRE Y CONTEXTO

INTRODUCCIÓN

Estamos rodeados de fenómenos sometidos a variabilidad e incertidumbre: hoy hace sol, mañana puede llover, pasado mañana, nevar. Mis parientes, mis amigos, los transeúntes que veo por la calle son de diversas alturas, pesos, color de ojos y cantidad de ingresos; y no por ello concluyo que son entes categóricamente distintos, ontológicamente aislados, sino que constituyen diversas ocurrencias de una misma clase de objetos, esto es, seres humanos, y a su vez juzgo que sus diferencias son una muestra de la variabilidad intrínseca a la especie (si se encontrase a un hombre con alas, deberíamos, o bien determinar que el individuo es de una especie distinta —–es decir, que constituye un fenómeno distinto al hecho de "ser humano"— o ampliar nuestra definición del fenómeno "ser humano" para incluir esta clase concreta de variación —tener alas— dentro del rango de variabilidad asumible de la especie). Desde los albores del pensamiento filosófico-científico se ha tratado de dar cuenta de las distintas clases de entes, objetos, representaciones, sucesos, procesos, fenómenos en fin (aunque llamarles de cada una de estas maneras ya tiene implicaciones filosóficas en las que no podemos abundar), pero no de la variabilidad dentro de cada fenómeno[1], la cual, por su carácter *azaroso* o *aleatorio*, se consideró durante mucho tiempo fuera del alcance del conocimiento, o, más bien, del conocimiento *respetable* (filosófico o científico). La ortodoxia filosófica, y en primer lugar Parménides (quien negaba la diversidad de los seres, e incluso la realidad de ningún proceso de cambio), al que siguió Platón (para el cual las diversas variaciones respecto de un prototipo o "idea" eran mera apariencia) y más tarde Aristóteles, estimaban la variabilidad de la naturaleza como asunto objeto de opinión (*doxa*) y no de verdadero conocimiento (*episteme*). Los fenómenos aleatorios, es decir, los menos ordenados y más variables de todos, entre los que incluían tanto los juegos de azar como el auge y caída de los imperios, pasaban directamente a ser asunto de la Diosa Fortuna, es decir, de tahúres, adivinos o sofistas como algunos partidarios de la indivisibilidad última de la materia y, posteriormente, los epicúreos, quienes (según la ortodoxia, claro) sostenían la absurda tesis de que la materia se había formado por colisiones casuales, fortuitas, de átomos **[14]**.

Tal postura prevalece hasta bien entrada la era moderna (según los cánones de los historiadores) ya se considere como obra fundacional de la teoría de la probabilidad la obra de Cardano, *De Ludo Aleae*, o la bastante posterior correspondencia entre dos "diletantes" y un jugador, (Pascal, Fermat y el caballero de Mère). Unos comienzos tan tardíos, junto con un desarrollo posterior bastante problemático, de una disciplina que versa sobre una propiedad de los fenómenos tan generalizada, tan evidente, y, en muchos casos, tan decisiva para la vida de muchas personas como es la variabilidad y el azar, junto con la humana —y a menudo, imperiosa— necesidad de actuar en circunstancias dominadas en todo o en parte por

1 Utilizaremos de ahora en adelante el significado operativo, o, más bien, sobreentendido, que se otorga a la voz "fenómeno" en la mayor parte de los textos científicos, es decir: lo realmente dado o acaecido, considerado independiente del hecho de ser observado, conocido, definido o comunicado por un sujeto, aunque tal concepto incluya por sí mismo varias hipótesis de principio que aún son materia de discusión en filosofía y en algunas campos científicos, como la mecánica cuántica.

la incertidumbre, la casi cotidiana tarea de adjudicar, calcular y decidir sobre la probabilidad de determinados sucesos (¿Debería cultivar trigo o centeno? ¿O vender mis tierras e invertir en la Compañía de los Mares del Sur?) constituye uno de los enigmas de la historia de la ciencia (o de la matemática), sobre la que se han propuesto diversas conjeturas **[30][14]**, entre las cuales se hace referencia a la resistencia opuesta por un modo de pensamiento determinista, impuesto por necesidades filosóficas en el mundo antiguo y consagrado por el cristianismo (vid. p.104) y por tanto, refractario al mismo concepto de azar o casualidad por principio. En apoyo de esta idea, y aduciendo el siguiente texto de quien, en cierto modo, representa el punto final del proceso de mteorización matemática rigurosa de la teoría de la probabilidad, es decir, A. N. Kolmogorov **[33]**:

> *Pero si de un proceso real abstraemos sus aspectos esenciales, nos quedamos con un cierto residuo que debemos considerar aleatorio.* [...] *Muy pocos fenómenos susceptibles de tratamiento matemático dejan, cuando la teoría es comprada con la observación, de mostrar la influencia de factores aleatorios ignorados. Este es más o menos el estado de cosas en la teoría del movimiento planetario bajo la fuerza de la gravedad: la distancia entre planetas es tan grande en comparación con su tamaño que su idealización como puntos materiales es casi perfectamente satisfactoria; el espacio en el que se mueven contine una materia tan dispersa que su resistencia al movimiento es casi infinitamente pequeña; las masas de los planetas son tan grandes que la presión de la luz apenas juega un papel en su movimiento. Estas circuntancias excepcionales explican el hecho de que la solución matemática para el movimiento de un sistema de n puntos materiales, cuyos estados se describen mediante* 6^n *parámetros que solo tiene en cuenta la fuerza de la gravedad, se conforma tan bien con las observaciones de los movimeintos de los planetas.*

concluimos que el pensamiento determinista quizá estaba *él mismo determinado* por los instrumentos matemáticos de la época, los cuales a su vez pudieron tener origen en el único fenómeno al alcance de la observación en época clásica que se ajustaba al estándar determinista, esto es: los movimientos planetarios. La única aplicación científica de la geometría antigua era la astronomía, y ello hasta tal punto que matemático y astrónomo eran términos casi sinónimos, por lo menos en la época de Copérnico: y no hay que olvidar que los *Principia Mathematica* contienen casi exclusivamente demostraciones geométricas, y que Newton sólo recurre al cálculo infinitesimal inventado por él mismo (o, como él lo llamaba, método de las fluxiones) cuando no le queda más remedio, casi como una construcción *ad hoc* en apoyo de su teoría **[32]**

Este cambio de mentalidad, si bien lento y problemático, tuvo que tener necesariamente lugar cuando las ciencias factuales comenzaron a considerar fenómenos demasiado complicados, demasiado *variables* como para encajar en los antiguos moldes deterministas. No obstante, en cierto modo el desarrollo teórico —y práctico— de las disciplinas cuyo tema principal es determinar el juicio y la acción en contextos de incertidumbre siguió un camino paralelo, y controvertido, respecto de las ciencias matemáticas y físicas. ¿Hasta qué punto esta distancia, sólo acortada muy modernamente, influyó en el desarrollo teórico de los problemas relacionados con la aleatoriedad, llegando a determinar los distintos papeles de las disciplinas que estudian la incertidumbre, como son la teoría de la probabilidad, la estadística matemática e inferencial? ¿Hasta qué punto este hecho tiene influencia sobre la didáctica de estas materias, o sobre el lugar que ocupan en los diversos currículos de matemáticas en enseñanazas medias?

La respuesta a los anteriores interrogantes daría cuenta de las razones *históricas* de ser del peso social y económico de las disciplinas reseñadas, y por ende de su peso en los divesos planes de estudio de enseñanzas medias, (hablando por el momento en términos transnacionales), y quizá más importante aún, del *enfoque* de los contenidos correspondientes a aquéllas dentro del área de las matemáticas en secundaria. Pero la razones históricas no agotan la cuestión. Aunque se deja entrever en los anteriores párrafos, señalamos ahora que el mismo núcleo de la idea del azar (o aleatoriedad) muestra características singulares, de carácter *epistemológico,* que no solo han influido en la historia de los anteriores saberes, sino que influyen decisivamente en el papel social actual que desempeñan, y sobre todo, en su aprendizaje: lo cierto es que nuestra misma intuición de lo aleatorio es problemática y proclive al error, como veremos, sobre todo gracias a la obra clásica de Kahneman y Tversky **[29]**, y a la gran cantidad de problemas, falacias y paradojas a las que ha dado lugar el tema general del azar, la variabilidad, y la incertidumbre, hasta el punto de producir casi un subgénero propio dentro de la literatura, tanto matemática (*vid.* **[22]**, **[16]**) como de psicología cognitiva (*vid. supra* **[29]** y **[6]**)(o epistemología, como la prefieren considerar algunos estudiosos, entre ellos, Piaget) Este sería pues, el segundo punto a considerar en la presente introducción.

Finalmente, el tercer elemento que ha configurado el papel de los distintos discursos, teorías, técnicas y profesiones relacionadas con la aleatoriedad es la *sobreabundancia de datos,* la cual, aunque tiene antecedentes históricos claros **[23]**, nunca ha sido tan predominante en la historia como en su período actual, debido al desarrollo galopante de la informática y las telecomunicaciones y la generalización de su uso a todo el espectro social, fenómenos que parten de la segunda mitad del siglo XX y que crecen exponencialmente hasta la actualidad, y que ha ido forjando paulatinamente las técnicas propias y la misma profesión del estadísitico como *analista de datos,* más que como teórico de una rama matemática aplicada o experto en diseño de experimentación científica. La aparición de titulaciones superiores de estadística no necesariamente englobadas como especialidad matemática ha ido pareja a la creación y desarrollo del estadístico como profesional liberal, cuya principal función es el tratamiento de datos, por delante de su papel como teórico de la probabilidad o la estadística matemática **[45]** **[10]** **[36]**. Como consecuencia, tanto los teóricos de la didáctica de la estadística, como los responsables de los marcos teóricos del informe PISA, llegando incluso hasta la última reforma de la ley de educación (LOMCE) ha adjudicado un papel cada vez mayor a los "datos" y sus herramientas de análisis entre los principios básicos de la enseñanza del bloque que nos ocupa.

Estos tres elementos de juicio que hemos expuesto y que inmediatamente desarrollaremos constituyen los principios que nos encaminaron a la elección y elaboración del presente trabajo, el cual seguirá las bases que aquí adelantamos para el lector impaciente:

1ª)Las áreas del currículo de matemáticas de enseñanza secundaria correspondientes a probabilidad y estadística pueden desplazar su centro de gravedad desde sus aspectos más teóricos, y sus aplicaciones algorítmicas más básicas, hacia las técnicas de representación, análisis y valoración de datos o de las organizaciones y representaciones de los mismos, en orden a formar juicios, tomar decisiones y ejecutar acciones de interés efectivo para el educando, sobre todo en la fase obligatoria (ESO) de aquellas enseñanzas. No tanto la resolución de problemas numéricos o simbólicos como la atención, la capacidad de formular hipótesis, emitir juicios y, en su caso, definir estrategias de acción conforman valores cívicos en la educación del alumno y posterior desarrollo personal como adulto.

2ª) Es el **contexto** el que da significado a los datos, y no sólo eso: también el que despierta el interés y el aprecio por la práctica de su análisis. Este contexto puede ser más o menos cercano a la cotidianeidad del alumno: sus fuentes pueden ser directas, o procedentes de los medios de masas con los que tiene fuerte contacto, pero también pueden escogerse de temáticas de interés social, económico, científico y/o histórico. En todo caso los datos deben detentar alguna clase de relevancia en un acaecimiento factual, y no sólo ocupar un lugar inserto en un sistema abstracto, lógico-simbólico o matemático.

Y que se justificarán una vez se hayan discutido los principios en los que se basan: cosa que haremos inmediatamente.

Apuntes históricos

Ante todo, sería bueno empezar considerando el origen independiente de las dos partes de esa pareja de hecho (o matrimonio morganático, según cuál sea la parte preguntada) que son la probabilidad y la estadística. Que su unión no es, de todas formas, contingente o casual en sí misma, hablan, aunque sea a regañadientes, los mismos teóricos de ambos campos. Por ejemplo, en el clásico de William Feller (*vid.* [**17**], p. 21) se habla, desde las primeras páginas, de probabilidad *estadística,* aunque el término es en su caso, casi un sinónimo de probabilidad frecuentista u objetiva frente al —~~y~~ en contra de— concepto intuitivo, relacionado con el "*razonamiento inductivo*", que lleva a "juicios *como el de* "Pablo es probablemente un hombre feliz" *o* "Este libro probablemente será un fracaso"". De modo parecido habla Kolmogorov [**33**]: tras establecer que la "ley de la probabilidad" depende de la "Ley estadística" según la cual determinado fenómeno experimentado un cierto número de veces muestra unas frecuencias que van acercándose a un cierto valor *p (*que denominamos como *probabilidad*) a medida que el número de experimentos es mayor, de lo que se concluye que:

> *Dado que no se puede dudar que las leyes estadísticas son de gran importancia, volvemos a la cuestión de los métodos de estudiarla. Primeramente uno piensa en una manera de proceder completamente empírica. Puesto que la ley* [función] *de probabilidad se muestra solo en procesos en masa, es natural pensar que para descubrir esa ley debemos ejecutar un experimento en masa.*
> *Tal idea, de todas formas, es solo parcialmente correcta. Tan pronto como hemos establecido ciertas leyes de probabilidad mediante experimentación, podemos proceder a deducir nuevas leyes de la probabilidad por medios lógicos o mediante cálculo, bajo ciertos supuestos generales.*

lo cual viene a confirmar que la teoría ortodoxa o "frecuentista" de la probabilidad busca sus raíces, no en la intuición, la inducción o la subjetividad, sino en un hecho estadístico empírico. No fueron éstos, sin embargo, los comienzos de la teoría de la probabilidad. los primeros teóricos de la misma eran, desde Pascal a Pierre Simon Laplace, estrictos deterministas que no hubieran reconocido la aleatoriedad de los fenómenos variables como un hecho, y a los cuales escapaba la verdad empírica de la ley de los grandes números (expresión de Poisson, ya en 1835). Su noción de la probabilidad, denominada por algunos historiadores *clásica,* era directamente dependiente de *la ignorancia* del sujeto conocedor, no del carácter aleatorio del suceso, que los deterministas negaban. De ahí que, si un sistema es capaz de asumir un número de estados posibles, ante *la ignorancia* de un investigador *racional* —es decir, carente de prejuicios a favor o en contra del acaecimeinto de un suceso— acerca de cuál de ellos se da de hecho, siempre que "no haya nada que nos induzca a pensar que uno de los estados es más probable que otro", esos distintos estados posibles tienen las mismas oportunidades de acaecer, es decir, se pueden considerar *equiprobables*[1]. Luego si un suceso

1 Esta concepción, si bien parte del grado de conocimiento del sujeto y no del carácter aleatorio del fenómeno en sí, podría llamarse "subjetiva" —pese a que nació en un contexto en que la "racionalidad" (del investigador) se tomaba como un hecho objetivo—, pero el término se presta a confusión, puesto que se usa más bien para designar otras interpretaciones, entre ellas la bayesiana— la cual se basa, no en la elección apriorística, "racional" de la probabili-

A investigado, y por definición n estados implican A, siendo m el número de estados posibles, entonces tenemos la regla de Laplace:

$$P(A)=\frac{n}{m}=\frac{\text{estados que implican suceso A}}{\text{estados posibles}}=\frac{\text{casos favorables}}{\text{casos posibles}}$$

que mantuvo su predominancia en la teoría de la probabilidad durante aproximadamente un siglo desde la publicación en 1812 de su *Théorie analytique des probabilités;* hasta Markov (1912)[**21**] sigue enunciándola de esta manera. Y de esta manera sigue estudiándose en secundaria, aplicada a una clase particular, sencilla, pero muy restringida, de experimentos formados por sucesos equiprobables. No obstante, algunos autores [**20**] [**23**]sostienen que, anteriormente a Laplace, e incluso el mismo Laplace, mantenían posturas incoherentes sobre la cuestión, fueran cuales fueran sus matemáticas y sus declaraciones filosóficas:

> *Los matemáticos que trataron de medir estas probabilidades de alguna manera no arbitraria llegaron a encontrar tres métodos: probabilidades iguales basadas en simetrías físicas; frecuencias observadas de sucesos; y grados de certeza o creencia subjetivos;la primera casaba bien con los artefactos propios de los juegos de azar, como monedas o dados, pero poco más; los segundos dependían de la* ***recolección de estadísticas bajo supuestos de estabilidad a largo plazo****;y los terceros se hacían eco de la práctica legal al uso, de contabilizar en diversas proporciones los grados de certeza hasta llegar a la evidencia.*
> *Los probabilistas de tiempos posteriores ven estas tres respuestas a la cuestión "¿Qué es lo que miden las probabilidades?" como bastante distintas.* [...] *En particular, una clara frontera divide ahora los dos primeros significados "objetivos" de probabilidad, que corresponde a estados del mundo, y el tercer sentido "subjetivo", que corresponde a estados de la mente. Y sin embargo, los probabilistas clásicos usaban "probabilidad" para significar estos tres sentidos, pasando de uno a otro con una despreocupación que descoloca a sus sucesores, más despejados.*

¿Estamos seguros de que en las aulas de secundaria, pese a la preocupación por la forma y la aplicación correcta de algoritmos, no se deja al alumno a oscuras sobre esta misma cuestión? El hecho de que las interpretaciones bayesianas no alcancen estas enseñanazas no significa que el mismo alumno, arrastrado por el ambiguo significado de la voz "probable" en su uso común, no acabe por confundir los conceptos de probabilidad objetiva y subjetiva, sobre todo cuando se le obliga a "navegar" conceptualmente entre los dos mares vecinos, afines, (y a veces no bien avenidos) pero no coincidentes de la probabilidad y la estadística.

Pero pasemos a observar la relación desde el otro extremo (el de la estadística), donde también se observan unos comienzos independientes y un desarrollo paralelo, pero diverso, que transcurre durante largo tiempo desde que el comerciante John Graunt, en 1662, sólo ocho años después de la primera carta de Pascal a Fermat sobre el famoso problema de los puntos propuesto por De Mère, sólo cinco desde la publicación del primer tratado impreso sobre probabilidad[2] por parte de Huygens, fundara la estadística inferencial mediante la publicación de sus tablas de mortalidad[3], construidas sobre la *estimación* aproximada de que

---

dad de un suceso, sino en la inducción limitada sobre un conjunto de las observaciones disponibles y la asignación de una probabilidad a posteriori en función del conocimeinto que estos aportan (también los bayesianos reniegan del término "subjetivo", prefiriendo la categoría "epistemológico")

2 *De ratiociniis in ludo aleae* (Sobre el cálculo en juegos de azar) donde no se mencionaba el término "probabilidad", sino que se usaba el concepto de "esperanza" (matemática), más tarde relegado por la teoría probabilística, pero recogido por la estadística matemática.

3 *Natural and Political Observations Made upon the Bills of Mortality,* obra que le sirvió para ingresar en la Royal Society, frente a la oposición de la mayoría de miembros de la misma; al fin y al cabo, era un simple mercero.

la misma proporción de muertes antes de la edad de seis años tenían lugar cada década, estimación basada en los boletines de mortalidad (*Bills of mortality*) londinenses: es decir, basada en una serie de muestreos (desde 1629) de la población de Londres. Ya desde le primer momento se aprecia la diferencia de propósito: práctico, concreto, factual, predictivo de Graunt (alertar sobre posibles rebrotes de la peste) frente a las consideraciones teórico-académicas de los "filósofos" —hoy diríamos "intelectuales"— Pascal y Fermat, o los matemáticos Leibniz, Huygens, Bernoulli, etc. Graunt confesaba sin ambajes su condición de iletrado en matemáticas, y apenas nada fuera de simple aritmética hay en su obra; Jakob Bernoulli, en su *Ars Conjectandi*[4] (1713), por el contrario, declara la ilustrada intención de mostrar a los hombres, mediante el cálculo, que en realidad es el ejercicio de la pura razón, donde se encuentran sus verdaderos intereses: es decir, lo que él llama "arte de conjeturar" es en esencia lo que hoy se conoce como teoría de la decisión racional ante incertidumbre.Y sin embargo, en él incluye su famoso teorema fundacional, *el primero en relacionar probabilidad,* —que Bernoulli mismo entendía como grado de certeza—, *con frecuencia;* su misma formulación es clásica y sabidísima: tomada una urna con una cierta cantidad de bolas, de las cuales una proporción p tiene una característica singular (por ejemplo, el color), y de la cual se retira con reemplazo N veces una bola, se da que a medida que N es mayor, la *Probabilidad P* de que la proporción m/N de bolas singulares (coloreadas) entre el número de ensayos se iguale a la proporción efectiva p de bolas coloreadas en la urna se acerca más a la certeza, de modo que en el límite (N→∞) la probabilidad se hace total (1)[5]. El hallazgo conceptual es revolucionario, teniendo en cuenta el contexto intelectual de los *philosophes* y las propias ideas sobre la probabilidad de Bernoulli. No mucho después, en la 2ª edición de la otra gran obra teórica sobre probabilidad hasta Laplace, *Doctrine of Chances* de De Moivre (1738), se realiza otro hallazgo fundamental de la estadística matemática: la primera enunciación —habrá que esperar hasta Laplace para que lo demuestre y para que amplíe su alcance— del más tarde llamado (por Poisson) Teorema Central del Límite, que relaciona matemáticamente, no empíricamente como Graunt, los parámetros de la distribución de la *muestra* (que sigue una binomial en De Moivre, ampliada a la normal por Laplace) con los de la distribución de la *población* correspondiente mediante aproximación dada por una distribución normal[6].

Podría esperarse que, desde el momento en que se estableció esta relación fundamental, estadística y teoría de la probabilidad quedarían irrevocablemente unidas. Lejos de ello, ambas se desarrollaron bajo premisas heterogéneas, y sus caminos tardan mucho en encontrarse. Durante todo el siglo XVIII se consideraba a la estadística (y de ahí su nombre) como equivalente a "aritmética política", y se relacionaba con la teoría de la probabilidad sólo indirectamente, en la medida en que aquella decía ofrecer instrumentos acerca de toma de decisiones "racionales", las cuales podían convertirse en prescripciones de carácter político. No obstante, los mismos estadísticos, si bien cada vez más atentos a las cantidades y la medición y menos a sacar conclusiones políticas, dieron de lado a la teoría de la probabilidad durante todo el siglo XVIII —por lo menos— debido principalmente a que la estadística se consideraba a sí misma una disciplina objetiva y ahistórica, hiperpositivista, basada en "pu-

4 *Arte de conjeturar*, un título que le acerca bastante al campo subjetivo, si bien sobre la base de una intersubjetividad de la razón, al menos entre los hombres "ilustrados".

5 La expresión analítica según el criterio actual es:

$$\lim_{n\to\infty} P\left(\left|p-m/N\right|<\varepsilon\right)=1 \quad \forall\varepsilon>0$$

6 El resultado fue generalizado para *todo tipo de distribución poblacional* por Lindeberg y Lévy (independientemente) en los años 20 del siglo pasado. Para una exposición del Teorema de de Moivre, *vid.* **[12]**, p. 198, por ejemplo.

ros hechos", mientras que la teoría de la probabilidad era contemplada como un ejercicio *de especulación teórica,* y no sólo eso: como una teoría antagónica, ya que su objeto de estudio era el azar o *chance* (hasta Laplace, la teoría de la probabilidad todavía era la *Doctrine of Chances)*, lo cual iba en contra de todos los descubrimientos estadísticos de la época, que venían a demostrar exactamente lo contrario: no el azar, sino la *regularidad*, una regularidad tanto más perfecta cuanto mayor es el número de datos recogidos, era la base de los fenómenos variables y la razón de ser de la recogida y ordenamiento de datos estadísticos. Baste para ello mencionar los títulos de las obras respectivas de John Arbuthnot: *Una prueba de la Divina Providencia tomada de la regularidad constante observada en el nacimiento de ambos sexos* (1710); y Johann Peter Süsmilch: *El Orden Divino en las circunstancias del sexo, el nacimiento, la muerte y la reproducción humanas* (1741)[7], ambos teólogos "naturales" y precursores de la demografía, y cuyo interés fundamental en la regularidad de los fenómenos observados, como las tasas de mortalidad, nacimientos, casamientos, suicidios, etc. consistía en probar mediante su recogida que, bajo la apariencia aleatoria, se escondía un orden cósmico promovido a instancias de la divinidad, un orden negador, en última instancia, de la casualidad, azar o alatoriedad. Que la obra de Bernoulli, De Moivre y la posterior reformulación general de Laplace (que incluía la probabilidad condicional y el Teorema de Bayes, entendido "racionalmente") *venían a darles, efectivamente, la razón*, parece no haber hecho mella en los estadísticos profesionales hasta la década de los 30 del siglo XIX.

Lo cierto es que la estadística del siglo XIX ignoraba casi completamente la inferencia puramente probabilística (matemática), es decir, no extraía conclusiones sobre una población mediante el análisis de muestras de la misma. En las raras ocasiones en las que iba más allá de la recogida de extensas cantidades de datos (casi siempre mediante censos, cuanto más completos, mejor), lo que practicaron asiduamente fue la inferencia de las causas de la regularidades que observaban. Pronto descubrieron que una nación podía ser descrita y diferenciada de otra por un número no muy grande de variables que mostraban frecuencias relativamente estables, de acuerdo a categorías tales como las tasas de criminalidad, suicidio o matrimonios. Estas regularidades eran "los hechos" sobre los que se podía fundar alguna clase de conocimiento; su objeto, además, se presentó pronto como algo más fundamental todavía que el "estado" que había dado nombre a la disciplina "estadística", pasando a ser la "sociedad". El individuo era, si bien no completamente aleatorio en sus actos y naturaleza, demasiado variable al fin y al cabo; las sociedades, en cambio, eran suficientemente estables como para ser estudiadas. Se hacía alusión episódicamente a una versión muy vaga de la ley de los grandes números, en el sentido primordial de Bernoulli, pero esa era toda la relación de la estadística con la probabilidad hasta el segundo cuarto del siglo XIX[**20**].

Por otra parte, el descubrimiento de que incluso actos considerados hasta el momento como puros exponentes de irracionalidad, y del arbitrio del azar, como los crímenes violentos o las revoluciones políticas, eran no obstante regulares en su camportamiento como agregados estadísticos llevó a la creencia de que, al igual que en la mecánica celeste, no había lugar en el universo para la casualidad, sino que el orden imperaba por doquier: solo era cuestión de empeño buscar las leyes que gobernaban ese orden: Quételet, en 1831, trató de fundar la por él lamada Física Social basándose en tales consideraciones. Su concepto central, *l'homme moyen* (el hombre medio), y no el hombre racional de la teoría probabilística

7 Respectivamente, *An Argument for Divine Providence Taken from the Constant Regularity Observed in the Birth of Both Sexes,* y *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts aus der Geburt, dem Tode und der Fortpflanzung desselben*

clásica (todavía muy relacionada con la teoría de la decisión; hasta el mismo Laplace escribió que: "el cálculo de probabilidades no era más que el sentido común expresado en números")era el prototipo de la especie, y la piedra angular sobre la que se basaría la nueva ciencia. Si bien sus esperanzas se vieron frustradas, no es menos cierto que Maxwell y Boltzmann, requeridos en defensa de su termodinámica , adujeron los mismos argumentos y pruebas de Quételet —entre otros— como ejemplos de que el comportamiento aleatorio de una masa de millones de partículas es mucho más estable y su conocimento mucho más seguro que aquél que se pueda construir sobre partículas individuales consideradas por separado **[20]** **[23]** .

Curiosamente, fueron las implicaciones morales y políticas de las ideas de Quételet y sus seguidores las que atrajeron, por vez primera, los recursos de la teoría de la probabilidad a la disciplina estadística. Algunos consideraban que bajo las ideas de Quételet sobre "leyes sociales" se encontraba un fatalismo que negaba la libertad humana; otros, que la existencia de estas leyes irresponsabilizaba, en cierta medida, al individuo: muchos objetaban que la variabilidad era, a pesar de todo, más definitoria del comportamiento humano que las simples medias. Uno de aquéllos, el alemán Wilhelm Lexis, fue de los primeros en estudiar exhaustivamente la dispersión de los datos; su pretensión primera fue la de buscar los "esquemas de probabilidad" de las conductas sociales. Posteriormente realizó uno de los primeros análisis serios de varianza. El planteamiento era el siguiente: si las leyes sociales eran verdaremente alguna cosa, entonces, al estudiar series temporales de ciertas variables importantes, las desviaciones respecto de las frecuencias medias reflejadas deberían deberse *completamente al azar*, es decir, deberían ser aleatorias. Su primer ensayo, acerca del sexo de los recién nacidos, dio resultado positivo: ahora bien, esto era un hecho natural, más que social. Inmediatamente probó otro número de variables "morales" (suicidios, crímenes,matrimonios, etc.) y encontró que las fluctuaciones de las series eran *mucho mayores que las que cabía esperar solamente del azar*: lo cual resultó en el abandono de la idea de "leyes sociales" ahistóricas, y, de manera más importante a nuestro asunto: incluyó el análisis de la dispersión como elemento esencial de la estadística y la teoría de la probabilidad como herramienta mediante la cual contrastar hipótesis.

A los efectos de la estadística, más que de la investigación social, aún no hemos hablado del mayor mérito de Quételet. No hay que olvidar que algunas de sus aportaciones son para el campo de la *astronomía,* ciencia que fue a estudiar a París en 1823. Allí aprendió que se utilizaba cierta fórmula para calcular la distribución de los errores respecto de los valores verdaderos en mediciones astronómicas. La fórmula en cuestión era:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{\left(-x^2/2\right)}$$

esto es, la hoy conocida como función de densidad de probabilidad de la *distribución normal estándar* (normal con media 0 y varianza 1), que, como ya hemos visto, introdujo De Moivre por primera vez en referencia a la distribución de muestras de una población con distribución binomial. No obstante, quien la popularizó fue el omnipresente Gauss, tras haberla deducido por su cuenta, de una forma más o menos heurística, del método de mínimos cuadrados[8] para el cálculo de errores en mediciones astronómicas, es decir, no inmediatamente relacionado con asuntos probabilísticos (relación que forjó más tarde Laplace,

8 Cuya invención se disputa con Legendre, quien publicó (y nombró) el método en un anexo de sus *Nouvelles Méthodes pour la determination des orbites des comètes* (1805), es decir, antes que Gauss (1809)

en 1812). Pues bien, Quételet, que concebía su hombre medio como una forma original, como un tipo (distinto para cada nación, eso sí) respecto del cual las demás manifestaciones individuales eran *desviaciones*, no tardó en relacionar la por entonces llamada curva de errores[9] con la distribución de las frecuencias de sus tablas sobre mortalidad, nacimientos, criminalidad, etc., relación que estableció en 1844. La estadística matemática estaba, en cierto modo, fundada desde ese mismo momento.

Pero tuvo que ser Galton quien recogiera la idea y expandiese sus aplicaciones a todo tipo de mediciones (algunas considerablemente extravagantes). Galton estaba primordialmente interesado en la eugenesia, y sus ideas sociales partían de cierta interpretación personal (y ciertamente aristocrática) de la teoría de la evolución recién enunciada por su primo, Charles Darwin. Ello no le impidió tomar la idea de Quételet, aunque dándole un sesgo peculiar: para él, el hombre medio no era un prototipo, sino simplemente, la clase más abundante de individuos. En su obra *El Genio Hereditario*[10] (un título que por sí mismo lo dice todo) sostiene que (tomado de **[20]**, p.):

> *Lo escaso de la capacidad para asumir el mando, junto con la vasta abundancia de la mediocridad, no es accidental, sino que sigue la ley de la necesidad propia de estas cuestiones* [trad. propia].

Galton asumió que todas la cualidades humanas, imaginables e inimaginales, se hallaban distribuidas a través de la especie según curvas normales propias. Hombre de una envidiable capacidad para el entusiasmo, aunque de un gusto algo dudoso para la metáfora, llegó a escribir de la distribución normal que "los antiguos griegos la hubieran personificado y deificado, si la hubieran conocido".

Pero lo que nos importa es su aplicación de los instrumentos matemáticos a la inferencia estadística, dentro del campo de la biología evolutiva —en su variante eugenésica—, camino que seguirán muchos otros de sus sucesores, desde Pearson hasta Fisher. Tras tentativas diversas, fue el primero en abordar cuestiones de correlación entre variables por medios probabilísticos: en concreto, la determinación de rectas de "regresión"[11] (a la mediocridad, su nombre completo original), que en realidad podría llamarse "convergencia a la media" ya que su propósito era relacionar cada valor de la variable escogida como independiente (o explicativa) con la *media* de todos los valores de la variable dependiente (o respuesta) que se corresponden con aquél. Representados sobre una gráfica, estos valores medios se acercaban a la forma de una recta[12] para una cantidad tan grande de variables así conjugadas, que el mismo Galton dio el atrevido salto conceptual de pensar que sus rectas de regresión no se aplicaban sólo a los índices biológicos o las características herediatrias de las especies, sino que podían emplearse para relacionar cualquier par de variables en cualquier campo de estudio. Y su visión tuvo éxito, en un primer momento: múltiples disciplinas rebosaban de fárragos de datos sin analizar, y las solas cifras parecían cada vez más insuficientes: era preciso indagar sobre las posibles relaciones de cada categoría recogida. Así, campos tan dispares

9 Más tarde "ley normal", para después pasar a ser la curva de campana o campana de Gauss (a despecho de De Moivre y Legendre) (*vid.* **[20]** , p. 53)

10 *Hereditary Genius: an Inquiry into its laws and consequences* (1869)

11 Por razones de precedencia lógica y facilidad didáctica, suelen presentarse en primer ugar el concepto de correlación entre variables y después los métodos para hallar rectas de regresión (sobre el modelo lineal simple); pero el desarrollo histórico de los conceptos fue a la inversa. Los coeficientes de correlación en su forma actual se los debemos a Pearson, alumno y sucesor de Galton.

12 En su obra *Natural Inheritance* (Herencia Natural), 1889. La línea de regresión se determina mediante el método gaussiano (o Legendriano) de los mínimos cuadrados. No nos detenemos en el particular por razones de espacio.

como la sociología, economía, medicina, agricultura, etc. comenzaron a utilizar los métodos promovidos por Galton en vistas a formalizar matemáticamente sus disciplinas.

Pero, quizá se pregunte el lector, mientras tanto, ¿Qué hacían los probabilistas?

Desde que la obra monumental de Laplace instituyera los principios ya vistos, junto con algunas adiciones de Poisson en sus *Investigaciones sobre la Probabilidad en los Juicios*[13], continuistas desde el punto de vista de la interpretación "epistemológica" de la decisión racional (y que fueron durísimamente criticados; es la época de Cauchy y el nuevo énfasis en el rigor de la conceptualización matemática), se vieron en un callejón sin salida: o mejor dicho, con una única salida: la interpretación frecuentista, conducente a la estadística matemática. La visión clásica recibió el oprobioso mote de "subjetiva", y con la definitiva quiebra de la confianza en la razón ilustrada, tal y como estaba planteada como teoría de la decisión basada en grados de creencia, se convirtió muy pronto en inaceptable desde el punto de vista matemático y filosófico. Por otra parte, el modelo basado en la regla de Laplace para sucesos equiprobables demostró tener aplicaciones severamente restringidas (casi exclusivamente en juegos de azar). Sólo el método de los mínimos cuadrados y sus implicaciones sobrevivieron, como ya se ha dicho, para que la estadística matemática se fundara sobre su base. En la segunda mitad del siglo XIX la teoría de la probabilidad atraviesa un hiato, mitigado tan solo por la escuela de probabilistas rusos de S. Petersburgo apadrinada por Chebyshev. No obstante, en esta escuela se forjaron los cimientos de la que acabaría renaciendo como teoría matemática aceptable: primero, con el nuevo enfoque sobre los procesos estocásticos con las cadenas de Markov (1912), pasando por otras muchas aportaciones, como la del mismo Chebyshev, Bernstein, Liapunov, Kinchine, hasta la llegada de la trascendental *axiomatización de la probabilidad* llevada a cabo por A.N. Kolmogorov en 1933, antes de la cual la teoría de la probabilidad ocupaba una tenebrosa tierra de nadie entre las matemáticas y la filosofía, pero sin pertenecer a ninguna de las dos. Tras esta fundamental aportación, que puso "en el mapa" de las matemáticas a la teoría de la probabilidad, proscrita desde hacía casi un siglo, y ayudada por las nuevas concepciones en física (termodinámica estadística, movimientos brownianos, mecánica cuántica) que concedieron a la probabilidad , no un papel meramente ancilar, sino un lugar en el centro mismo de la teoría (según algunos, de la misma realidad), la teoría de la probabilidad avanzada recibió un impulso espectacular y un crecimiento exponencial consecuente: pero considerar estos progresos está claramente fuera de los objetivos del presente trabajo y del currículo de secundaria.

Así pues, solo nos queda proseguir con la estadística, que además de inferencial, se hacía cada vez más matemática con la generación siguiente a la de Galton: Edgeworth, Karl Pearson, Yule. En concreto, Karl Pearson desarrolló y "matematizó" el concepto de correlación ya implícito en las rectas de regresión investigadas por Galton, de quien fue sucesor directo y primer ocupante de una "cátedra de eugenesia" fundado con parte de la herencia de Galton en la Universidad de Londres, sobre cuya base Pearson fundó el primer departamento exclusivamente dedicado a la estadística aplicada. Aunque sus intereses estaban del lado de las aplicaciones (sobre todo a la biología), puede considerarse a Pearson como el verdadero fundador de la estadística matemática: fue él quien estableció las bases conceptuales (probabilísitcas) de la inferencia paramétrica: estimación por el método de los momentos, definición por medio de parámetros de los modelos de distribución poblacional, coeficiente r de correlación, clasificación y generalización de las distintas funciones correspondientes a

13 *Rerecherchés sur la probabilite des jugements,* 1837

modelos útiles para la inferencia, el primer acercamiento al concepto de p-valor, test de ajuste a los modelos de distribución (chi-cuadrado), entre otros. La posición central de Pearson en la disciplina y su amplio espectro de intereses le llevaron a prefigurar, en cierto modo, el diseño de experimentos; fue él quien apoyó y ayudó a la formalización matemática de la distribución t de Student, hallada por Gosset para resolver el problema de inferencia en base a muestras pequeñas, un problema de planteamiento y ejecución de experimentos, además de análisis de datos.

Pero el verdadero fundador, por derecho propio, del campo del diseño de experimentos aleatorios fue Ronald Aylmer Fisher, cuya labor en estadística da comienzo en 1912, y transcurre sobre los hitos que constituyeron sus obras *Métodos estadísticos para investigadores*, y *El diseño de experimentos*[14]. En ellos se definen las características esenciales y los métodos a llevar a cabo para la experimentación de fenómenos no deterministas tal como aún se observan hoy en día: formación aleatoria de grupos de experimentación y control, eliminación de factores de influencia, experimentos dobles ciegos, y test de significación (mediante el hallazgo del p-valor —introducido por Pearson— en orden a confirmar o refutar la hipótesis inicial o nula). Su obra parte de las consideraciones prácticas de la ejecución de experimentos o estudios observacionales, y de intereses científicos (agricultura, biología evolutiva, genética, etc. contempladas de nuevo desde el punto de vista eugenésico), en función de los cuales aborda las matemáticas. Su obra constituye, por tanto, otro giro "aplicado" de la estadística inferencial, que se acentúa con las posteriores correcciones de sus tesis elaboradas por Neyman y E. S. Pearson, quienes introdujeron —corrigiendo la concepción de Fisher— la estimación por intervalos de confianza y de una nueva concepción del contraste de hipótesis (errores de tipo I y II, potencia de contraste)[15]. Sobre estos últimos desarrollos de la inferencia (paramétrica) cabe comentar el hecho, de consecuencias para la educación estadística —y no solo en enseñanzas medias— que Gigerenzer *et. al.* **[20]** designan como "modelo híbrido" de la enseñanza de estadística; según esta visión, ambas escuelas "objetivas", la de Fisher y la de Neyman-Pearson, conformadas en torno a procedimientos distintos (test de significación en un caso, intervalos de confianza y estudio de la potencia de contraste en el otro) responden también a presupuestos inferenciales heterogéneos, —con consecuencias sobre la misma actividad científica, graves según algunos, teniendo en cuenta que la convalidación de muchos experimentos depende decisivamente de la elección de uno de los modelos de contrastación a debate— y, en el fondo, incompatibles, cosa que no se refleja en los textos introductorios de estadística en secundaria (y, según nuestra experiencia, tampoco en muchos de los correspondientes a los primeros cursos universitarios). ¿Es esto hilar demasiado fino en el contexto de la educación secundaria? Es posible: máxime teniendo en cuenta que el modelo de Neyman-Pearson parece haberse impuesto definitivamente en el currículo de enseñanzas medias (según la LOMCE**[3]**): no obstante, no es poco probable que, dada la inercia del profesorado y de los textos escolares existentes, amén de la volatilidad de los diversos planes de estudios de enseñanzas medias españoles, el alumno se encuentre con ambas concepciones. Posteriormente, según su elección universitaria, puede tener que ver con ambos al mismo tiempo; por lo cual, podría ser interesante, más que la elección de

14 *Statistical Methods for Research Workers* (1925), y *The design of experiments* (1935)

15 El desarrollo paralelo y contemporáneo de la escuela "subjetiva", o inferencia Bayesiana (Jefferys, Savage, De finetti, *et al.*)basada en la elección del modelo de distribución poblacional *condicionada* a la información disponible (limitada) sobre esa población, y que en los últimos tiempos ha pasado de ser una postura más o menos "herética" a cobrar un gran auge en determinadas ciencias humans y aplicadas (economía, epidemiología, sociología, psicología) no se trata en el presente trabajo, dado que el curriculo de enseñanza secundaria (y no escasa parte de la universitaria)no lo abarca (por ahora).

principio entre uno y otro, proponer el planteamiento y discusión de ambos puntos de vista inferenciales en las mismas aulas, asociados a las diferentes aplicaciones de los mismos, y a los distintos resultados arrojados.

Valga este comentario, en el cual quizá nos hemos extendido demasiado, como excusa para detenernos en este punto y señalar que este carácter híbrido recorre el tratamiento curricular del "bloque" entero de probabilidad y estadística (últimamente reformado como "estadística y probabilidad", y no por accidente). Que estos apuntes no pretenden ser ni exhaustivos ni especialmente lúcidos, sino más bien *expresivos* de las relaciones problemáticas existentes entre las dos ramas de estudio tan comúnmente aparejadas, y no obstante tan disímiles en muchos sentidos, como decíamos al comienzo del capítulo. Que esta divergencia ha afectado y afecta a la enseñanza de estas materias en secundaria a lo largo de los diversos planes de estudio, lo ha constatado el autor mismo de estas líneas, por observación personal y por testimonios ajenos: perteneciente a la generación de la EGB, el BUP y el COU, recuerda perfectamente (sin necesidad de analizar los contenido en la ley misma : ¡Tan fácilmente transgredidos y mutilados en la práctica!) que, todavía en la fase obligatoria (EGB) comenzaba a estudiarse combinatoria, aplicada a problemas de contaje exclusivamente, sin ninguna referencia a la probabilidad. Posteriormente, en BUP, en la opción A llamada científico técnica (Matemáticas I), no volvió a saber nada de la materia, estando los contenidos casi completamente absortos en el análisis, con algo de geometría analítica y álgebra, para, después, en COU, tratar somerísimamente, casi a las puertas de selectividad, conceptos primordiales, básicos de probabilidad, como la regla de Laplace, la definición de probabilidad condicionada y el teorema de Bayes; por supuesto, absolutamente nada de estadística, que solo llegaban a conocer (también tarde, también a trasmano) en COU los alumnos que escogían Matemáticas II (correspondientes a las opciones C:ciencias sociales y D: humanísitca/linguística), ya que en esos itinerarios las matemáticas eran *opcionales*. De forma y manera que una gran parte de los alumnos ni siquiera rozaban la estadística, quedando para los de las opciones A y B (esta última, Biosanitaria) una fárrago deslavazado de teoría de la probabilidad. El lector interesado puede comprobar que las experiencias relatadas se deducen directamente delos contenidos curriculares del plan correspondiente (LGE **[1]**, y sus respectivos desarrollos para educación General Básica y Bachillerato)

La tan denostada LOGSE **[2]** puso, no obstante, a la estadística en el mapa curricular de forma más seria, y las posteriores reformas no han hecho sino acentuar su presencia, hasta esta última ley de inminente aplicación, cuyos contenidos para 2ª de Bachillerato en la opción de ciencias sociales, al menos en principio, podrían corresponderse con los del comienzo de un semestre universitario. No obstante, el desfase en su introducción y formalización curricular con las otras ramas clásicas del currículo de matemáticas como el álgebra, el análisis, la geometría sintética y analítica, etc. es apreciable. ¿A qué puede ser debido este hecho?

Por un lado, la disciplina misma de la teoría de la probabilidad ha tenido, si se compara con el de otras áreas matemáticas, un desarrollo tardío y accidentado, sometido a contingencias varias presentadas por el contexto social, ideológico y moral de su época, así como por el estado de desarrollo de disciplinas en primera instancia ajenas a la misma, como la estadística primitiva (no inferencial), las ciencias sociales, la biología y posteriormente, los avances "posclásicos" de la física; lo que ha provocado paréntesis profundos en la definición de su objeto, junto a profundos "cambios de rumbo" en su fundamentación como teoría matemática, por no decir nada de sus implicaciones filosóficas, aún muy discutidas. Sus dificultades comienzan desde el simple hecho de ser aceptada como área matemática válida: que esta aceptación no siguió inmediatamente a la publicación por parte de Kolmogorov de su axiomática lo averiguamos en el clásico de Feller (por ejemplo), el cual, en la primera frase del primer parrafo del prólogo a su segunda edición en inglés, (1967) declara (*vid.* **[17]**, p. 7):

> *Cuando se concibió este libro por primera vez (hace más de 25 años), había* ***pocos matemáticos****, fuera de la Unión Soviética, que* ***reconocían la teoría de las probabilidades como rama legítima de las matemáticas****. El panorama era limitado, y al trabajar con problemas individuales se solía llegar a complicaciones increíbles. Debido a estas circunstancias, el libro no se escribió* [...] *sino con la esperanza de atraer atención sobre aspectos de la probabilidad poco conocidos, de forjar vínculos entre varias de sus partes, de desarrollar métodos unificados y de indicar aplicaciones potenciales. Debido al interés cada vez mayor por la teoría de las probabilidades, el libro encontró muchos lectores inesperados fuera de las disciplinas matemáticas.*

Los planes de enseñanza, que son los últimos en recoger los cambios programáticos de las ciencias, los subsiguientes estudios en la teoría didáctica correspondiente, etc. por razones evidentes, todavía se resentían, en nuestra opinión, del efecto de esta tardía incorporación de la probabilidad al *corpus mathematicum,* del mismo modo que cierto miembro nuevo de un club tiene que sufrir durante algún tiempo el "vacío" contra él practicado por mucho exponentes de la "vieja guardia", de los campos matemáticos clásicos.

Asimismo, las aplicaciones de la teoría de la probabilidad *que no tienen que ver con la estadística* (nos permitimos enfatizar el anterior subrayado), no son inmediatas en la mayor parte de las ciencias por sí mismas, sino sólo en unas muy concretas especialidades, a menudo solo para cierta clase restringida de problemas. De tal modo que su enseñanza fuera de un contexto estadístico parecía fútil, o cuando menos, irrelevante respecto, por poner el caso, el análisis o el álgebra lineal, de aplicación directa en ciencias e ingenierías. La única forma de encontrar acogida en las enseñanzas medias era vincularla con la estadística, al menos la estadística matemática. Pero si ese era más menos el status de la teoría de la probabilidad, ¿qué decir de la estadística matemática? Para muchos, probabilistas incluidos, no es sino una rama aplicada del análisis, y no de las más interesantes (*vid.* el comentario de David Aldous referido por Moore en **[10]**, p. 803). De la estadística inferencial, ya no había ni que hablar, pues no pertenecía a las matemáticas de ninguna de las maneras. De hecho, al menos en España, tenía problemas para consitutirse como disciplina independiente. Eran consideradas como herramientas ancilares multiusos: así se enseñaban en las carreras de humanidades, o ciencias sociales, en biología, medicina, economía, etc. En esto el caso español también es significativo de un cierto retraso del contexto internacional: no obstante, este hecho ha cambiado profundamente, debido a las causas que se expondrán en la última parte de la introducción, y que tienen que ver con el contexto actual científico y tecnológico, y que ha determinado el mismo estatus profesional (al menos, en los países punteros en ciencia y tecnología) del estadístico. Por el momento, en relación a la exposición anterior, puede afirmarse que la obra que parte de Galton y culmina con Fischer y sus émulos dio un impulso inusitado a toda clase de experimentación en contextos no deterministas. Ello no podía sino acrecentar la importancia de la labor del estadístico, como teórico y como profesional independiente, en multitud de ciencias humanas como las antes reseñadas, y en las cada vez más importantes aplicaciones en ingeniería (optimización, control de calidad, etc.)

Que ello requería una independencia y seguridad profesional para el estadístico lo señala el mismo Fisher [**18**]:

> *Cuando se llama al estadístico una vez concluido el experimento, se actúa como cuando se llama al médico forense; lo máximo que podrá hacer el estadístico será indicar la causa de la muerte [del experimento]*[trad. propia]

Las necesidades de las ciencias sociales y biomédicas, y de las industrias afines, elevaron el papel del estadístico como figura profesional, e influyeron en el diseño de los estudios correspondientes: no obstante, el paso decisivo de la estadística matemática e inferencial clásica al análisis exploratorio de datos, la inferencia no paramétrica, la estadística descriptiva, la utilización preferente de estadísticos robustos, etc. y, sobre todo, su decisiva influencia en la didáctica y en la enseñanza práctica de la estadística, vendrá dada por un cambio todavía más fundamental que aún está en proceso a día de hoy, y que dejamos para la tercera parte de la introducción. Pero antes debemos pasar a exponer las dificultades cognitivo-epistemológicas relacionadas con la incertidumbre, que incluye conceptos probabilísticos y estadísticos, y que, aunque pueden deducirse a su vez de lo conflictivo del desarrollo de ambas teorías expuesto anteriormente (problemas en la interpretación de la probabilidad condicional, confusión entre frecuencia y probabilidad, erróneo entendimiento de la ley de los grandes números, percepciones distorsionadas de la dispersión de una distribución, etc.) creemos que merecen una exposición aparte, por su carácter intrínsecamente cognitivo, más que histórico, por su general extensión a toda clase de indivuduos, educados matemáticamente o no, y, más importante que nada, su sistemática desatención desde el punto de vista de la enseñanza práctica, que no de la didáctica.

Dificultades cognitivas

Veremos en una de las actividades propuestas que, lejos de resultar *demasiado intuitivas* —como a menudo se consideran las cuestiones de la disciplina desde ciertas alturas— grandes intelectos (D'Alembert, Leibniz, etc.) han hallado su peculiar talón de Aquiles matemático en problemas probabilísticos de desarmante simplicidad (*vid. infra,* pp. 103, 131). Sin embargo es de señalar en descargo de D'Alembert que se detecta en sus textos una beligerancia efectiva (constatada por la historia) contra la misma idea de fundar una ciencia sobre el azar que pudo hacerle perder el norte. No es el caso de lo que sigue a continuación.

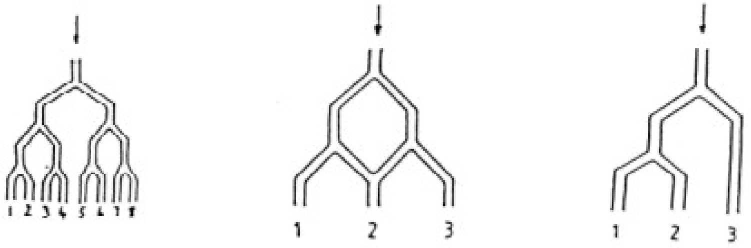
En primer lugar cabe considerar los estudios, pertenecientes al campo de la psicología cognitiva desde elpunto de vista filogenético: esto es, las que buscan en los primeros estadios del desarrollo intelectual del niño la raíz de las posteriores confusiones (o errores aprendidos, como sotiene uno de los autores). Siguiendo a Batanero **[6]**[1]**,** encontramos las tomas de postura principales en los autores siguientes:

·Piaget e Inhelder: Según su punto de vista, el niño (es decir, antes del período de las operaciones formales, que los autores estiman da comienzo aproximadamente a los 11 años) no está dotado de intuiciones primeras correctas sobre el azar o la probabilidad. En primer lugar, no diferencian bien las situaciones deterministas de las aleatorias. El niño, en la etapa preoperacional (2-7 años) no entiende bien que una combinación de diversas causas puede dar lugar a un resultado imprevisible, el cual además no puede ser revertido haciendo operar estas mismas causas en sentido contrario. Batanero parece tener razón al señalar que los mismos autores parecen tener una idea muy determinista del azar, entendido como tal el efecto incontrolable de un número dexcesivo de causas —idea análoga ala de la ilustración— , de las cuales solo puede calcularse la cantidad de estados posibles a los que pueden dar lugar. Sea como fuere, (lo que abogaría, indirectamente, por la tesis del fallo intuitivo en probabilidad, ¡Por parte de los mismos analistas de estos errrores!), como el niño no posee una intuición precisa de la combinatoria subestima el número de estados posibles y confía siempre en a reversibilidad de un desorden creado por causas externas. Si esto no sucede, atribuirá la causa del hecho a una instancia exterior, del mismo modo que, para el hombre primitivo, todo fenómeno estaba dirigido, determinado, por la divinidad (o una divinidad conreta para cada tipo de fenómeno)y no había lugar para el azar; hasta la suerte era asunto de los dioses. En el período de las operaciones concretas (7-11 años), aunque comienza aentenderse la irreversibilidad de los fenómenos debidos a cadenas causales, sigue sin comprenderse la independencia (o la dependencia) entre fenómenos, lo que da lugar a especulaciones sobre causas bastante arbitrarias. Según estos autores, solo en elperíodo de operaciones formales, tras los 11 años de edad, comien-

1 Nos hubiera gustado, como en elresto de casos, haber consultado las fuentes primarias y realizar citas de primera mano: no obstante, nos ha sido imposible encontrar las obras de los autores que la autora describe.

za a formarse la idea de cálculo combinatorio, y con la idea de que sólo el conjunto de estados posibles es manejable desde el punto de vista teórico.
Asimismo realizaron experimentos para detectar la intuición de la ley de los grandes número en niños; se les pidió a los mismos que señalasen sobre un papel cuadriculado en qué lugar caerían las gotas de lluvia asumiendo que el papel es un suelo embaldosado. Los niños de edad más temprana disponen las gotas de modo sistemático, rellenando cada vez una baldosa: sólo a medida que la edad va aumentando las disposiciones se vuelven más irregulares, aunque generalmente no los suficientemente irregulares: hay períodos más cortos en estados con el mismo número de gotas en cada baldosa de lo que cabría esperar de una distribución aleatoria (en este caso, de Poisson). Se podría objetar a estos autores que la generalidad de los adultos tampoco poseen una idea “natural” de la ley de los grandes números, como veremos más adelante.

·Fischbein: Frente a estos autores se coloca Fischbein, para quien las intuiciones básicas del infante sobre probabilidad son aproximadamente adecuadas, pero pueden echarse a perder sin una eduación consecuente con la existencia de fenómenos aleatorios. Según este autor, la educación, y quizá la misma cultura (occidental) adolecen de un cierto sesgo determinista que la investigación especializada en materia deprobabilidad no atenúa en el ciudadano no experto. Para empezar, el niño menor de 7 años es capaz de discriminar la probabilidad de sacar una bola de determinado color con los ojos cerrados en una urna con más casos favorable que en otra con menos, siempre que el número sea pequeño. Esta capacidad no debe desatenderse, o la cultura en la que está inmerso le condoicionarñan a buscar causas a toos los efectos, y a formarse una idea de la mala o la buena suerte, la cual, según este autor, no es fundamental en el hombre, sino creada, derivada. De igual manera comprobó que el niño se adapta a un fenómeno que marca frecuencias relativas marcadas, y que esta operación mejora a medida que crece, aunque en mayor medida cuanto más cotidiano es el fenómeno y más consecuencias prñacticas tiene para el sujeto.
Según este autor, la educación probabilística en el niño es factible: se puede incluso promover la intuición de la convergencia a la distribución normal (teorema central del límite) con esquemas simplificados. Por ejemplo, pare este autor el famoso Quincux de Galton, es decir, el dispositivo que distribuye bolas en casilleros según una distribución normal, por el simple expediente de arrojarlas sobre un plano dotado de unos vástagos cuyos intersticios tienen la misma probabilidad de ser recorridos por las bolas es, debido a su mecánica, conceptualmente complicado para los niños: pero se puede optar por simplificaciones como los canales propuestos por Green[1]:



Fuente: **[6]**

1 Esquemas tomados de **[6]**

De los cuales, como se comprobará, el de la izquierda ofrece casillas de salida equiprobables, pero no así los otros dos: discutir con el alumno (Y comprobar mediante un artilugio sencillo) que, por ejemplo, en el esquema central la casilla de salida número dos recibirá el mismo número de bolas que las casillas 1 y 3 juntas desde la etapa preoperacional (antes de los 7 años) es factible para nuestro autor, y fomentará que posteriores equívocos no tengan lugar en el adulto. Caso contrario, cuando socialmente haya "aprendido" a considerar todos los sucesos como determinados, deberá desaprender estas nociones para construir el conocimiento verdadero en su lugar.

Mas sea cual sea el origen —innato o aprendido— de estos fallos de la intuición de fenómenos aleatorios, lo cierto es que son muy abundantes y están generalizados al conjunto de la población, incluida aquella que posee formación superior e incluso formación estadística: el análisis detenido y el cálculo la eliminan, pero las reacciones instantáneas en contextos cotidianos y con información limitada demuestran poseer los mismos sesgos en toda clase de indivduos. Los ya clásicos Kahneman y Tversky [**29**] realizan una taxonomía exhaustiva de estos fallos, o, más bien, de las heurísticas seguidas que conducen a tales fallos, agrupados en categorías generales, que pasaremos a repasar someramente:

· <u>Representatividad:</u> Según estos autores, bajo esta categoría incluyen todas las heurísitcas (sesgadas) que concurren cuando se trata de evaluar la pertenencia de cierto objeto o suceso a una clase determianda; en otras palabras, la clase de errores que se cometen cuano se pregunta: ¿Qué probabilidades hay de que el objeto i pertenezca a la case A? Los autores detectan que el grado de representatividad: es decir, la facilidad con la que se mustra a la imaginación la información pfrecida influye decisivamente en las respuestas a la anterior pregunta. En preguntas sobre personas se detecta un sesgo en las respuestas que responde a un estereotipo cultural, cuya "representatividad" es mayor que optras consideraciones; por ejemplo, cuando, en base a la descripción de un individuo, se pide a los sujetos respondentes que evalúen la probabilidad de que pertenezca a una serie de profesiones determinadas. La fuerza del estereotipo es tal que, pese a probablemente la experiencia del sujeto no le ha informado en ningún sentido sobre el particular, seguirá juzgando que una persona "tímida, ordenada y pasiva" ostenta una probabilidad mayor de ser "bibliotecario" a ser "granjero" o "piloto de avión". Hasta aquí los autores solo han detectado la fuerza de la ideología cultural subyacente; lo interesante viene cuando los sujetos son a la vez informados de las frecuencias relativas de cada pofesión en relación a la población a la que pertenecen: pues bien, se demuestra que esta nueva información no provoca que los sujetos construyan una probabilidad a priori o inicial en base a esos datos; el sesgo del estereotipo se mantiene. Los autores denominan al fenómeno con la etiqueta de "insensibilidad al efecto de la probabilidad a priori". No es la única de tales subclases. También se analizan casos correspondientes a "insensibilidad al tamaño muestral", consistentes en juzgar igual probabilidad en la medida de un resultado, sin tener en cuenta el tamaño de la muestra del que resulta: por ejemplo, se plantea a los sujetos la pregunta: En una ciudad hay dos hospitales, uno más grande que el otro; de manera que el número de nacimientos es mayor en el más grande; imagínese que se registran los días en los que el número de nacimientos correspondientes a un sexo frente al otro es mayor del 60%. ¿Cuál de los hospitales registrará un número mayor de entradas de esa clase?;

efectivamente, la variabilidad será mayor en el hospital pequeño, dado que una diferencia más grande entre los datos en más probable en una muestra pequeña; no obstante, esta no fue la respuesta de la mayoría de los sujetos a estudio, que se dejaron guiar por el tamaño de los hospitales como información más fácilmente "reprsentable" y dieron mayoritariamente la respuesta errónea.

Bajo la denominación "Ley de los pequeños números" los autores detectan otra variante de las falacias provocada por los juicios basados en representatividad: es decir, que se juzgue más probable una secuencia de resultados de la forma (con C= cara y X= cruz): CXCCXC que CCCXXX o CXXXXX, deduciéndolo de que la representatividad de los dos resultados C y X, es la misma. Este error es la base de la falacia del jugador, consistente en otorgar mayores probabilidades , digamos, a C, tras una racha de repeticiones de X: A XXXXXXX debe seguirle C con una probabilidad cada vez mayor.

Provocada por esta clase de heurística descritas se producen diversas consecuencias, entre ellas, la ilusión de validez. Los sujetos muestran una gran confianza en sus juicios basados en representatividad, sobre todo en los personales: el estereotipo anteriormente descrito sobre el bibliotecario es tremendamente fuerte, y capaz de resistir grandes dosis de información objetiva que se suministren al sujeto. Por otra parte, toda información que confirme el primer juicio, aunque sea redundante o esté correlatada con la misma (con lo cual, no es verdadera información) será tomada por el sujeto como fuentes de apoyo de su primera intuición. Muy al contrario de lo que nos dice la estadística, en el sentido de que fuentes de información independientes o variables no correlatadas con las primeras que apoyen una afirmación aseguran el grado de confiabilidad de la hipótesis primera. En dependencia de ello tiene lugar lo que suele llamarse la "falacia de la regresión": por ejemplo, cuando se consideran los resultados académicos de diversos estudiantes sometidos a una serie de exámenes: si se selecciona un grupo de alumnos por su buen resultado en el primer examen y se comparan sus resultados con los del segundo, la comparación suele ser decepcionante: en realidad, si se toma el grupo de peores calificaciones en el primer examen, generalmente se observará una mejoría (de media) en el segundo. Ello se debe a que los resultados extremos tiene que ver en mayor grado con la desviación de la variable "resultados de la prueba"que con el verdadero "nivel académico medio" de cada alumno. No obstante, esta relación no llega a ser percibida por los sujetos, ya que una gran parte de contextos no se prestan a ser analizados de este modo (intuitivamnete, se entiende); y cuando se percibe el efecto de la regresión, (que no la causa estadística del mismo), suele explicarse en base a causas es púreas de carácter especulativo. Esta atribución a causas irrelevantes puede tener efectos muy perniciosos: por ejemplo, en el campo de la educación, siguiendo el caso anterior; el profesor puede escoger dentro del conjunto de incidencia que acompañan a la realización de los exámenes alguna que sea totalmente causal, y no obstante, tomarse como correlatada: por ejemplo, practicar una clase de exámenes de una clase o de otra: premiar al alumno o castigarle, etc. alguna de estas variables puede ser tomada como la causa y crear uno de los juicios erróneos que puede echar a perder la docencia de ese curso.

· <u>Disponibilidad:</u> Cuando se evalúa la frecuencia determinada a la ocurrencia de una propiedad en un objeto o suceso (pertenencia del mismo a una clase)o se adjudica un valor de probabilidad consecuente a esa frecuencia estimada, la mayor parte de los sujetos atribuyen una frecuencia o probabilidad mayor a un suceso que puede se más fácilmente “convocado” mentalmente: a diferencia de la heurística considerada anteriormente, la ahora considrada depende de la facilidad de cálculo o de ocurrencia de un objeto cuantitativo, no de la representación imaginativa de una clase de objetos o sucesos. Esta disponibilidad de ejemplos convocables cuantitativamente pueden tener que ver con varias características: la familiaridad del sujeto con los objetos en cuestión es una de ellas. Por ejemplo, al ser presentada una lista de personalidades públicas conocidas a un sujeto, de tal manera que las más conocidas para él pertenezcan a uno de los exos, al ser preguntado por la frecuencia en la que se le han mostrado el sujeto decidirá sobre la base de los individuos que mejor recurda e indicará como más frecuente el sexo correspondiente, aunque en realidad la muestra esté igualmente repartida a este respecto.

Esta peculiar heurística tiene gran incidencia en la percepción del riesgo; el conocimiento de que varios parientes o amigos han sufrido un ataque al corazón predispone a suponer una probabilidad mayor a tal clase de eventos que la que se adjuducaría en caso de no conocer tales ejemplos. Asimismo, la percepción de la probabilidad de sufrir un accidente de tráfico aumenta acorde al número de accidentes contemplados efectivamente por el sujeto. Pero donde se ve más claramente el sesgo provocado por esta estrategia intuitiva es en aquellas situaciones en las que alguna clase de cálculo o ejercicio de imaginación cuantitativa tienen lugar; así, puede pedirse al sujeto que evalúe la frecuencia en la que la letra r sea la inicial de una palabra, frente a la frecuencia de que sea la tercera letra. En Inglés, la frecuencia es mayor en el último caso, pero al ser más sencillo traer a la memoria las palabras según su inicial que según su tercera letra, la mayoría de los sujetos señalaban la frecuencia equivocada. Asimismo una par de características pueden ir frecuentemente correlatadas según el juicio de muchos sujetos, simplemente porque saltan a la vista más fácilmente que otras.

· Ajuste y anclaje: A menudo se estima una probabilidad mediante una primera aproximación inducida, la cual realiza en la mente una suerte de “anclaje” que hace que el sujeto, ante una mayor información, no sea capaz de corregir del todo su primera asunción, la cual arrastra la respuesta final a valores cercanos a la misma. Digamos que se pregunta a los sujetos cuál es, por ejemplo, la proporción racial de los estados unidos, o el porcentaje de población mundial asiática. En primer lugar se les dice una cifra aleatoria entre 0 y 100, y se pregunta si el valor verdadero es mayor o menor que el ofrecido. Tras la respuesta, se les pide una estimación más exacta. Pues bien, una gran cantidad de sujetos respondían con cifras que se aproximaban más a la sugerida, no importaba cuál fuese el valor verdadero de lo preguntado; una sugerencia de una cifra baja provocaba respuestas generalmente bajas, y viceversa.

Otro experimento preguntaba a un grupo de sujetos por el resultado, mediante cálculo mental en un breve espacio de tiempo, de la siguiente multiplicación:

1 x 2 x 3 x 4 x 5 x 6 x 7 x 8

mientras que a otro grupo se preguntaba por una estimación rápida de esta otra:

8 x 7 x 6 x 5 x 4 x 3 x 2 x 1

Los del primer grupo hicieron una estimación sensiblemente inferior a la emitida por los del segundo.

> En el campo de la intuición probabilística se dan abundantes ejemplos de esta clase de estrategia mental equivocada: por ejemplo, en la consideración de la probabilidad en procesos que los autores dividen en independientes, conjuntivos y disyuntivos. El experimento consiste en comparar composiciones de sucesos que presenten probabilidades similares y comprobar cuánto se diferencia de la percepción intuitiva en cada caso. El experimento está citado dpor los autores de Bar-Hillel. Dénse tres distintas combinaciones de sucesos: 1) Sucesos independientes: extraer ua bola de determinado color de una bolsa con la mitad de las bolas de ese mismo color. 2) Combinación conjuntiva de sucesos: extraer (con reemplazao) siete veces consecutivamente una bola de un color determinado de una bolsa con bolas repartidas al 90 % con ese color; 3) Combinación de sucesos disyuntivos: extraer *al menos una vez* en siete intentos una bola de determinado color de una bolsa cuya proporción de bolas de tal color fuera del 10%. Pues bien, en el primer caso, la probabilidad es del 50%; en el segundo, es de $(0,9)^7 \cdot 100 \approx 48\%$; en el tercero, $[1- (0,9)^7] \cdot 100 \approx 52\%$. Pues bien, ofrecidos los tres procedimientos a la apuesta por parte de los sujetos, la mayor parte escogió la combinación conjuntiva; la disyuntiva, cuya probabilidad es menor, fue escogida en último lugar. Este fallo en la estrategia de decisión también tiene repercusiones en planeamiento de proyectos o el diseño de procesos industriales. Generalmente el reultado buscado es producto de una combinación de muchos sucesos, todos de una probabilidad alta; pero la factibilidad del proyecto (o la fiabilidad del proceso) eran sistemáticamente sobrevaloradas debido a su conjuntividad.

El trabajo que hemos resumido continúa con un análisis más pormenorizado de diferentes tópicos relacionados con la primera clasificación junto a la descripción de experimentos para valorar su incidencia y escala. Aquí basta con haber referido lo suficiente como para apoyar nuestro enunciado, acerca de la particular dificultad, desde el plano cognitivo, que los tópicos diversos de probabilidad y estadística presentan al público, y por ello al estudiante de secundaria. Una dificultad que no depende del grado de atención, ni de la capacidad para manejar lógicamente entidades abstractas, es decir, para realizar operaciones formales, ni de su visión espacial, como la de otras ramas de la matemática incluidas en el currículo de secundaria, sino de dificultades intrínsecas a la materia de estudio, que no es equívoca en su definición (aunque históricamente, como hemos visto, este punto ha tardado mucho tiempo en resolverse), pero que, aunque aprentemente los problemas que presenta son de fácil representación mental, así como de relativamente sencilla exposición lingüística, y aunque las nociones estudiadas tienen una contraprestación directa en fenómenos fácilmente observables por el alumno —como la extraccción de bolas de una urna o la tirada consecutiva de dados— , esta facilidad, como decimos, es superficial; ya sea innata, ya aprendida, nuestra percepción de las posibilidades es torpe, mucho más que nuestra intuición numérica o espacial.

¿Cómo afecta la cuestión a la enseñanza de la materia? En el sentido siguiente: es preciso realizar un esfuerzo, adicional respecto de otra ramas matemáticas, de fijación de heurísticas adecuadas. Para ello, trabajos como el estudiado tratan de detectar en la raíz y de encontrar los puntos comunes a los diversos errores encontrados, para, mediante el análisis del funcionamiento de las intuiciones y estrategias que llevaron al error, ser capaces de corregirlas. De lo que se deduce que, en la enseñanza de la probabilidad y estadística, el énfasis en los procesos de cálculo y el enfoque dirigido a la resolución de problemas debería dar paso (o considerar como paso previo) *la detección de falacias y malentendidos probabilísticos*, y la comprensión profunda de las implicaciones de los mismos. De una manera informal, aunque en nuestra opinión, exacta, y desde un punto de vista cercano a la enseñanza elemental práctica, Kai Lai Chung **[9]** hace el siguiente comentario:

> *Muchos principiantes encuentran difíciles los problemas de probabilidad y combinatoria porque han sido habituados en otros cursos elementales de matemáticas a problemas de tipo "receta", tales como: "resolver la ecuación x^2-5x+10=0" o "derivar (quizá dos veces) la función xe^{-x}", etc. Tales problemas pueden resolverse sin más que aprender de memoria ciertas reglas; apenas es necesario pensar por cuenta propia. Evidentemente, también hay en combinatoria problemas de este tipo* [...] *hay, por ejemplo, una famosa fórmula para resolver el problema "de la mesa redonda"*[2] *: " ¿De cuántas maneras distintas se pueden sentar a una mesa redonda 8 personas?" Conociéndola se puede resolver el problema sin saber qué significa aquí la palabra "distintas", lo cual es el meollo del problema. Ahora bien, basta una pequeña variación para ponernos en dificultades. La verdad es, y es una verdad de Perogrullo, que la comprensión profunda del problema no puede sustituirse por nada."*

Por nuestra parte constatamos la realidad de la situación reflejada en el párrafo anterior: e incidimos en que la dificultad de este y otros problemas relacionados con la incertidumbre dependen de lo inapropiado de nuestras intuiciones primeras. Una intuiciones que, por cierto, persistirán en el alumno que no se especialice en esta rama de las matemáticas si no se practica un ejercicio sistemático de desvelarlas y corregirlas en las aulas de enseñanza obligatoria hasta la adultez, influyendo de manera perjudicial en su capacidad para el juicio y la acción ante situaciones de incertidumbre: esto es, la inmensa mayor parte de las decisiones que afectan a la vida de un sujeto.

2 El que escribe estas líneas ha visto (y deplorado) la tal ecuación en un libro de texto de secundaria, fabricada *ad hoc* para uso exclusivo en problemas de mesas redondas, y a la que presume una gran número de potenciales aplicaciones en negociaciones de convenios colectivos, tertulias radiofónicas, organización de bautizos, bodas y comuniones, etc.

Datos y más datos

Hemos visto ya una muy somera —y parcial— aproximación histórica al estado actual de las disciplinas de probabilidad y estadística, comunes denominaciones del bloque del currículo considerado. En nuestras primeras páginas no realizamos esta clásica distinción, sino que preferimos inopinadamente términos como “azar”, “fenómeno aleatorio”, “situaciones de incertidumbre”, etc. No era sólo por afán de generalidad, sino porque el desarrollo moderno de las anteiores disciplinas ha trastocado varias nociones clásicas y ha formado nuevos agrupamientos. Los responsables del informe PISA, en su última edición completa (2012)(*vid.* [**38**], p. 19) consideran problemas en cuatro bloques temáticos de los cuales el último se denomina “Incertidumbre y Datos”. Una nota nos informa, además, de que la inclusión de la palabra “datos” es una novedad de la tal edición. De la incertidumbre y sus distintos abordajes respecto de la teoría de la probabilidad —y su rama correspondiente de teoría de la decisión— y la estadística matemática e inferencial, con respecto a la enseñanza en educación secundaria, se ha tratado desde una pespectiva histórica en páginas previas. Abordaremos ahora este otro aspecto en relación a la enseñanaza, y a la vez dependiente en gran parte de la historia reciente y de la situación tecnológica, económica y social de la inmediata actualidad: los datos.

Ya se han expuesto los orígenes y desarrollo diversos de la probabilidad y la estadística matemática e inferencial, y las influencias mutuas que ambas sufrieron en los momentos en que se han producido intersecciones más o menos profundas y constantes entre los dos modos de estudiar lo casual. Recapitulando: la teoría de la probabilidad, nacida como teoría del juicio y la decisión racional declinó, en buena parte por indefinición de su objeto y por las dificultades epistemológicas y filosóficas que suscitaba. La suposición de equiprobabilidad Laplaciana vegetó, en una especie de limbo desde el cual ni resultaba fértil teóricamente, ni encontraba aplicaciones en otras disciplinas que motivasen su desarrollo, salvo en problemas derivados de juegos de azar. Las probabilidad condicionada y el teorema de Bayes reformulado por Laplace esperaban su turno, que llegaría en el siglo XX, para que filosofía, lógica y, también parte de la estadística la interpretaran de modo que se relacionase con la inducción lógica, el criterio de validez de las hipótesis científicas y la perspectiva bayesiana de la inferencia[1], provocando vivos debates que aún persisten y una escuela alternativa de estadísticos frente a la ortodoxia frecuentista, aunque de cada vez mayor aceptación. Tan sólo el enfoque que unía probabilidades con frecuencias realmente observadas de sucesos derivó en un campo de estudio vivaz, la estadística inferencial, y solo en relación con la estadística empírica (o positiva), aunque por medio de algunas partes estrictamente matemáticas de la obra de Bernoulli, los resultados sobre muestras de De Moivre, los estudios sobre error de mediciones basados en Legendre y popularizados (o quizá inventados) por Gauss y la formulación matemática de su distribución, mediante técnicas en un primer momento heurís-

1 Una de las razones de que indiquemos que se peca de parcialidad en la introducción histórica es la desatención consciente de estos conceptos, debida a su ausencia de los currículos de matemáticas d enseñanzas medias.

ticas y analíticas, (como el método de los momentos y el de mínimos cuadrados), como principales aportaciones se convirtió, en manos de Galton y Pearson principalmente, en estadística *matemática*.

Nos detuvimos en el momento en que R. A. Fisher extendía el campo de la estadística desde la inferencia (paramétrica) respecto de un conjunto de datos observacionales al diseño estadístico de experimentos, único viable para fenómenos que presentan gran variabilidad y efectos de reducida magnitud (y complicada medición). No obstante, tras la segunda guerra mundial tuvieron lugar profundos cambios sociales y económicos que hicieron pasar el análisis de datos a un primer plano. Primeramente, la eclosión y exponencial aumento de la influencia de los medios de comunicación de masa, paralela a la nueva sociedad de consumo que se comenzaría a fraguar tras la guerra mundial. Uno de sus efectos en relación a nuestro asunto fue la democratización de la estadística. Si bien Mark Twain, a comienzos de siglo ya peroraba contra su abundancia y sus perjuicios, hay que entender que ni él ni aquellos para los que escribía eran todavía las masas, sino la clase media, entendida al modo dieciochesco; muy distinta de la *middle class* que se considera actualmente. La obra sintomática de Huff **[28]** muestra hasta qué punto su extensión era un hecho cotidiano ya en los años cincuenta. De lo que no da cuenta Huff es que las mismas masas estaban siendo consideradas como datos, no solo como sus consumidores, pero desde un punto de vista más amplio que el de los primeros estadísticos, que era el punto de vista del estado, político. En una sociedad de consumo, necesariamente las técnicas de mercado requieren de datos, y de intrumentos de análisis de los mismos, en orden a orientar la producción y la venta.

Pero este es solo uno de los aspectos del uso, ya ni matemático ni científico, sino meramente *industrial* de la estadística. En la lógica de la producción y la venta, la investigación de modelos es secundaria: lo que se requiere son instrumentos de juicio y acción, de toma de decisiones ¿Qué fabricar, cuántas unidades, a quién vendérselo, de qué manera maximizaré el número de compradores? estos son los interrogantes para los cuales los modelos clásicos se mostraban inoperantes o farragosos. A la lógica de la industria no le interesa tanto averiguar el modelo que sigue la distribución de una o varias variables, ni mucho menos la inferencia causal al modo de la ciencia, sino la *realización de predicciones*, entre otros aspectos, en apoyo de la toma de decisiones, en muchos casos costosas y/o irreversibles; en fución de ello, la economía financiera fue pronto ávido reclamo de tales instrumentos.

Otros campos con mayor enjundia teórica provocaron en algunos estadísticos el cambio de enfoque del análisis y estudio de modelos hacia el análisis, y muy pronto la *exploración* de datos. Por ejemplo, en el campo de la bioestadística: en el estudio estadístico clásico las tablas de las que se partía tenían unos listados ingentes de muestras en referencia a una pocas variables: las especiales condiciones de las ciencias biológico-sanitarias —incluyendo medicina, farmacología, genética, etc.— provocaban la situación contraria: un número limitado de individuos y una cantidad ingente de variables consideradas por cada individuo. En esta clase de análisis se desarrollaron técnicas matemáticas de análisis cierto, pero que no hubieran podido resultar útiles sin la participación de la otra característica fundamental de esta fase histórica conducente al estado actual de cosas, y que es: la aparición y expansión de la informática y sus consecuencias para la computación y el tratamiento de datos. No obstante, la temática de análisis multivariante queda tan lejos de los contenidos de enseñanza secundaria que no la trataremos aquí. Aunque la emergencia de las nuevas tecnologías informáticas, y más tarde las nuevas técnicas de telecomunicaciones, vienen al caso al respecto del estudio del campo de la *exploración de datos*, que repasaremos a continuación[**45**]:

> *Este libro es sobre análisis exploratorio de datos, sobre mirar los datos para ver qué es lo que quieren decir. Se concentra sobre todo en aritmética sencilla y gráficos fáciles de dibujar. Considera cualquier apariencia que hallamos reconocido como descripciones parciales, y trata de mirar debajo en busca de nueva inspiración. Le concierne la apriencia, no la confirmación.*[trad. propia]

He aquí el casi primer párrafo del prefacio de la obra fundacional de Tukey, *Exploratory Data Analysis (1977)*, y que, como si dijéramos, inventó todos los instrumentos gráficos modernos de estadística descriptiva orientada al análisis, más que a la representación de los datos: los diagramas de tallos y hojas, de caja y bigotes, la utilización preeminente de los estadísticos robustos como la mediana los cuartiles y los percentiles (la descripción mediante 5 valores de la distribución de los datos). El tono de intención didáctica es más que evidente. Un poco más adelante se puede leer:

> *El análisis exploratorio de datos es un trabajo de detectives; trabajo de detective numérico; o trabajo de detective cuantitativo; o trabajo de detective gráfico.*
> *Un detective investigando un crimen necesita herramientas y conocimiento. Si no tiene polvo de detección de huellas dactilares, no podrá encontrar esas huellas. Si no conoce dónde es más probable que el criminal haya puesto los dedos, no sabrá buscar en los lugares adecuados. Es el propósito de este libro proveer algo de ambos.*

Estos párrafos de tintes tan *naïf* indicarían, en primer lugar, una intención clara de introducir sus contenidos en capas más amplias del público que el estadístico profesional o el *scholar* (universitario), hasta entonces el destinatario de esta clase de obras. Y, en segundo lugar una intención preeminentemente práctica: los datos dicen cosas, más allá, o, mejor dicho, dicen cosas *antes* de su transformación en objetos matemáticos. En un tono igualmente desenfadado, Persi Diaconis [**15**] escribe:

> *El análisis exploratorio de datos buscan revelar la estructura, o hallar descripciones sencillas, en los datos. Cualquier búsqueda de estructura es un ritual azaroso que puede aproximarse* ***al pensamiento mágico****, a menos que permitamos la posibilidad de que la estructura revelada pueda haber surgido por azar. Incluso así, la investigación exploratoria,* ***cuando está controlada adecuadamente y equilibrada con propiedad por más métodos de estadística clásica****, ha sido un importante contribuyente al progreso científico.* […]
> *La estadística matemática clásica* […] *busca interpretar los patrones como fluctuaciones aleatorias. En su formulación más rígida, decidimos sobre los modelos e hipótesis* ***antes de ver los datos.*** *Después calculamos estimaciones y llevamos a cabo nuestros tests sobre nuestros presupuestos. La estadística clásica nos ofrece un antídoto para algunos de los problemas del pensamiento mágico. No obstante, como descripción de lo que un científico real hace cuando se enfrenta a una real, rica base de datos, parece tan lejana al sunto como un ritual primitivo.*[...]
> *En cierto sentido la estadística clásica facilita el pensamiento mágico: algunas veces la "persona con los datos" mira al "estadístico" como alguien que le ofrecerá todas las respuestas y despejará todas las incertidumbres, exactamente como el primitivo hombre de la tribu mira al chamán local como a quien practica los rituales adecuados y hace que se retire la enfermedad. Tukey (1969) discute el uso de la estadística clásica como una suerte de ritual de santificación.*

Se echa de ver en el anterior texto, si se coteja con la evolución de los currículos y de los textos de secundaria, la influencia que opiniones como éstas han tenido sobre la confección de los currículos de enseñanzas medias, no solo en la presencia esencial de la estadística descriptiva, sino en lo referente al orden de presentación de los contenidos; así, en primer lugar se realiza la enseñanza de la estadística descriptiva, previamente a la teoría de la probabilidad y la estadística matemática e inferencial.

Otro aspecto importante a destacar es su caracterización, un tanto beligerante, de los métodos estadísticos clásicos, entendidos en su versión más formal, como un proceso opaco, en cierta forma ajeno a la práctica real de la experimentación: una suerte de *caja negra* donde se introducen los datos y se recogen los resultados, pero en el cual el mecanismo conceptual —la maquinaria en el interior de la caja— es desconocido, y casi místico, para el no especialista. Sea o no exagerada, veremos como estas posturas han determinado, no solo el papel social y científico del estadístico en la actualidad, sino las reflexiones sobre la enseñanza de la disciplina, con un nuevo enfoque no exclusivamente matemático.

¿Se huabrían impuesto la utilización de estos criterios, el planteamiento y las herramientas descriptivas sin el positivo apoyo de los ordenadores, que pueden realizar operaciones con ingentes cantidades de datos en meros instantes, además de ofrecer la posibilidad de construir representaciones gráficas basadas en cualquier grupo de premisas de investigación con la sola variación de unos pocos parámetros de algún *software* especializado? Una pregunta tan extensa solo puede ser retórica. En realidad sucedió una retroalimentación mutua: fueron las primeras computadoras las que posibilitaron los primeros análisis exploratorios, y aquellos impusieron la introducción de las nuevas tecnologías en la labor del estadístico, y en la enseñanza de la estadística.

Finalmente todas estas influencias —la formación de una figura profesional del estadístico como profesional liberal de una disciplina independiente, las necesidades prácticas de la industria, las finanzas y los estados de análisis de bases de datos cada vez más abultadas, la aparición de las nuevas tecnologías, la generalización de los métodos gráficos y la utilización de estadísticos robustos— desembocaron en las nuevas perspectivas de educación estadística; a este respecto, es central la figura de David S. Moore, cuyos artículos (uno de ellos, en colaboración con Cobb) **[10]** **[36]** son rotundos al respecto, y no solo eso; también centrales: incluso los marcos teóricos PISA 2012 incluyen, en apoyo a su inclusión la voz "datos", junto a la "incertidumbre", (que catalogan uno de los cuatro grupos temáticos de problemas demostrativos de la adquisición de competencias básicas en educación obligatoria) los anteriores artículos, donde se pueden leer declaraciones como las siguientes **[10]**:

> *¿Cómo difiere el pensamiento estadístico del pensamiento matemático? ¿Cuál es el papel de las matemáticas en la estadística? Si se purga a la estadística de su contenido matemático, ¿Qué sustancia intelectual queda?* [...]
> [...] *el foco último del pensamiento matemático son los patrones abstractos: El contexto es parte de los detalles irrelevantes que deben ser hervidos, separados mediante destilación en el fuego de la abstracción en orden a revelar el previamente escondido cristal de pura estructura.* ***En Matemáticas, el contexto oscurece la estructura****. Como en matemáticas, el análisis de datos también busca patrones, pero en últmo extremo, en análisis de datos, si los patrones tienen significado y si tienen valor, depende de cómo estén entrelazados con los hilos complementarios de la* ***historia, el relato*** *al que pertenecen a los datos.* ***En análisis de datos, el contexto provee de significado****.* [trad. propia]

Esta postura militante acerca del papel de las matemáticas para el estadístico resulta, en parte, reflejo de la separación definitiva entre teoría de la probabilidad y estadística, habiendo protagonizado la primera una expansión sin precedentes desde Kolmogorov y gracias a la física cuántica, y otras ramas de las matemáticas como la teoría de juegos, teoría de la computación y de la comunicación, etc. cuya aparición y fomento fueron propiciadas, asimismo, por el avance de la informática, las nuevas técnicas de predicción en economía y estudio de mercados, etc. Otros autores afines también constatan esta separación en perspectivas, obejetivos y metodología; por ejemplo, leemos en **[20],** p. 110 y ss.:

> *Los métodos que el estadístico usa son matemáticos y abstractos. Pero no proceden de una sola idea, sino de una red de ideas interconectadas. La mayoría, pero no todos, usa el concepto de probabilidad. Muchos, pero no la mayoría, usan el concepto de modelo estadísitico. Sin embargo, estos dos conceptos,* ***probabilidad y modelo, son centrales en la red de conocimientos del estadístico****, desde el mismo momento en que los métodos que no los usan, como el análisis exploratorio de datos y los test estadísticos no paramétricos, a menudo son definidos precisamente por el hecho de que no los usan.*[...]
> *El concepto de modelo estadístico fue introducido por Fisher, pero no lo definió formalmente; no estaba interesado en la matemática abstracta, sino que usó un acercamiento conceptual más bien intuitivo.*[...] *Fue Harald Cramér quien, entre 1930 y 1950, meditando sobre la ciencia estadística tal y como se concebía enlos países anglosajones –basada en la tradición más empírica y experimental– y el rigor matemático de la teoría de la probabilidad propia de la Europa Continental. Intentó unificar estas dos tradiciones en su libro de 1946* Métodos Matemáticos de Estadística **[12].** ***Esta unificación nunca fue completa****, y uno todavía encuentra la teoría de la probabilidad trabajando en los departamntos de matemáticas, mientras que los estadísticos, que aplican sus resultados en inferencia, tiene sus propias unidades. Característico de ello es la escisión, en 1973, de la publicación* Annals in Mathematical Statistics (Anales de Estadística Matemática) en Annals of Statistics (Anales de Estadística) y Annals of Probability Theory (Anales de la Teoría de la Probabilidad) [trad. propia]

Sin duda los autores reflejan una situación más bien perteneciente al mundo anglosajón, donde el estadístico tiene un campo de actividad más abierto (debido a la potencia y necesidades de la gran industria allí localizada), no estrictamente coincidente con el académico, como ha sido el caso del continente europeo hasta hace relativamente poco. No obstante el hecho tiene gran influencia en la estadística en general, dada la hegemonía desde el punto de vista de la investigación y la publicación científica de los Estados Unidos, y en su didáctica en particular, incluyendo los estudios de organizaciones internacionales como la OCDE, responsable de las pruebas PISA, o incluso las directivas europeas sobre los currículos de enseñanza obligatoria, que afectan a las leyes de educación del entorno europeo (como la LOMCE). Moore también da cuenta de esta especie de "cisma" entre probabilidad y estadística en **[10]**, p. 803:

> [...]*pero aunque la estadística no puede prosperar sin matemáticas, la recíproca no es cierta. Que la estadísitca no es una parte necesaria de la formación de un matemático está impícito en la declaración por parte del eminente probabilista David Aldous en la que dice que él "está interesado en las aplicaciones de la probabilidad a todos los campos* ***excepto la estadística****"*[trad. propia]

El reflejo de esta situación en la didáctica estadística es claro, si se observa el creciente papel en los currículos de las leyes sucesivas, donde la estadísitica descriptiva e inferencial gana cada vez más contenido, mientras que la teoría de la probabilidad se va aligerando, por ejemplo, en la parte de combinatoria (bastante dificultosa para el alumno de secundaria). En particular, lo que Moore y la escuela de la nueva didáctica en estadística, asumida de hecho por los más recientes planes de estudio para enseñanza secundaria, internacionales y españoles **[3]**, y de derecho por los responsables de los marcos teóricos de la prueba PISA **[38]** —como se desprende de su lectura sobre nuestro asunto particular— se deduce de las siguientes directrices expuestas en **[36]**, u otras de carácter análogo:

> **Más datos y conceptos:** *Los primeros cursos deberían ofrecer experiencia trabajando con planteamientos de problemas reales. Deberían concentrarse en aquellas cosas que no pueden (al menos todavía* [1997]*) automatizarse, tales como interpretación de gráficos, estrategias efectivas de exploración de datos, diagnósticos básicos como preliminares de la inferencia, y el significado conceptual del "p-valor", "confianza", y "significancia estadística".* […] *acometer un razonamiento de inferencia es más importante que cuántos procedimientos individuales toquemos*[...]
> **Enfatizar los conceptos estadísticos:** "La estadística tiene su propia substancia, sus propios conceptos distintivos y modos de razonar. Estos deberían ser el corazón de la enseñanza de estadística a principiantes a todos los niveles de sofisticación matemática". […] *A un estudiante que sale de su primer curso de estadística sin apreciar la diferencia entre observación y experimentación, y la importancia de los experimentos aleatorios comparativos, por ejemplo, se le ha engañado.*
> **Automatizar el cálculo y la representación gráfica:** *Quizá después de un ejemplo corto hecho a mano por motivos pedagógicos, las medias y las desviaciones estándar son botones en una calculadora y los diagramas de puntos con líneas de regresión son ítems en el menú de un software. La automatización del cálculo es un asunto controvertido entre los matemáticos; pero es mucho menos controvertido entrelos estadísticos, puesto que refleja la práctica real de nustra disciplina.*[trad. propia]

Todo lo cual nos conduce, finalmente, a la justificación del presente trabajo.

Conclusiones y Justificación del trabajo

En páginas precedentes adelantamos las que habrían de ser las conclusiones en las que nos hemos apoyado para acometer la redacción de este trabajo, de un modo quizá no muy coherente desde el punto de vista lógico formal, —ya que primero habría que haber fundamentado bien esas bases— pero con la intención de dotar nuestro discurso de un marco previo útil sobre el cual ubicar los diferentes enunciados expuestos. Volvemos a exponerlas a continuación, para después pasar a demostrar cómo se siguen de los puntos considerados en los apartados previos de esta introducción:

1ª) Las áreas del currículo de matemáticas de enseñanza secundaria correspondientes a probabilidad y estadística pueden desplazar su centro de gravedad desde sus aspectos más teóricos, y sus aplicaciones algorítmicas más básicas, hacia las técnicas de representación, análisis y valoración de datos o de las organizaciones y representaciones de los mismos, en orden a formar juicios, tomar decisiones y ejecutar acciones de interés efectivo para el educando, sobre todo en la fase obligatoria (ESO) de aquellas enseñanzas. No tanto la resolución de problemas numéricos o simbólicos como la atención, la capacidad de formular hipótesis, emitir juicios y, en su caso, definir estrategias de acción conforman valores cívicos en la educación del alumno y posterior desarrollo personal como adulto.

2ª) Es el **contexto** el que da significado a los datos, y no sólo eso: también el que despierta el interés y el aprecio por la práctica de su análisis. Este contexto puede ser más o menos cercano a la cotidianeidad del alumno: sus fuentes ser pueden directas, o procedentes de los medios de masas con los que tiene fuerte contacto, pero también pueden escogerse de temáticas de interés social, económico, científico y/o histórico. En todo caso los datos deben detentar alguna clase de relevancia en un acaecimiento factual, y no sólo ocupar un lugar inserto en un sistema abstracto, lógico-simbólico o matemático.

Respecto del punto 1º): un propósito explícito en los planes de estudio de enseñanza secundarioas *obligatoria* (ESO), y la misma existencia y forma de aplicación de las pruebas PISA, cuyo reconocimiento e influencia internacionales son indiscutibles, consiste en el fomento de las *competencias.* Abundantemente descritas e interdependientes, resulta un poco más difícil buscar el propósito *implícito* en las que se basa su definición, que tratan de responder a los siguientes interrogantes: ¿Cuál debe ser el nivel de formación matemática *mínimo* que debe poseer un ciudadano, sea cual sea la orientación que elija para su carrera profesional y su desarrollo personal? ¿Cuál debe ser el enfoque de esta formación mínima, esto es, qué posibilidades de efectiva realización en la cotidianeidad, en la práctica de labores no necesariamente científicas, en la toma de decisiones vitales, brinda la formación matemática considerada?

No en vano las pruebas PISA se realizan a la edad estimada en la que los países desarrollados suelen dar por concluida la fase de escolarización obligatoria: de nuevo, implícitamente este hecho nos dice que la preocupación principal de los responsables de tal informe es el total de la ciudadanía, no solo aquella que orientará su desarrollo intelectual y su labor profesional hacia las ciencias intrínsecamente matemáticas o para las cuales las matemáticas son una parte decisiva y fundamental. En los planes de estudio se refleja esta distinción claramente al dividirse la misma introducción, antecedentes y definiciones básicas en dos partes diferentes, una correspondiente a la ensñanza obligatoria y otra al bachillerato. El supuesto subyacente principal, es, por supuesto que para la ESO deben definirse las características que para una gran proporción de la población serán el último contacto con la educación matemática (al menos, formal); mientras que el bachillerato definirá unos contenidos cuyo propósito será preparar al alumno para estudios superiores.

La enseñanza en ESO debe tener, por tanto, un carácter *autocontenido en respuesta a unos objetivos mínimos predefinidos.* En cambio, la enseñanza de Bachillerato debería tener un crácter *propedéutico,* es decir, preparatorio para distinos grados de desarrollo previsibles: por lo cual debería ser más versátil y abierto, pero también más concentrado en las materias desde el punto de vista formal, puesto que el alumno se resentirá de la falta de formalidad no bien acuda a un centro de enseñanza superior. Que estas distinciones sean efectivamente logradas constituye en buena medida, el éxito o el fracaso de los planes de estudio respectivos.

En respuesta a estos requerimientos, el trabajo presente se enfoca en mayor grado a la perspectiva antes descrita sobre educación obligatoria, con la intención de formar competencias que sean al mismo tiempo básicas, transversales y útiles para el ciudadano no especialista en matemáticas o en estadística: en este sentido, no solo se ha acudido preferentemente a fuentes en estrecho contacto con el común de la ciudadanía (y, a veces, del mismo profesional estadístico), sino también en busca de asuntos que tengan verdadera relevancia para la vida del ciudadano: la incertidumbre es omnipresente, y los medios de abordarla que posee la ciudadanía son escasos y en buena parte, conducentes a error, como hemos tratado de ilustral en el capítulo previo acerca de dificultades cognitivas. Grupos de interés político y económico conocen estas deficiencias y tratan sistemáticamente de aprovecharlas; esperamos, con los ejercicios propuestos, que este punto sea puesto de relieve en las aulas de enseñanza obligatoria y, al mismo tiempo, brindar herramientas tanto de carácter matemático como de análisis exploratorio de los datos — en lo posible, uniendo los dos enfoques— para *la evaluación de situaciones y la formación de juicios sobre la información* (los datos), así como la toma de decisiones subsiguiente, más que para la realización de cálculo simbólico o numérico o la resolución de problemas derivados de definiciones exclusivamente matemáticas.

Respecto del punto 2º): Hemos comprobado ya que los avances modernos de la teoría de la probabilidad rebasan, con mucho, el nivel de la enseñanza secundaria. Luego, el peso correpondiente a los resultados que contribuyen a la construcción de la estadística matemática debe ser mayor, indiscutiblemente en ESO, de manera un tanto menos preponderante en bachillerato.. Es por ello que la temática escogida y el enfoque de los temas de probabilidad deben, sin por ello quedar desnaturalizados, orientarse a la disciplina de la estadística (matemática). Pero otro tanto cabe decir de esta última disciplina: la enseñanza de sus objetos y relaciones no puede ser el centro de la enseñanza de la estadística en su nivel más básico: son los datos, y el contexto de estos datos, lo que debe constituir el centro de la enseñanza estadísitica en ESO; en Bachillerato, dado su caráter preparatorio y en base a que se supone que la educación obligatoria ha configurado el esquema inicial de la disciplina, y en atención a los alumnos cuya formación posterior transcurra por el camino científico-técnico[1], pueden ofrecerse contenidos más formales. Pero solo en el supueesto de que estos son posteriores a la introducción de contenidos de análisis y descripción de datos cuyo valor depende del interés del contexto real del que proceden.

Debido a ello se han tomado, en las actividades constituyentes del presente trabajo, el punto de vista preferentemente *contextual* reseñado por Moore: más allá de que los ejemplos para las actividades estén tomados de los medios de comunicación, de informes científicos o técnicos, o de la misma historia de la matemática, el hilo conductor que los une es que responden a un ***relato*** que da carácter, significación y relevancia a los datos presentados y a los problemas propuestos. Es decir: no se han *fabricado los datos* para ilustrar tal o cual objeto o procedimiento de carácter probabilístico o estadístico, sino que se ha *acudido en busca de datos, de forma que se consigan los siguientes objetivos:* que porporcionen una materia interesante de reflexión, no solo matemática, al alumno; que sean *relevantes*, es decir, que respondan a problemas reales —y a veces serios— a los que la sociedad en su conjunto, y junto ella el alumno y posterior adulto, deben saber enfrentarse.

1 En este sentido, resulta chocante comprobar que en los nuevos currículos de LOMCE se incluye la axiomática de Kolmogorov en el bachillerato correspondiente de **Ciencias Sociales**. Por este y otros detalles que muestran cierta incoherencia en la eleccion de los contenidos, prefrimos seguir un punto de vista independiente de los mismos, y más atento a los principios y objetivos en los que dice pretender inspirarse, así como en la filosofía general expresa en los marcos teórico del la prueba PISA, que el mismo ministerio de educación publica; sobreentendemos que sus principios generales son, por tanto, asumidos como propios por la administración pública española.

INCERTIDUMBRE Y CONTEXTO

TÓPICOS Y ACTIVIDADES

Nota Preliminar

La serie de actividades que se proponen a continuación no responden a una ordenanción sistemática, exhaustiva, de los contenidos del currículo de enseñanza secundaria. En primer lugar, porque la organización del mismo es amplimante redundante, por motivos didácticos; en segundo, porque resulta a veces incoherente con los principos que dice observar; y en tercer lugar, porque en dependencia de los planes de estudio en sus realizaciones concretas —como su historia refleja desde 1990—, resultan extremadamente volátiles, no así sus principios, en los que hemos preferido fijarnos en términos muy generales. Otra de las razones básicas de su configuración es la heterogeneidad de las fuentes de las que se alimenta; si hubiéramos dado preferencia a criterios de ordenación y sistematicidad, habríamos debido desechar capítulos interesantes (a nuestro modo de ver) e incluir en cambio otros de carácter más rutinario o prosaico. Una ordenanción consecuente habría puesto de relieve en mayor grado los "huecos" en el tratamiento de contenidos, así que, antes de tomar la decisión de un Mendeleiev, que construyó su tabla periódica de los elementos químicos teniendo en cuenta los "huecos" correspondientes a elementos aún no descubiertos en la época de su confección—tal era su confianza en su teoría— hemos preferido seguir un hilo conductor más cercano al del ensayo, más, si se nos permite decirlo, literario, en el que concepciones más o menos subjetivas de la disposición de los elementos ayuda a su construcción del conjunto como un todo, en cierta manera, *narrativo*.

Y si se diera el caso de que mis explicaciones no resultaran convincentes, podemos aún aducir un último argumento: el trabajo presente tiene por tema los fenómenos aleatorios, el azar. ¿De qué manera mejor organizar su contenido que siguiendo un criterio igualmente azaroso?

Problemas de muestreo

Un rasgo importante de la educación matemática, o, mejor dicho, de la educación en general, consiste en fomentar el espíritu crítico. Sobre este particular, la enseñanza de la estadística puede desempeñar un papel fundamental, dada su omnipresencia en medios de comunicación, así como su uso abusivo o tendencioso, como ya señalamos. Ahora bien, la estadística consiste en el análisis de los datos; sucede que hay aspectos de los datos *que no son estrictamente matemáticos.* ¿Debemos renunciar a elaborar metodologías de enseñanza que contemplen estos aspectos? Según el currículum de ESO y de Bachillerato (CC. SS.) **[4]**, entre los criterios de evaluación, se encuentra, en formulaciones muy similares por curso, lo siguiente:

1. *Utilizar el vocabulario adecuado para la descripción de situaciones relacionadas con el azar y la estadística, analizando e interpretando informaciones que aparecen en los medios de comunicación.*

El caso es que los medios rebosan de informaciones cuyo enunciado es cuantitativo y que se pueden —y se deben— criticar desde un punto de vista cualitativo, dado que el supuestamente objetivo dato numérico central de la noticia se ha obtenido mediante un método susceptible de incluir en sus resultados diversos tipos de *factores de influencia*.

El método de la estadística inferencial consiste en inducir propiedades de un conjunto de objetos (población) a partir de porciones más pequeñas (muestras) de tal conjunto. Para que las propiedades de la muestra y de la población sean comparables es preciso que la muestra sea lo suficientemente grande[2], pero, sobre todo, que sea *verdaderamente representativa,* lo cual necesita de algún tipo de aleatoriedad en el mecanismo de selección; sólo así puede evitarse la presencia de sesgos. Veamos unos ejemplos:

Encuestas de participación voluntaria

Uno de los errores de muestreo más frecuentes e insidiosos: un periódico, un partido político, un colectivo defensor de tal o cual causa o tal o cual ley, etc. convoca una encuesta de opinión, por medio de llamadas telefónicas, correo electrónico o convencional, o, últimamente, convocando una encuesta en su página web, a la cual el encuestado accede *voluntariamente.* Nos remitiremos al siguiente comentario en **[35]** :

> *El problema principal de las encuestas de participación voluntaria* [Call-In Polls] *es que los participantes están completamente autoseleccionados: sólo esas personas que ven u oyen ese programa de radio o televisión, o leen ese periódico, puede estar incluidas. Por añadidura, la gente que se toman la molestia de participar son, frecuentemente, muy diferentes de los que no. Eso es porque los que*

2 El margen de error no aumenta para un grado de confianza dado si la muestra se hace mayor que determinada cantidad, por muy grande que sea la población, siempre que esta sea de un tamaño 100 veces superior a la muestra. *Vid.* **[37]**

> *participan están normalmente más interesados por el tema o tienen sentimientos más fuertes sobre él.*[...]
>
> *Otro gran problema es que estas encuestas están abiertas a la manipulación de cualquier individuo o grupo que tengan intereses en el asunto. Sin límite en el número de llamadas, la gente puede realizar múltiples participaciones, y ciertos grupos pueden organizar operacions más sofisticadas para inundar las líneas telefónicas con llamadas en apoyo de su punto de vista.*[...]
>
> *Aunque los que tienen alguna formación en la investigación por sondeos saben que estas encuestas de acceso voluntario no deben tomarse en serio, muchos miembros del público no hacen la distinción entre estas pseudoencuestas y las de verdad. De hecho, las pseudoencuestas pueden ser vistas incorrectamente como más creíbles que las encuestas reales porque a menudo ofrecen cifras de participación mucho más grandes.*[traducción propia]

Debería resultar obvio que de esta manera se introduce un factor de influencia que destruye cualquier viso de aleatoriedad, y por tanto cualquier parecido de la muestra con la población. No obstante, su uso sigue estando bastante extendido, como se verá en las actividades.

ACTIVIDADES

1º) En la página web http://elpais.com/encuestas/ [consultado Mayo 2015] *se muestran una serie de encuestas propuestas a los lectores por el periódico El País. Abre una cualquiera de ellas y adjunta una imagen del cuestionario y otra de los resultados.*

Encuestas: Debate del Estado de la Nación.Votación [recurso en línea] (24 Feb. 2015) [cons. Mayo 2015]El País (edición digital). Disponible en: http://elpais.com/encuestas/html/2015/02/24/1424799658.html

Encuestas: Debate del Estado de la Nación.Resultados[recurso en línea][consultado 20 Mayo 2015)El País (edición digital). Disponible en: http://elpais.com/encuestas/html/2015/02/24/1424799658.html

a) A primera vista, ¿Te parecen los resultados fiables?¿Por qué?
b) ¿Crees que las "28028 respuestas" son suficientes para constituir una muestra representativa de la población Española? ¿Es este criterio suficiente para saber si un muestreo es representativo?
c) Investiga la metodología de la encuesta: ¿Puede producir muestras verdaderamente representativas? ¿Por qué?

a) Con esta pregunta se tratará de averiguar qué conceptos básicos sobre muestreo tiene el alumno (aunque probablemente también se recibirá una respuesta partidista por su parte, dependiendo de la notica elegida)

b) Como este es uno de los argumentos que se sostienen para intentar dar validez a esta clase de encuestas, resulta una buena ocasión para introducir una primer método de estimación paralelo o previo a la teoría, a saber, que, para un nivel de confianza del 95%[3] , la verdad sobre la *población* (en este caso, cada proporción p de votos sobre el total) estará dentro de un intervalo de valores cuyo "radio" será, en proporción respecto del valor del parámetro, $r=1/\sqrt{n}$ *(vid.*[**37**]*, pp.42-48)* , siendo n el tamaño de la muestra, es decir:

3 la fórmula $r=1/\sqrt{n}$ es una simplificación del llamado **margen de error para una proporción**, y que suele expresarse como $\pm\sqrt{\frac{p(1-p)}{n}}\cdot z_{\alpha/2}$, donde $z_{\alpha/2}=1.96$ para $1\text{-}\alpha=0{,}95$ (95% de nivel de confianza) y p=1/2 → p(1-p) es máximo.

- Que si repitiésemos 100 veces esta misma encuesta con este mismo método de muestreo, entonces, 95 de cada 100 veces, $\Rightarrow$
- $\Rightarrow$ los valores de los porcentajes ***reales p de la población*** (parámetros), estarían dentro de un intervalo de "centro" el valor de los porcentajes **estimados** (estadísticos), en nuestro caso, $\hat{p}_{\text{rajoy}} = 22{,}38\% \hat{p}_{\text{sánchez}} = 69{,}24\%$ y de "radio" un valor igual a $r = 1/\sqrt{n}$, siendo n=28028 de donde r≈0,6 % $\Rightarrow$
- $\Rightarrow$ los valores reales de las proporciones buscadas tienen una probabilidad del 95% de estar tan cerca como un 0,6% por arriba y por abajo (±0,6%) del valor estimado.

Una exposición como la anterior debería llevar al alumno a pensar que la encuesta es fiable, después de todo, pero resulta que es ¡Todo lo contrario! Nada infunde más respetabilidad a un discurso (mediático, político, y, en ocasiones, también pretendidamente científico) que el ropaje de las matemáticas: cifras, porcentajes, índices, precisión hasta el cuarto decimal... el truco es tremendamente efectivo —no hay más que atender un poco a la prensa, escrita, radiada, televisiva, o en red)—, y por eso debemos sacar al alumno del error, en primer lugar, inculcando que la forma matemática no siempre encierra contenido matemático.

¿Es el tamaño de la muestra, por tanto, adecuado para medir lo que se propone? Lo es, ciertamente, pero ¿Es este un criterio suficiente? En absoluto. El tamaño de la muestra influye en la anchura del intervalo alrededor del estadístico, no en la calidad del estadístico mismo; si éste padece de un sesgo, el hecho dependerá del método de muestreo, es decir, dependerá de si este método es capaz de producir muestras verdaderamente representativas[4].

c) La encuesta tiene un *sesgo sistemático,* debido a las razones expuestas más arriba **[35]**.

2º) Fíjate en la nota al pie de los resultados de la noticia que hayas escogido y adjunta la imagen: analiza el texto (por lo demás, bastante vago) ¿Podrías explicar en mayor detalle lo que quiere decir?

Esta encuesta no es científica, responde tan sólo a las respuestas de los lectores que desean exponer su opinión.

Para reforzar ideas, convendría incidir en que la clave del párrafo son las palabras "lectores" y "desean"; esto es, que la muestra está autoseleccionada por lo menos de dos maneras:

· Primera, porque la muestra está escogida entre los lectores del periódico en cuestión; luego no es una muestra aleatoria de la población que nos interesa (y que se supone que son todos los españoles en edad electoral).

· Segunda, que de entre la primera selección, sólo aquellos que han *deseado* responder lo han hecho: es decir, es otra selección no aleatoria dentro de la primera.

4 El estudio de técnicas de muestreo alternativas al aleatorio simple (*vid. infra*) no pertenece al curriculo de enseñanza secundaria.

Encuestas basadas en muestreos de conveniencia

En el anterior caso explorábamos las características no aleatorias de las muestras conseguidas mediante acceso o participación voluntarios; pero no son éstos, ni mucho menos, los únicos ejemplos de selección sesgada. Si las anteriores constitutían el ejemplo más claro de *autoselección*, en lo que sigue veremos que, a menudo, la muestra está *seleccionada* de principio por el encuestador.

Tomar muestras aleatorias simples para estudios basados en encuestas es caro, difícil o decididamente impracticable: por ello, se utilizan determinados métodos de muestreo más sofisticados [**40**]: por otra parte hay métodos, como el de estratificación, que consiguen muestras más representativas que la simplemente aleatoria. No obstante, ya sea por comodidad, por economía, por negligencia o, sencillamente, por ocultar los intereses de la parte que finacia o publica determinados resultados, a menudo esta selección no obedece a un criterio probabilísticamente válido (en el cual, cualquier miembro de la población tiene las mismas probabilidades de pertenecer a la muestra), sino, como se suele denominar, a la *conveniencia* —palabra que, pese a ser técnica, encierra una divertida ambigüedad— , que suele interpretarse cortésmente como la facilidad de acceso del investigador a los datos. Por ejemplo, el caso del encuestador que se aposenta en un mercado al aire libre o en una estación de tren para realizar su cuestionario electoral, ¿Puede esperar que su muestra sea representativa del total de la población con derecho a voto? Evidentemente no, ya que la clase de personas que con más frecuencia acuden a mercados al aire libre o a estaciones de tren son una muestra *seleccionada*, sesgada hacia cierto tipo socioeconómico, de género, de edad...Igualmente la encuesta a domicilio *no concertada de antemano, o con revisita por ausencia*[**40**][**41**], sino la que se conforma a partir de las respuestas obtenidas puerta a puerta: efectivamente, dependiendo del tramo horario en que se realice la encuesta, se estará privilegiando a cierto grupo social u a otro: si la encuesta se realiza, digamos, a media mañana, se obtendrá una mayor proporción de personas paradas, jubiladas, amas de casa, etc. pero veamos algún ejemplo que demuestra que el pensamiento estadístico — o el análisis crítico de datos en general, si se prefiere— debe saber trascender las tablas llenas de variables y de frecuencias, y empezar siempre por preguntarse acerca de la procedencia de esos datos.

El Barómetro[5] del Centro de Investigaciones Sociológicas (CIS)[6] es uno de los recursos más expuestos a la luz de la prensa, consistente en un estudio estadístico de periodicidad mensual (excepto Agosto) basado en encuestas cuyo propósito es la medición del *"estado de la opinión pública del momento"*, con lo cual cabe razonar que uno de sus pocos resultados cuantitativos que responden a categorías no especulativas —y la razón principal del interés de los medios por él— es la intención de voto. Suele dispensársele un trato mejor que a la mayor parte de los sondeos y una respetabilidad mayor que a la del resto de sondeos, quizá por el hecho de que, al ser un organismo autónomo pero de carácter administrativo[7], se supone que cuenta con más medios y mejores datos a su disposición.

No obstante lo cual, o quizá debido a ello, —las voces críticas suelen aparecer con más

---

5 http://www.cis.es/cis/opencm/ES/11_barometros/index.jsp [consultado Mayo 2015]
6 http://www.cis.es/cis/opencms/ES/index.html [consultado Mayo 2015]
7 Cuyo funcionamiento está regulado por ley: Real Decreto 1526/1990, de 8 de noviembre [consultado Mayo 2015] http://www.cis.es/cis/export/sites/default/-Archivos/Legislacion/RD_1526-1990.pdf

frecuencia en las entidades serias— podemos leer el siguiente extracto[8]:

> *Siguiendo a H. Lawrence Ross (1963), se entiende por inaccesibilidad la imposibilidad de contactar con un posible entrevistado porque el portero físico de un inmueble impide la entrada al edificio de viviendas. También puede ocurrir que una persona que contesta a la llamada del entrevistador no desea recibir explicaciones sobre el motivo de su visita; y también, y muy importante, porque la ausencia del domicilio por motivos laborales, de ocio, etc., es un problema cada vez más habitual.* ***Todo ello favorece el aumento de la tasa de no respuesta y, con ello, la probabilidad de introducir sesgos que desvirtúen la calidad de los datos obtenido****s. [...]*
>
> *El problema principal de este porcentaje de ausencias es que se está dando mayor probabilidad de ser elegidas a aquellas personas que están en casa cuando el entrevistador realiza su trabajo, como pensionistas, jubilados, amas de casa, estudiantes, etc. Sánchez Carrión (2000) sostiene que «el uso de las sustituciones debe hacerse de manera controlada, teniendo en cuenta que la teoría del muestreo no las justifica, máxime si para cada entrevista hay que hacer seis o más contactos (sustituciones)». Las tres cuartas partes de las entrevistas se hacen entre las 12 y las 20 horas, y 2 de cada 10 durante el fin de semana. Éstos son los días y las horas en los que aumenta la probabilidad de encontrar en el hogar a todos los colectivos que deben estar representados en la muestra. En algunos casos hay ciertas cuotas para las que es necesario emplear tres horas hasta completar el sexo y la edad buscada.*
>
> *Díaz de Rada (2000) señala que la mejor estrategia para solucionar el problema de las ausencias prolongadas del hogar es incrementar el número de visitas. Pero ¿qué ocurre cuando la duración del trabajo de campo es de seis días y en ese espacio de tiempo no es posible volver a visitar el hogar?* ***El muestreo por cuotas no contempla la revisita para solucionar el problema de las ausencias*** *y, finalizado el trabajo de campo, el número de entrevistas conseguidas y diseñadas es prácticamente el mismo. Se completa la muestra a costa de «barrer» prácticamente en su totalidad las secciones electorales.* ***¿Estamos sobrerrepresentando a aquellas personas que están siempre en casa****, como sostienen Sánchez Carrión (2000), Jesús Alderete (1996) y otros autores?*

ACTIVIDADES

1º) Visita la página http://www.cis.es/cis/opencm/ES/11_barometros/index.jsp *y abre el enlace correspondiente (inf.) a "metodología"* http://www.cis.es/cis/export/sites/default/-Archivos/NotasdeInvestigacion/NI004_MetodologiaBarometros_Informe.pdf [consultado Mayo 2015] *Relacionado con tus conocimientos actuales en estadística*[9]*,*

a) ¿Hay algún párrafo que te resulte familiar? Adjunta una imagen.
b) ¿Puedes explicar por qué? Es decir, ¿Qué conceptos que conozcas aparecen en ese párrafo?
c) Sabrías comprobar, con tus conocimientos actuales, si los valores expresados para esos conceptos son correctos?

8 A. Núñez Villuendas, *Incidencias de la entrevista personal en la investigación mediante encuesta,* Reis(CIS) [revista en línea] (2005) [cons. Mayo 2015] 109/05 pp. 219-236. Disponible en: http://www.reis.cis.es/REIS/PDF/REIS_109_091168259907774.pdf
9 *Vid.* [37]

d) El texto en cuestión, ¿Es suficientemente explícito respecto de la metodología utilizada? ¿Te queda alguna duda de cómo se ha realizado efectivamente la encuesta tras su lectura?

a) Suponiendo que el alumno ha realizado las actividades anteriores, es de suponer que extractará el siguiente párrafo[10]:

> - El error de muestreo, bajo el supuesto de muestreo aleatorio simple, se sitúa en ± 1,9% para el conjunto de la muestra, para un nivel de confianza del 95,5% (dos sigmas) y la situación más desfavorable (P=Q).

b) En él aparecen el márgen de error o "error de muestreo", y también se menciona el grado o nivel de confianza: además, el valor de 95,5% es muy similar al manejado anteriormente.

c) Para comprobar lo pedido es necesario —como señalará algún alumno avispado— un dato que no aparece en el párrafo en cuestión, esto es, el tamaño de la muestra, (que no es difícil de encontrar en el documento) n=2500. Entonces, aplicando $r=1/\sqrt{n}$, tenemos que $r=1/\sqrt{2500}=0{,}02=2\%\simeq 1{,}9\%$ para un nivel de confianza muy similar: 95% ≈95,5%.

d) La mayor parte de los alumnos respodería que sí, que es incluso exhaustiva, fijándose sobre todo en este párrafo[11]:

> - Procedimiento de muestreo polietápico, estratificado por conglomerados, con selección de las unidades primarias de muestreo (municipios) y de las unidades secundarias (secciones) de forma aleatoria proporcional, y de las unidades últimas (individuos) por rutas aleatorias y cuotas de sexo y edad. Los estratos se han formado por el cruce de las 17 comunidades autónomas con el tamaño de hábitat dividido en siete categorías: menor o igual a 2.000 habitantes; de 2.001 a 10.000; de 10.001 a 50.000; de 50.001 a 100.000; de 100.001 a 400.000; de 400.001 a 1.000.000; y más de 1.000.000 de habitantes.

Llegado este momento, es cuando el profesor debe señalar la humilde frasecilla que resume todo el trabajo de campo[12]:

> - Encuesta personal realizada en hogares.

Muy bien, pero...¿Cuál es realmente el individuo objetivo (parte de la muestra), según el texto, es decir, el diseño de la encuesta? ¿Un individuo determinado al que se va a visitar? ¿El primero que abra la puerta? ¿Se ha seleccionado un individuo en particular o un domicilio? Y si se ha seleccionado un individuo, ¿Se ha concertado una cita previa? ¿Se vuelve al día siguiente si no se ha conseguido contactar con él? Y si, finalmente, no se consigue, se elige otro individuo objetivo? ¿Y qué sucede si finalmente, el individuo se presenta pero se niega a contestar a una o a todas las preguntas?

---

10 Centro de Investigaciones Sociológicas, *Nota de investigación sobre la metodología general de los barómetros mensuales del centro de investigaciones sociológicas.*[recurso en línea] Disponible en [consultado Mayo 2015]: http://www.cis.es/cis/export/sites/default/-Archivos/NotasdeInvestigacion/NI004_MetodologiaBarometros_Informe.pdf
11 *Vid.* 9
12 *Vid.* 9

Todas estas dudas y más deben plantearse al alumno en orden a que reflexione por sus propios medios, y a que, ante cualquier clase de discurso, no se deje apabullar en un primer momento por la terminología técnica y sepa conservar la suficiente distancia intelectual como para ser capaz de elaborar un análisis crítico, *incluso* en una situación en que no se poseen *todas* las herramientas para examinar *todos* los elementos de un discurso.

2) *Visita la página* http://www.cis.es/cis/opencm/ES/11_barometros/index.jsp, *abre uno cualquiera de los* barómetros *y localiza la ficha técnica.*

a) Adjunta el texto correspondiente al epígrafe "Procedimiento de muestreo"
b) Comprueba, según tus conocimientos, si el margen de error expresado es correcto.
c) Visita la página correspondiente a la metodología de la EPA (Encuesta de Población Activa) en http://www.ine.es/daco/daco43/notaepa.htm [cons. Mayo 2015] *y abre el fichero correspondiente a* "descripción de la encuesta". *Localiza los capítulos denominados* "Diseño de la muestra" y "Calidad de las estimaciones"*¿Cuál es la diferencia principal que, a primera vista, puedes establecer entre estos capítulos y el epígrafe anteriormente adjuntado ("procedimiento de muestreo") del barómetro del CIS? Indicación: no es preciso leerlo entero, ni tan siquiera entender todos sus puntos: sólo basta considerar la extensión de ambos textos.*
d) Dentro de el capítulo "Calidad de las estimaciones" (EPA) *localiza el epígrafe* 9.2 "Errores ajenos al muestreo". *Aparte de la extensión, ¿Se puede deducir del epígrafe adjuntado en a) que se cumplen las condiciones del que ahora estamos estudiando? ¿Cuál de los dos estudios te parece más fiable y por qué? ¿A qué puede deberse, si existe, la diferencia en definición y fiabilidad de ambos estudios?*

a) El texto que se pide es el siguiente[13]:

Procedimiento de muestreo:

Polietápico, estratificado por conglomerados, con selección de las unidades primarias de muestreo (municipios) y de las unidades secundarias (secciones) de forma aleatoria proporcional, y de las unidades últimas (individuos) por rutas aleatorias y cuotas de sexo y edad.

Los estratos se han formado por el cruce de las 17 comunidades autónomas, con el tamaño de hábitat, dividido en 7 categorías: menor o igual a 2.000 habitantes; de 2.001 a 10.000; de 10.001 a 50.000; de 50.001 a 100.000; de 100.001 a 400.000; de 400.001 a 1.000.000, y más de 1.000.000 de habitantes.

Los cuestionarios se han aplicado mediante entrevista personal en los domicilios.

Error muestral:

Para un nivel de confianza del 95,5% (dos sigmas), y P = Q, el error real es de ±2,0% para el conjunto de la muestra y en el supuesto de muestreo aleatorio simple.

b) Procedimiento análogo al descrito en la actividad anterior 1.c)

13 Centro de Investigaciones Sociológicas, *Estudio CIS nº 3045.Barómetro de Noviembre*[2014].*Ficha Técnica* [recurso en línea] Disponible en [consultado Mayo 2015]: http://www.cis.es/cis/export/sites/default/-Archivos/Marginales/3040_3059/3045/FT3045.pdf

c) Aunque hacemos hincapié en la extensión, en realidad lo que se debería poner de manifiesto ante el alumno es que lo que realmente importa es la *definición:* en el caso de la EPA es muy considerable, mientras que la definición del barómetro del CIS es mínima; y lo que es peor, es fallida, da lugar a ambigüedad, como se expresará en el siguiente apartado. (adjuntamos extracto de la EPA para comodidad del lector[14])

Generalmente se calcula para ambos sexos y para cada uno de ellos por separado. Los intervalos de edad suelen ser quiquenales o decenales.

- Tasa específica potencial de actividad.

La tasa específica potencial de actividad para un intervalo de edad determinado es el cociente entre la población el potencial de activos (activos + activos potenciales) de esas edades y la población correspondiente al intervalo.

8 Diseño de la muestra

La Encuesta de Población Activa es una investigación estadística de tipo continuo enmarcada dentro del diseño de la Encuesta General de Población, modelo utilizado por el INE como marco para las encuestas dirigidas a los hogares. Los aspectos más importantes del diseño son:

8.1 TIPO DE MUESTREO. ESTRATIFICACIÓN

El tipo de muestreo utilizado es un muestreo bietápico con estratificación de las unidades de primera etapa.

Las unidades de primera etapa están constituidas por las secciones censales. Las unidades de segunda etapa son las viviendas familiares principales y los alojamientos fijos (chabolas, cuevas, etc.).

Dentro de las unidades de segunda etapa no se realiza submuestreo alguno, recogiéndose información de todas las personas que tengan su residencia habitual en las mismas.

Las secciones censales se estratifican atendiendo a un doble criterio:

Criterio geográfico (de estratificación):

Las secciones se agrupan en estratos de acuerdo con la provincia y el tipo de municipio (según importancia demográfica) al que pertenecen.

Criterio socioeconómico (de subestratificación):

Dentro de cada estrato geográfico las secciones censales se agrupan en subestratos atendiendo a la categoría socioeconómica de los hogares ubicados en la sección.

8.2 TAMAÑO DE LA MUESTRA. AFIJACIÓN

El tamaño de la muestra actual es de 3.484 secciones, investigándose un número teórico de 18 viviendas por sección. La muestra final se eleva a unas 65.000 viviendas al trimestre.

La distribución de secciones por provincias se ha realizado utilizando una afijación de compromiso entre la uniforme y la proporcional. Dentro de cada provincia la fijación entre estratos se ha basado en la proporcional.

8.3 SELECCIÓN DE LA MUESTRA

La selección de la muestra se ha realizado de tal forma que dentro de cada estrato cualquier vivienda tenga la misma probabilidad de ser seleccionada, es decir, se obtengan muestras autoponderadas. Para ello las secciones se seleccionan con probabilidad proporcional al número de viviendas familiares existentes en las mismas, de acuerdo con los datos del último Censo o Padrón. Dentro de cada sección se seleccionan las viviendas mediante la aplicación de un muestreo sistemático con arranque aleatorio.

8.4 DISTRIBUCIÓN DE LA MUESTRA EN EL TIEMPO. TURNOS DE ROTACION

Cada periodo de encuesta dura un trimestre, estando la muestra distribuida uniformemente a lo largo de él, por lo que cada una de las secciones es visitada en una de las 13 semanas del trimestre.

La totalidad de la muestra está dividida en tres submuestras mensuales independientes, representativas cada una de ellas de to-da la población.

En los sucesivos trimestres, la muestra de secciones permanece invariable. Sin embar-

14 Instituto Nacional de Estadística, *EPA. Descripción de la Encuesta* [recurso en línea] Disponible en [consultado Mayo 2015]: http://www.ine.es/epa02/descripcion%20encuesta.pdf

go, con objeto de evitar el cansancio en las familias que colaboran, la muestra de viviendas se renueva parcialmente en una sexta parte.

A estos efectos, la muestra total de secciones se encuentra dividida en seis grupos, que denominamos **turnos de rotación** renovándose cada trimestre las viviendas de las secciones que pertenecen a un determinado turno.

Por tanto una vivienda seleccionada permanece en la muestra y es entrevistada a lo largo de seis trimestres consecutivos.

8.5 ESTIMADORES

Se utilizan estimadores de razón, tomando como variable auxiliar las Proyecciones Demográficas de Población elaboradas cada trimestre por el INE.

A partir del primer trimestre de 2002, se aplican a los estimadores **técnicas de reponderación** con objeto de ajustar las estimaciones obtenidas a partir de la muestra, a la información procedente de fuentes externas.

Esta técnica consiste en modificar los factores de elevación originales deducidos del diseño, de acuerdo a las siguientes condiciones: en primer lugar, la población de 16 y más años estimada por sexo y grupo de edad en cada comunidad autónoma y esa misma población en cada provincia, deben coincidir, en ambos casos, con los correspondientes efectivos de las proyecciones de población, es decir, las cifras de población estimadas coinciden con las reales. En segundo lugar, debe verificarse que la variación de los factores de elevación reponderados respecto a los factores originales deducidos del diseño sea mínima.

9 Calidad de las estimaciones

Las estimaciones de toda encuesta realizada por muestreo vienen afectadas por dos tipos de errores:

a) Errores de muestreo

b) Errores ajenos al muestreo

Los errores de muestreo aparecen como consecuencia de la utilización de una parte de la población para estimar características de todo el conjunto poblacional. En general, a medida que aumenta el tamaño de la muestra disminuye el error de muestreo, aunque también influyen en él las características del diseño y la naturaleza de la variable investigada.

Los errores ajenos al muestreo incluyen toda una serie de errores que se producen en las distintas fases de realización de una encuesta, desde el diseño del cuestionario hasta la publicación de los resultados finales, (definiciones deficientes, errores en el marco, falta de respuesta en las unidades informantes, errores de codificación y grabación, etc.).

9.1 ERRORES DE MUESTREO

Un buen indicador del error de muestreo de un estimador es su desviación típica.

En la Encuesta de Población Activa para el cálculo del error de muestreo de una determinada característica se utiliza el método de las semimuestras reiteradas.

La fórmula que permite estimar el error de muestreo es:

$$D(\hat{X}) = \sqrt{\frac{\sum_{i=1}^{r}\left(\hat{X}_i - \hat{X}\right)^2}{r}}$$

siendo:

r = el número de reiteraciones utilizado

$\hat{X}_i$ = la estimación obtenida con la i-ésima reiteración aplicando el mismo proceso de estimación que para la muestra completa

$\hat{X}$ = la estimación obtenida con la muestra completa

Cada reiteración es una submuestra formada por un número de secciones equivalente al 50 por ciento de la muestra total. Se utilizan en la EPA 40 reiteraciones que se han formado emparejando las secciones dentro de cada estrato y asignando aleatoriamente en cada par la primera sección a 20 reiteraciones y la segunda a las 20 restantes. De esta forma cada sección figura en la mitad de las reiteraciones.

En las tablas aparece el coeficiente de variación en porcentaje, cuya expresión es:

$$C.V.\left(\hat{X}\right) = \frac{D\left(\hat{X}\right)}{\hat{X}} \cdot 100$$

El error de muestreo nos proporciona un intervalo numérico en el que existe una cierta confianza, medida en términos de probabilidad, de contener al verdadero va-lor de la característica estimada. El interva-lo de confianza más utilizado es el comprendido entre la estimación menos dos veces el error de muestreo y la estimación más dos veces el error de muestreo. Este intervalo tiene una confianza del 95% de encontrar en él al valor verdadero que se está estimando.

Toda estimación con un error de muestreo elevado debe ser tomada con reservas; aunque debe ser el usuario el que, de acuerdo con el grado de fiabilidad que precise, determine si un dato con un cierto error de muestreo le es útil o no para la to-ma de decisiones.

ERRORES AJENOS AL MUESTREO

En este grupo se incluyen todos los errores que no se producen como consecuencia de la utilización de muestras para la estimación de una determinada característica.

En la EPA se realiza un estudio de la falta de respuesta originada por las negativas de las unidades informantes, así como de los errores de cobertura y contenido.

9.2.1 Falta de respuesta

En caso de falta de respuesta parcial (faltan algunos datos de la persona encuestada) se realiza una imputación automática. La falta de respuesta total (vivienda encuestable que no ha podido encuestarse por ausencia, negativa o por ser inaccesible en el momento de la entrevista) recibe un tratamiento distinto según se trate de la primera entrevista o de una posterior.

En primera entrevista las viviendas ausentes y las inaccesibles son visitadas nuevamente hasta conseguir la información. Si la vivienda sigue siendo inaccesible o la familia sigue ausente se vuelve a visitar al trimestre siguiente.

Las negativas en primera entrevista se sustituyen aleatoriamente por otras viviendas de la misma sección. Si la negativa se produce en segunda o sucesivas entrevistas no se sustituye, aunque sí se vuelve a visitar los trimestres posteriores, por si hubiera habido un cambio de actitud o el grupo humano fuera otro distinto.

Las viviendas no encuestables (destinadas a fines distintos a los de vivienda familiar, vacías y de temporada) se visitan los trimestres posteriores por si han pasado a ser encuestables.

El tratamiento de la falta de respuesta total en entrevistas posteriores a la primera consiste en repetir los datos de la vivienda obtenidos en el trimestre anterior. Esta imputación se realiza sólo la primera vez que se produce la incidencia.

9.2.2 Errores de cobertura y contenido

- Errores de cobertura: son los originados por la omisión o inclusión errónea de unidades en la encuesta.
- Errores de contenido: se producen cuando alguna o algunas características del entrevistado están erróneamente recogidas en el cuestionario.

Para evaluar ambos errores, se realiza una encuesta de evaluación consistente en seleccionar una submuestra de viviendas en las que se repite la entrevista. Los datos obtenidos en la primera entrevista o entrevista original (E.O.) se cotejan con los obtenidos en la segunda entrevista o entrevista repetida (E.R.).

A partir de esta comparación se obtienen indicadores tanto de la cobertura de personas y viviendas, como de los errores de contenido de las características más importantes analizadas en la encuesta.

d) De la frase “los cuestionarios se han aplicado mediante entrevista personal en los domicilios” en realidad, se puede deducir muy poca cosa. Cabe hacerse todas las preguntas que expresamos en 1.d) y quizá algunas otras más. En cambio, en el epígrafe correspondiente de la EPA, encontramos una verdadera definición del método seguido en el trabajo de campo, conducente a obtener una muestra de plausible aleatoriedad, es decir, no influida:

> 9.2.1 Falta de respuesta
>
> En caso de falta de respuesta parcial (faltan algunos datos de la persona encuestada) se realiza una imputación automática. La falta de respuesta total (vivienda encuestable que no ha podido encuestarse por ausencia, negativa o por ser inaccesible en el momento de la entrevista) recibe un tratamiento distinto según se trate de la primera entrevista o de una posterior.
>
> En primera entrevista las viviendas ausentes y las inaccesibles son visitadas nuevamente hasta conseguir la información. Si la vivienda sigue siendo inaccesible o la familia sigue ausente se vuelve a visitar al trimestre siguiente.
>
> Las negativas en primera entrevista se sustituyen aleatoriamente por otras viviendas de la misma sección. Si la negativa se produce en segunda o sucesivas entrevistas no se sustituye, aunque sí se vuelve a visitar los trimestres posteriores, por si hubiera habido un cambio de actitud o el grupo humano fuera otro distinto.
>
> Las viviendas no encuestables (destinadas a fines distintos a los de vivienda familiar, vacías y de temporada) se visitan los trimestres posteriores por si han pasado a ser encuestables.
>
> El tratamiento de la falta de respuesta total en entrevistas posteriores a la primera consiste en repetir los datos de la vivienda obtenidos en el trimestre anterior. Esta imputación se realiza sólo la primera vez que se produce la incidencia.

Podría decirse en conjunto que la encuesta del CIS es más básica, más “casera” que la EPA (también más barata, probablemente) lo cual quizá venga determinado por la diferencia de medios y supuestos legales de los organismos que las ejecutan: la EPA no depende del CIS, sino del Instituto Nacional de Estadística (INE). Otra posible causa es que la EPA está armonizada por directiva europea con la ECFT[15](Encuesta Comunitaria de Fuerza de Trabajo), debiendo ser compatible en metodología con esta última; esto es, que la EPA cumple requisitos normalizados (de la misma manera que se cumplen la normas UNE en la industria) y el barómetro del CIS no tiene esta obligación.

15 Instituto Nacional de Estadística, *La EFCT y la EPA* [recurso en línea] Disponible en [consultado Mayo 2015]: http://www.ine.es/daco/daco42/daco4211/ecft_epa.pdf

Media VS. Mediana

Bajo el mismo título que el utilizado por Hoaglin, Mosteller y Tukey en **[26]** comenzamos el presente capítulo, dedicado a un viejo malentendido; aunque, también, a una vieja controversia: entre estadística matemática y descriptiva, entre inferencia paramétrica y no paramétrica.

Como se sabe, la media (aritmética) o bien,

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

con x_i cada una de las realizaciones de una variable *x,* (o cada uno de los valores de un conjunto de datos, desde el punto de vista más descriptivo) y *n* el número de tales datos, es la medida de centralización (o de localización, si se prefiere) clásica sobre la que se ha construido la estadística desde sus comienzos (en estrecha dependecia del método de los momentos), y es la más adecuada si los datos son homogéneos; pero es una medida muy sensible a los resultados anómalos, ya vengan dados por errores de medición o por valores atípicos *(outliers)*. La mediana, esto es, un valor tal que es mayor que los valores del 50% de un conjunto de datos ordenados ordenados de menor a mayor, por contra, es una medida de centralización robusta, esto es, que se ve poco influida por la dispersión de conjuntos pequeños de datos o valores *excéntricos*, ya se hayan obtenido por error o no.**[39]**

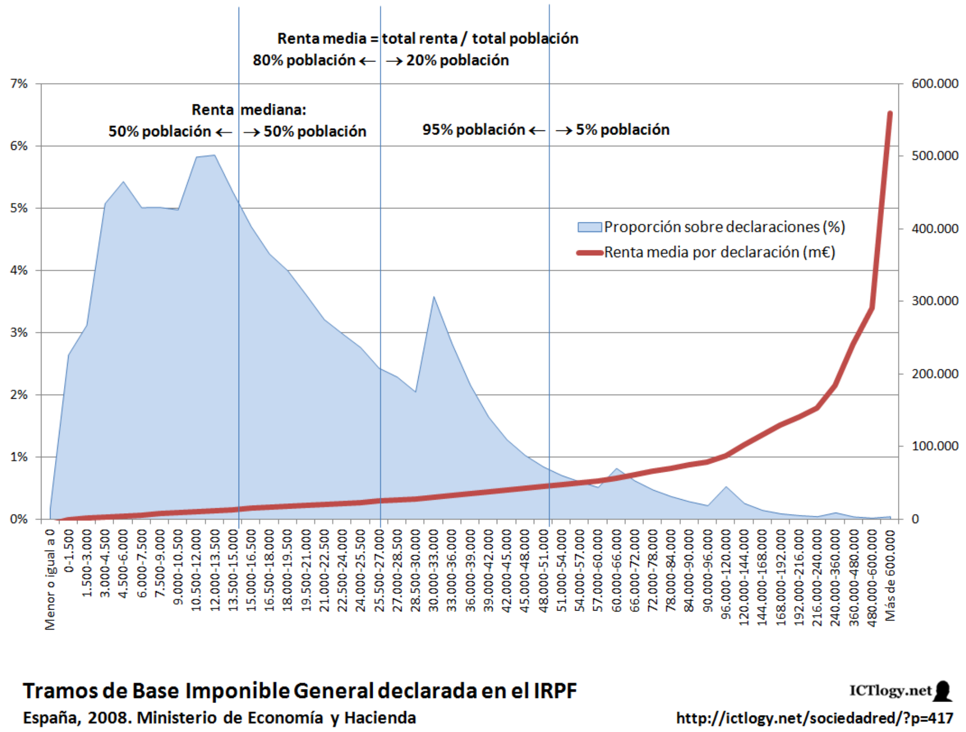
Además, por tener en cuenta sólo el orden de los datos, la mediana se mantiene pese a que se practiquen transformaciones no lineales sobre los datos, es decir, es posible recuperar su valor del conjunto de datos transformados, cosa que sólo puede hacerse aproximadamente con la media. Por razones como estas, entre otras, la mediana es preferida por los estadísticos orientados al análisis exploratorio de datos y a la inferencia no paramétrica (generalmente procedentes de los campos de las ciencias aplicadas o del comportamiento, donde las distribuciones de probabilidad suelen no ser asimilables a las canónicas,y aparte de mostrarse asimétricas, e intratables en general), llegando a sostenerse que deberían preferirse siempre en escalas ordinales **[42]**.**[45]**. Pero los estadísticos afines a las disciplinas más relacionadas con la probabilística (mecánica estadística, procesos estocásticos, control de calidad, etc.) encuentran que, en muchas ocasiones, la mediana es mucho más difícil de calcular y de tratar matemáticamente que la media, y que ésta última cuenta, además, con la ventaja de que puede completarse con la información dada por la forma de distribución de la variable; y que es esa distribución de probabilidad correspondiente al fenómeno determinado la que debe investigarse, en vistas a la formalización teórica, no empírica, del mismo. En realidad no son muchos los probabilistas y estadísticos matemáticos que se molestan en tener la mediana en consideración, o acaso mencionarla, salvo excepciones **[12]**.

Lo que subyace al debate es la diferencia entre la estadística descriptiva y la probabilística**[11]**: la que se plantea como fin último la obtención de resultados o predicciones de base empírica y la que pretende modelizar matemáticamente la totalidad de un fenómeno (ya sea físico, económico, social...). Esta discusión afecta no sólo al problema que describimos, sino a la misma personalidad científica y/o profesional del estadístico: lo cual también concierne directamente al enfoque del currículo matemático correspondiente al bloque de estadística en enseñanzas medias (vid. supra). Por nuestra parte, en lo que sigue nos atendremos a lo que un alumno de educación secundaria pueda aprovechar.

La distribución del ingreso (renta)

No obstante el anterior debate, la vida cotidiana abunda en ámbitos donde es relevante practicar la distinción entre media y mediana, y en los que, caso de elegir sólo una entre ambas, ésta última debería ser la más tenida en cuenta. El ejemplo más típico de distribución asimétrica, sobre la cual el conocimiento de la media desnuda puede llevarnos a graves equívocos, es el de la distribución de la renta familiar. Se han realizado —y se siguen realizando— múltiples intentos de parametrizar la distribución de rentas familiares según determinadas funciones de densidad **[13]**, empíricamente o en orden a contrastar diversos modelos subyacentes hipotéticos, aunque parece que la Distribución Gamma **[8]** es la que mayor grado de adhesión consigue, por lo adaptable de su forma y por estar definida en un dominio de valores positivos **[7]**[1].

Mas todo esto cae fuera del contenido curricular que afecta al alumno que tenemos en mente: lo que quizá no esté tan lejos del dominio de su experiencia viene reflejado en el siguiente gráfico[2]:



más, abajo, el autor del artículo explica:

1 La distribución de renta familiar presenta, como se verá, la típica asimetría o sesgo "hacia la derecha" típico de variables definidas en un dominio acotado inferiormente (en nuestro caso, 0 € de ingreso (renta)

2 Ismael Peña-López, *Clase media, desigualdades, recortes e impuestos para los ricos* [entrada de blog] 30 Sept. 2011 [consultado Mayo 2015], en *ICTlogy* (ictlogy.net). Disponible en: http://ictlogy.net/sociedadred/20110930-clase-media-desigualdades-recortes-e-impuestos-para-los-ricos/

Hay algunos números que vale la pena resaltar.

*El primero es el valor de la **renta media** española, es decir: el total de rentas declaradas dividido por el total de declarantes. Esa renta media se situó en 2008 en 26.379 € (aproximadamente). Sin embargo, este dato es engañoso y suele ejemplificarse con el conocido "si yo tengo dos manzanas, y tú no tienes ninguna manzana, de media tenemos una manzana cada uno". En realidad, en España vemos que hay muchísima gente que está por debajo de la renta media y muy poca que está por encima. En concreto, **el 80% de la población está por debajo de la renta media**.*

*Mucho más realista que el dato anterior es la **renta mediana**, es decir, cuánto cobra la persona que tiene a la mitad de la población cobrando más que ella y la otra mitad menos. Esta renta mediana se situó en España en 2008 en 14.921 €, es decir, la mitad de los españoles declararon cobrar menos de 14.921 € brutos al año.*

Por si hay algún escéptico (como el que redacta estas líneas) que no considera los datos fiscales fiables a la hora de medir el verdadero ingreso, sobre todo de las rentas más altas, el mismo autor ha explicado ya antes:

*[...]E**sta medida es muy grosera para cuantificar los ingresos de una persona**, y su precisión disminuye a medida que las personas tienen mayor proporción de rentas no provenientes del trabajo (compensaciones en especie, ingeniería financiera, evasión fiscal, etc.). También quedan fuera personas sin la obligatoriedad de declarar, normalmente porque sus rentas son demasiado bajas. Sin embargo, el análisis, aún (insisto) con sus debilidades, es suficientemente válido para la reflexión que queremos hacer aquí.*

A lo que yo añadiría que, supuesto el caso de que obtuviésemos cifras reales, es de suponer que la gráfica que resultase de las mismas estaría aún mucho más sesgada a la derecha, de la media nos daría entonces una impresión mucho más distorsionada del verdadero estado de cosas. Por otra parte, observando detenidamente la gráfica se parecia que *los intervalos a partir de 30.000€ tiene el doble de amplitud (3.000€)*, lo que *vuelve a suceder* al llegar a los 60.000€ (6.000€) y a los 96.000€ (24.000€!!)lo que indica que la gráfica ya es, en parte, deshonesta (o, más bien, creemos, es *la selección de los datos* por parte de la Agencia Tributaria —en los cuales se basa la gráfica— la que es deshonesta) En las actividades siguientes propondremos a los alumnos que se hagan preguntas sobre una gráfica análoga(y contemporánea) a la anterior, hallada en otro artículo[3], y que obvia las explicaciones sobre media y mediana.

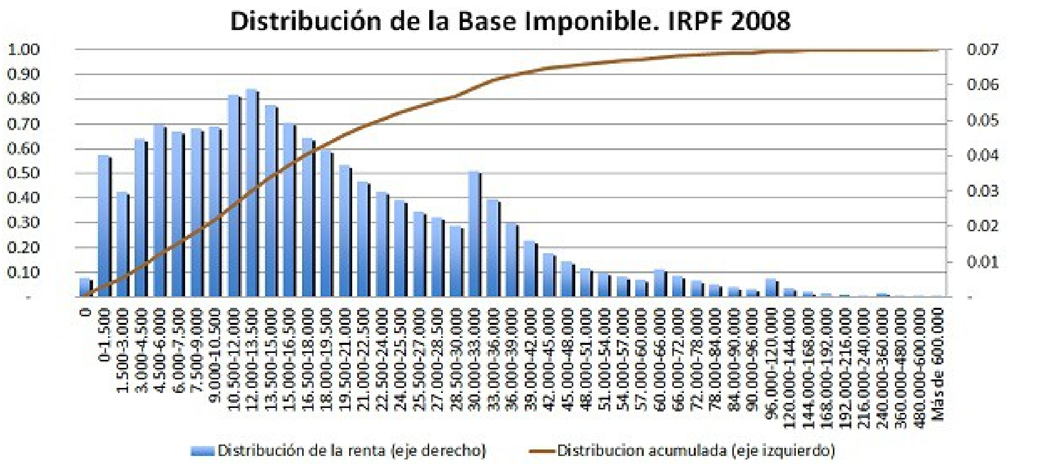
ACTIVIDADES

1º) *Observa la siguiente gráfica, correspondiente a la distribución de la renta de los españoles, realizada sobre la* base imponible *general declarada en el* Impuesto sobre la Renta de las Personas Físicas *(IRPF) de 2008. Ten en cuenta que:*

· *El eje de abscisas se ha dividido en intervalos de determinadas cantidades de €.*

3 J.I. Conde-Ruiz y J.R. Ramírez, *NeG Visual y Básico. La Distribución de la Renta en España según los datos oficiales del IRPF* [entrada de blog] 20 Marzo 2012 [consultado Mayo 2015] en (*Nada es Gratis) NeG. Un blog de economía casi siempre en español (nadaesgratis.es).* Disponible en: http://nadaesgratis.es/rubio-ramirez/neg-visual-y-basico-la-distribucion-de-la-renta-segun-la-bases-imponibles-del-irpf

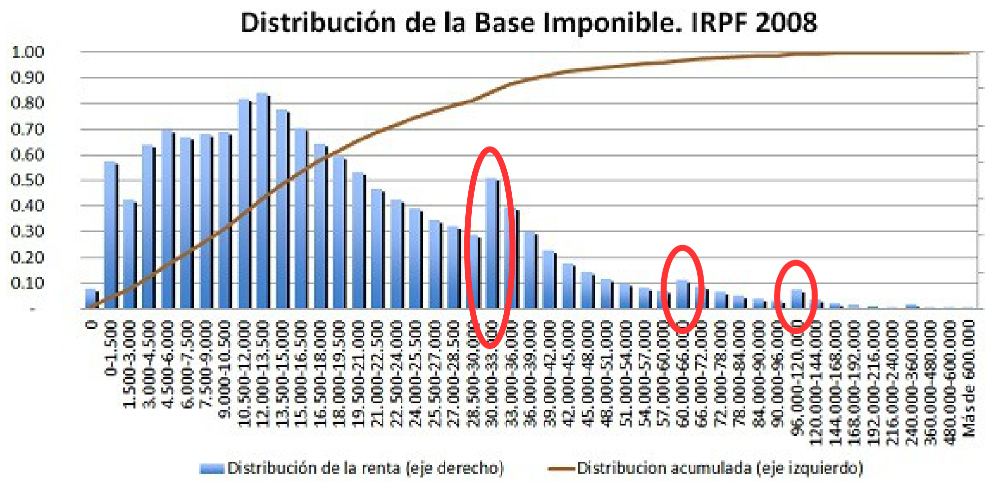
· *<u>El eje de ordenadas **derecho**</u> corresponde a la proporción de la población total que gana una cantidad entre los valores extremos del intervalo, esto es, la altura de cada barra.*
· *<u>El eje de ordenadas **izquierdo**</u> corresponde al porcentaje de personas (declarantes) que va acumulándose en el sentido del eje de abscisas: es decir, es el eje de ordenadas <u>de la línea de color ocre</u>.*



Fuente:nadaesgratis.es http://nadaesgratis.es/wp-content/uploads/tabla-neg10.jpg

a) Se supone que la gráfica representa los ingresos brutos de cada español que hace la declaración de la renta; es decir, "lo que gana" cada español (que gana algo). Calcula la media aritmética aproximadamente tomando los valores de la gráfica hasta 600.000€ (recuerda que las barras expresan frecuencias relativas)

b) Observa los "picos" que aprecen en la gráfica:



¿A qué crees que son debidos?

c)Como habrás podido observar, a partir de 30.000€ los intervalos duplican su amplitud (pasan de 1.500€ a 3.000€). ¿Vuelve a suceder esto más veces en la gráfica? ¿Qué le sucedería al diagrama de barras si lo reformásemos con intervalos de la mis-

ma amplitud? ¿Y a la posición de la media aritmética sobre ella? ¿Qué efecto tendría tal reforma en la percepción de la distribución de renta: que hay más españoles que ganan más dinero, o que la diferencia entre lo que gana un español de los de la izquierda de la tabla y lo que gana uno de la derecha es mucho mayor?

d) Ahora fíjate en la línea de color ocre: ¿Qué medida de localización que conozcas puede ubicarse con su ayuda? ¿Cuál es su valor?

e) Entre la media aritmética y el valor anteriormente hallado, ¿Cuál crees que refleja mejor "lo que ganan de media los españoles" y cuál "lo que gana un español medio"? ¿Qué valor te da más, o mejor información sobre la situación económica de los ciudadanos españoles en su conjunto?

f) Si tuvieras que jugarte una cena con un amigo a que una cifra que tú digas está más cerca de lo que gana una persona cualquiera que pasa por la calle (y suponiendo que os diga la verdad si se lo preguntáis) que la cifra que diga tu amigo, ¿Alrededor de qué valor la escogerías: de la media, o de la mediana antes halladas?

a) Como se sabe, la media aritmética $\overline{X}$ es igual al sumatorio de los productos de los valores de marca de clase $\left(x_i = \left(x_{i\inf} + x_{i\sup}\right)/2\right)$ por sus frecuencias relativas correspondientes:

$$\overline{X} = \sum_{i=1}^{n} x_i \cdot f_{ri} = 750 \cdot 0{,}04 + 1750 \cdot 0{,}03 + 2750 \cdot 0{,}045 + \ldots + 0{,}0003 \cdot 588000 \approx 26500 \pm \varepsilon$$

donde ε suponemos que será el error de medida debido al hecho de que la gráfica no tiene mucha definición.

b) Los "picos" se deben a que la amplitud de los intervalos se multiplica, en 30.000€ por 2(3.000€), en 60.000€ otra vez por 2 (6.000€), y en 96.000€ por 4 (24.000€)

c) Si modificásemos la gráfica con intervalos de la misma amplitud, los "picos" desaparecerían, la gráfica sería mucho más "larga" y mucho más sesgada hacia la derecha: la línea de color ocre, en lugar de presentar ese aumento de pendiente a la altura de 30.000€ sería mucho más "suave" y "tendida". La media aritmética estaría, relativamente, mucho más lejos del centro de la gráfica reformada de lo que está en la gráfica sin modificar. Por supuesto, si modificásemos la gráfica en el sentido que se pregunta, podría percibirse que la gran masa de los españoles *gana mucho menos dinero* que aquellos que disponen de rentas altas; tal y como está configurada la gráfica del ejercicio, este hecho está encubierto o disimulado.

d) La línea de color ocre marca, en su intersección con los valores de las ordenadas, que reflejan la frecuencia relativa acumulada, los puntos de abscisas que corresponden a los distintos deciles de población: es decir, donde la línea corta a la horizontal marcada como 0,01 tenemos como abscisa el primer decil, en 0,02 el segundo, etc. Gracias a esa representación, podemos fijar los valores de la mediana y los cuartiles:

· mediana ≈ 14.250€
· 1^{er} Cuartil ≈ 7.500€
· 3^{er} Cuartil ≈ 24.750€

e) La media aritmética podría definirse con la primera de las expresiones entrecomilladas, pero eso sólo me informa sobre la totalidad de lo que ganan los españoles tomados en su conjunto y sobre la cifra absoluta de población que gana dinero (o que hace la declaración de la renta). Si quiero hacerme una idea sobre lo que gana la proporción más representativa de la población, esto es, lo que gana un ciudadano cualquiera, tendré que fijarme en la mediana. Luego, si lo que me interesa es conocer el status económico y social de los españoles, la mediana es más informativa, pues refleja un valor válido para la mayoría de los españoles.

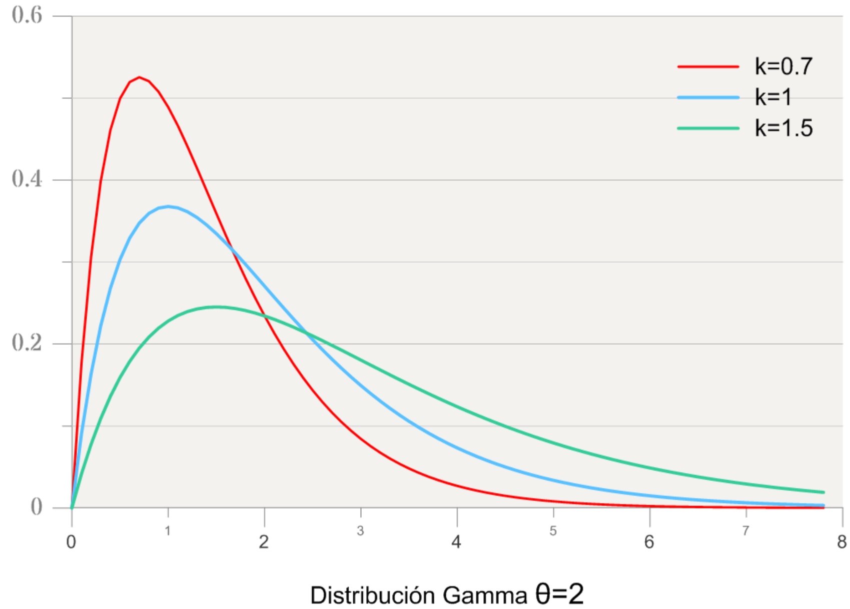
f)Aunque es bastante inseguro, escogiendo la mediana se escoge el valor de los ingresos de la parte más representativa de la población.

Tasa de Supervivencia

O "La Mediana Tampoco es Suficiente", como quizá también podríamos haber llamado a este apartado. Ya veremos por qué. Ahora nos ocuparemos del tema al que se refiere concretamente el título, que consiste en la esperanza de vida o Tasa de supervivencia (Survival rate) de los diagnosticados de diversas enfermedades graves sin cura conocida, pero con tratamiento más o menos asequible.

El tópico en cuestión pertenece a la familia de problemas de tiempo de espera, que se prestan a ser modelizados**[5]** con distribuciones muy parecidas a las del caso anterior[4], aunque clásicamente suele asociársele la distribución exponencial (o exponencial negativa, según otros autores) o alguna de sus relacionadas[5]. Y, como en el caso anterior, la mediana da una mejor información desde el punto de vista descriptivo[6], es decir, acerca del tiempo de supervivencia de un individuo tras el diagnóstico.

Fig.1 Distribución Gamma para algunos valores parámetricos [elab. propia]



4 A menudo se han propuesto modelos conceptuales para la distribución de la renta por familias sobre parámetros basados en tiempos de espera. *vid. [13]*

5 La distribución exponencial constituye un caso particular tanto de la Gamma como de la Weibull, por ejemplo.

6 Lo cual no quiere decir que el estudio científico de la tasa de supervivencia por cada clase de enfermedad no se deba realizar por medios paramétricos, a fin de comprender el fenómeno en su totalidad, y no solo desde el punto de vista del interés inmediatamente informativo.

Llegados a este punto referiremos un caso peculiar en el que parece —insistimos: parece— que la mediana tampoco resulta del todo informativa: un caso referido al conocido paleontólogo, biólogo e historiador de la ciencia, amén de famoso divulgador de temas relacionados con la biología evolutiva[7] Stephen Jay Gould, referido por él mismo[8]:

> *En 1982 supe que sufría de mesotelioma abdominal, un raro y grave caso de cáncer normalmente asociado a la exposición al amianto* [...]*La literatura* [médica] *no pudo haber sido más brutal: el mesotelioma es una enfermedad incurable, con una* ***mortalidad mediana de ocho meses*** *después de su descubrimiento.*[...]
> [...] *¿Qué significa* "mortalidad mediana de ocho meses" *en lenguaje ordinario? Sospecho que la mayoría de la gente, sin especial educación estadística, leería esa frase como: "Probablemente habré muerto en ocho meses"*[...]
> [...] *Pero cualquier biólogo evolucionista sabe que la variabilidad es la verdadera e irreductible esencia de la naturaleza. La variabilidad es la dura realidad, no un conjunto de medidas imperfectas de la tendencia central. Las medias y medianas son las abstracciones.*
> [...]*Cuando supe lo de la mediana de ocho meses, mi primera reacción intelectual fue: bueno, la mitad de la gente vivirá más, así que cuáles son mis probabilidades de estar en esa mitad.* [...] *Yo poseía todas las características que conferían una probabilidad de vivir más tiempo: Era joven, mi enfermedad había sido reconocida en una fase relativamente temprana, y recibiría el mejor tratamiento médico de la nación*[...]
> *Otro detalle técnico me dio un consuelo adicional. Inmediatamente me di cuenta de que la distribución de la variabilidad alrededor de la mediana de ocho meses sería casi seguramente lo que los estadísticos llaman* "sesgada hacia la derecha".[...] *después de todo, la izquierda de la distribución ofrece un irrevocable límite en cero (dado que el mesotelioma sólo puede ser diagnosticado al momento de la muerte o antes de esta)*[...] *la distribución debe de estar sesgada hacia la derecha, y yo necesitaba saber hasta qué punto se extendía la cola, puesto que había llegado a la conclusión de que mi perfil favorable me hacía un buen candidato a esa parte de la curva.*
> *La distribución estaba, efectivamente, fuertemente sesgada hacia la derecha, con una larga cola (estrecha en altura, no obstante) que se extendía varios años más allá de la mediana de ocho meses. No vi razón por la que no pudiera estar en esa estrecha cola, y lancé un gran suspiro de alivio*[traducción propia]]

En el postscriptum, Steve Dunn, el administrador de la página, refiere el hecho de que el señor Gould sobrevivió *20 años* a su diagnóstico de cáncer, hasta la edad de 60, es decir, *multiplicando por 30* la famosa mediana de ocho meses, con el curioso añadido de que su muerte no fue producida por el cáncer original, sino posiblemente por algún otro, quizá derivado del tratamiento de aquél.

Ahora bien, ¿Qué valor exacto tiene este testimonio? De hecho, la pregunta puede hacerse más general: ¿Qué valor exacto tiene esta clase de testimonios? Desde una perspectiva estadística, es decir, científica, y que el señor Gould nos perdone desde donde quiera que vayan las almas de los evolucionistas después de la muerte, su valor es **nulo.** ¿Por qué? Porque un solo caso particular, aunque sea el de un afamado científico y reconocido autor, sigue siendo solamente eso, *un caso particular*, una anécdota. Sin duda es un hecho disculpable, dado el carácter dramáticamente personal con que se presentaron en esta ocasión los datos. Podemos, asimismo, reconocer la honestidad con que el autor reconoce sus reacciones, pero eso no debe cegarnos sobre el hecho de que estas reacciones del señor Gould respecto

---

7 S.J. Gould, *La Estructura de la Teoría de la Evolución,* Tusquets, Barcelona, 2004. Trad. de A. García Leal
8 S.J. Gould, *The Median Isn't The Message,* [artículo en línea] 31 Mayo 2002 [consultado Mayo 2015], en *CancerGuide* (cancerguide.org), S. Dunn (adm.). Disponible en: http://cancerguide.org/median_not_msg.html

de su enfermedad fueron muy poco científicas. Si el análisis de datos tiene alguna utilidad, ésta sería, como mínimo, la de rebatir argumentos tales como "pues yo tengo un amigo/familiar/vecino/ que..." Moore y Notz, por ejemplo, nos previenen contra ello **[37]** (pp. xvii-xviii):

> *Una anécdota es una historia impactante que permanece en la memoria precisamente porque es impactante. Las anécdotas humanizan un asunto, por lo cual los informativos normalmente empiezan (y a menudo acaban) con anécdotas. Pero las anécdotas son una base débil para tomar una decisión: con frecuencia dan lugar a equívocos, precisamente porque son impactantes. Pregunta siempre si una declaración está respaldada por datos, no sólo por una atrayente historia personal* [...]
> *Estamos tentados a añadir:*"Los datos rebaten a los autoproclamados expertos" [...] *Si realmente te interesa un asunto, trata de averiguar qué dicen los datos y si son realmente buenos. Muchos problemas permanecen sin resolver, pero muchos otros sólo están irresueltos en las mentes de aquellos a los que no les preocupa la evidencia.* [traducción propia]

Por otra parte, resulta de interés señalar un detalle que, aunque alusivo a un tema que tendremos en consideración más adelante, no quisiéramos dejar del todo de lado en relación con lo anterior. Tiene que ver con el siguiente extracto del artículo de Gould:

> *Yo poseía todas las características que conferían una probabilidad de vivir más tiempo: Era joven, mi enfermedad había sido reconocida en una fase relativamente temprana, y recibiría el mejor tratamiento médico de la nación*

Más adelante veremos que correlación no es equivalente, ni mucho menos, a causación: el análisis de regresión es, por propia naturaleza, un acercamiento operativo para establecer una primera forma de relacionar variables mediante parámetros, en orden a ratificar o descartar las hipótesis (de corto alcance, en general; aunque no en todos los casos) que dieron pie a un estudio (experimental u observacional) concreto[9]. O, al menos, así deberían ser tratadas por todas esas disciplinas que, más o menos aplicadas, más o menos experimentales, todavía se consideran a sí mismas científicas[10]. De manera que una correlación no pasa nunca de ser hipotética, y por tanto, está siempre, como si dijéramos, "bajo sospecha". Sin la base de un modelo teórico de mayor "altura" epistemológica, y mucho más comprehensivo que aquella que tratamos de confirmar (o lo contrario) con nuestro análisis, el efecto de una o varias de las llamadas "variables ocultas" podría estar desvirtuando por completo nuestros resultados, y dando pie al establecimiento de conclusiones falsas.

Por causa de todo lo anterior, nos atrevemos a aventurar una pequeña hipótesis sobre el caso Gould, que podría hacerse extensiva al problema general de la tasa de supervivencia de enfermedades graves, de largo y complicado tratamiento. No sabemos (o no se nos dice) qué clase de estadísticas consultó el señor Gould; no sabemos si estas incluían consideraciones particulares sobre la edad, el sexo, o el grupo humano tratadas *por separado*. De la lectura de la frase extractada parece desprenderse que no fue así: contando con ello, ¿Cuáles son las variables realmente implicadas por los datos que se consultaron? En otras palabras, ¿Cuál es la verdadera causa de que la vida de un paciente de cáncer se prolongue años y años o se consuma en pocos meses? Si volvemos al texto, no tenemos más que parafrasear a su autor: la edad, la fase más o menos temprana en que fue diagnosticada la enfermedad, y la calidad del tratamiento posterior. Supongamos, como hace Gould, que las diferencias en-

9 Con la notable excepción de las ciencias que se ocupan de procesos estocásticos, y donde la misma variabilidad (y su correlato matemático, la probabilidad) forma parte del núcleo axiomático de la teoría.

10 *E.g. vid.* M. Bunge, *La Investigación Científica: su Estrategia y su Filosofía,* (2ªed. corr.), Ariel, Barcelona, 1983

tre sexos y otros definitorios biológicos son irrelevantes al asunto: las tres categorías que menciona tienen que ver con una variable fundamental: el *status* socioeconómico, es decir, en último término, **los ingresos del paciente.**

Las dos últimas estan claras: se podría hacer alguna objeción a la primera, la edad. Pero sucede que la edad *a la que se es diagnosticado* es completamente dependiente de la *capacidad de gasto* en un país como los EEUU, (y máxime en 1982) donde la sanidad es casi completamente privada: donde hay que pagar por cada análisis, por cada revisión, por cada chequeo: donde la calidad de los seguros médicos es directamente proporcional a la calidad (en salario) del puesto de trabajo que se ocupa.

En conclusión: imaginemos que tomamos una gráfica como la de las actividades anteriores *(vid. p.52)*, y la colocamos al lado de alguna de las gráficas de las que probablemente está hablando Gould, debidamente escalada, tales como la fig.1 *(vid. p. 54)* ¿Sería descabellado decir que, en gran parte, *ambas están representando la misma relación?* ¿Es decir, que ambas nos indican que la esperanza de vida tras diagnóstico de cáncer es *directamente proporcional* —factor de escala mediante— a los ingresos del paciente?

ACTIVIDADES

1º) *Lee el texto que se adjunta.* [Aquí se incluiría un resumen traducido del texto de Gould[8]]

a) ¿Qué concepto estadístico es el eje principal del artículo?¿Qué tipo de medida es?¿Qué otras medidas del mismo tipo conoces?

b) Según lo que has aprendido, para caracterizar una distribución hacen falta, al menos, dos tipos de medidas: ¿Cuál es la que el artículo no menciona?¿Ese tipo de medida que no se menciona, no hubiera informado mejor al autor sobre sus verdaderas opciones?¿Cuál de estas medidas suele acompañar a la mediana en la descripción de una distribución?

c) ¿Qué opinión te merece el artículo? ¿Estás de acuerdo con él? En concreto, comenta la frase: "la variabilidad es la verdadera e irreductible esencia de la naturaleza. La variabilidad es la dura realidad, no un conjunto de medidas imperfectas de la tendencia central. Las medias y medianas son las abstracciones" *¿No crees que el autor habría estado mejor informado sobre la "variabilidad" si hubiera conocido la medida que se preguntaba en el apartado b)?*

a) Evidentemente, la mediana, una medida de localización (o centralización). La otra medida del mismo tipo que estudia en ESO es la media (aritmética); según los libros de texto que se manejen o el leal entender del profesor, también pueden enseñarse las medias geométrica y armónica.

b) El artículo no menciona ninguna medida de dispersión. Precisamente la "variabilidad" que tanto se menciona se trata de cuantificar por este tipo de medidas. La medida que suele acompañarlo es el Rango Intercuartílico (RIC), definido como el intervalo de valores cuyos extremos son el primer y el tercer cuartil, siendo el primer cuartil un valor tal que es mayor que los valores del 25% de un conjunto de datos ordenados de menor a mayor, y el tercer cuartil mayor que el 75% de esos mismos datso igualmente dispuestos.

c) Las "historias personales" no deberían suplir el análisis de datos científico, y esto incluye sus instrumentos matemáticos, como la mediana, que quizá sean mejorables, pero que resumen el conjunto disponible de datos infinitamente mejor que considerar *un sólo dato*. La media y la mediana sos abstracciones, sí, pero de la misma manera en que son abstracciones las entidades de las ciencias (duras), como "fuerza", "campo electromagnético", "momento angular", dependientes en su definición a su vez de abstracciones deliberadas: los términos matemáticos en los que están definidos.
Reconocer la variabilidad de la realidad es equivalente a reconocer la "verdad" del fenómeno físico; pero lo cierto es que no somos capaces de comprender, definir, tratar y utilizar los objetos fenoménicos, tal y como se presentan a nuestra percepción: sólo somos capaces de tratarlos de alguna manera cuando reducimos algunas de sus propiedades a abstracciones matemáticas manejables. Desechar estas "abstracciones" es desechar la ciencia en bloque: incluidos los tratamientos del cáncer, estudiados por medios estadísticos de forma preponderante. Por otra parte, si hubiera considerado el rango intercuartílico, por ejemplo, o hubiese consultado un diagrama de caja y bigotes, estas abstracciones le habria informado bastante mejor de sus opciones que utilizando solo la mediana.

Media, mediana y moda

No obstante lo visto anteriormente, existe otro tipo de situaciones en las cuales ni la media ni la mediana de los datos brutos sirven como indicadores de ninguna tendencia central; estamos pensando en las poblaciones (o muestras) con distribuciones bimodales o multimodales. La moda, definida como el valor x_{mo} de la variable aleatoria X que maximiza la función de (masa de) probabilidad si X es discreta, o que hace máximo el valor de la función de densidad de probabilidad para X continua, pasa a asumir dos valores, lo cual provoca que tanto la media como la mediana asuman valores que no son en modo algunos los de mayor frecuencia relativa (no son los más *típicos*).

Estas distribuciones suelen ser bimodales (dos modas) cuando la distribución estudiada es resultante de la agregación de datos procedentes de dos distribuciones unimodales[11], cuyos ejemplos en biología son numerosos, debido a diversas clases de dimorfismo **[8]** **[34]**, sobre todo de carácter sexual, que se presentan en muchas especies. Como se da el caso de que el *homo sapiens* también presenta esta característica, ello nos permite la realización de la siguiente actividad.

11 Aunque no siempre: El famoso ejemplo del geyser *Old Faithful*, en el parque de Yellowstone, en Wyoming, USA, ofrece una distribución de frecuencias relativas de tiempos de espera entre erupciones que sigue una distribución bimodal al parecer intrínseca. Vid. A. Azzalini y A. W. Bowman, *A Look at Some Data on the Old Faithful Geyser* [artículo en línea]1990 (pub. en línea por *www.jstor.org*) [cons. Junio 2015] Applied Statistics, Vol. 39, No. 3, pp. 357-365. DOI: 10.2307/2347385 Disponible en [previo registro]: http://www.jstor.org/stable/2347385?origin=JSTOR-pdf

ACTIVIDADES

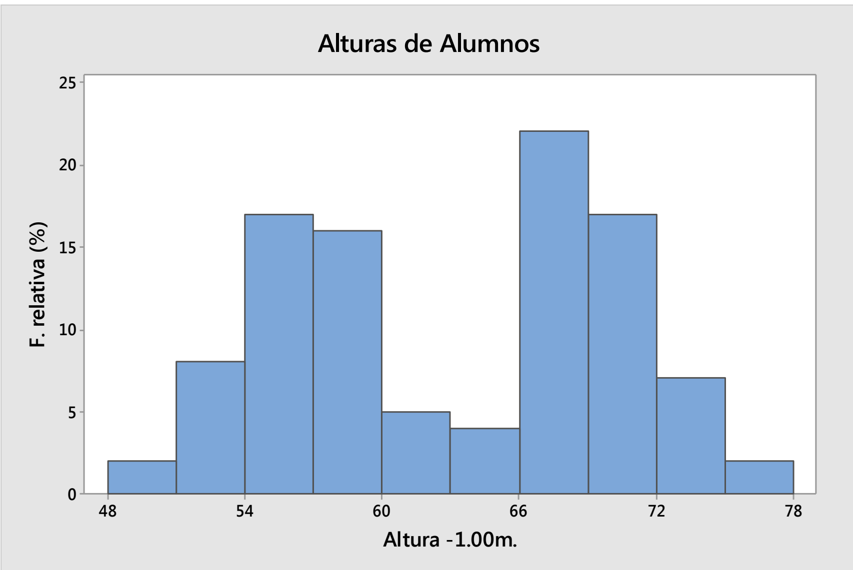
1º) Divididos en grupos de trabajo, realizad una encuesta para averiguar la altura (metro-centímetros) de todos los alumnos de vuestro mismo curso que haya en el centro. Recopilad los datos tanto numéricos como de identificación (nombre y apellidos de cada alumno) y repartidlos entre todos los alumnos de la clase. Realizad individualmente un histograma de frecuencias relativas o porcentajes. ¿Qúe amplitud de intervalos de valores de la altura habéis escogido, y en función de qué concepto?

La amplitud del intervalo se escoge normalmente en relación con el número de elementos de la población (o muestra) y del rango o recorrido de los datos (la diferencia entre el mayor y el menor de los valores observados) Existen algunas reglas empíricas, tendentes a que el histograma sea efectivamente ilustrativo, para hallar el número de intervalos en que debe dividirse el rango ; puede escogerse un redondeo del valor $\sqrt{n}$ con n el número de realizaciones de la variables (en este caso, el número de alumnos total), o bien la siguiente tabla ([**34**], p. 25)

Tamaño de muestra	25	50	100	200	500
Número de clases	6	7	8	9	10

De modo que si el rango es (cm.) 178-148= 30, para 100 alumnos (30/8= 3,75) tendríamos 8 intervalos de 3,75 cm., y según la regla de $\sqrt{n}$ tendríamos 10 intervalos de 3 cm. Escojamos esta última y pongamos por caso que resulta un histograma de la siguiente manera:

Fig. 2: Histograma de frec. rel. de todo un curso de ESO [elab. propia]



2º) Observa atentamente el histograma resultado de la encuesta:

a) ¿Encuentras algo extraña la gráfica, a la luz de lo que sabes sobre distribuciones normales? ¿Cómo la describirías?

b) Halla la media y la mediana de la distribución de frecuencias (relativas). Son sus valores típicos*? Es decir, si quisieras saber qué altura tiene un alumno cualquiera de tu curso, ¿Te serían de utilidad?*

c) Además de la media y la mediana, también has estudiado (o deberías hacerlo) otra medida de centralización distribuciones. ¿Cuál es? ¿Y cómo se refleja en el histograma que estamos estudiando?

a) Desde luego, no se parece a las gráficas de distibuciones normales que se han visto hasta el momento en clase: donde debería estar el máximo, o el “pico” central, ahora hay un “valle”. Y tampoco parece una distribución sesgada a la izquierda o a la derecha; ni tan siquiera se podría decir que *está a la vez* sesgada a la izquierda y a la derecha, porque entonces los valores en esos lados de la gráfica serían extremos. Lo que más bien parece e una distribución normal, pero con dos “picos” en vez de uno.

b) La media es $\overline{X}$= 63,053+100= 163,05 cm., y la mediana X_{me}= 64+100= 164 cm.[12] Pero ciertamente no me serían de utilidad, ya que ambos valores caen en los intervalos o clases del centro que forman el valle, de manera que sólo corresponden, respectivamente, al 5% y al 4% de la población total.

c) Se supone que también se ha estudiado la moda, y la manera de hallarla. Dado que la definición de moda incluye los valores localmente máximos, se entiende que la distribución tiene dos modas, X_{mo1}= 155 cm. y X_{mo2}= 168 cm., valores los cuales están incluidos en los intervalos con frecuencia más alta del anterior histograma

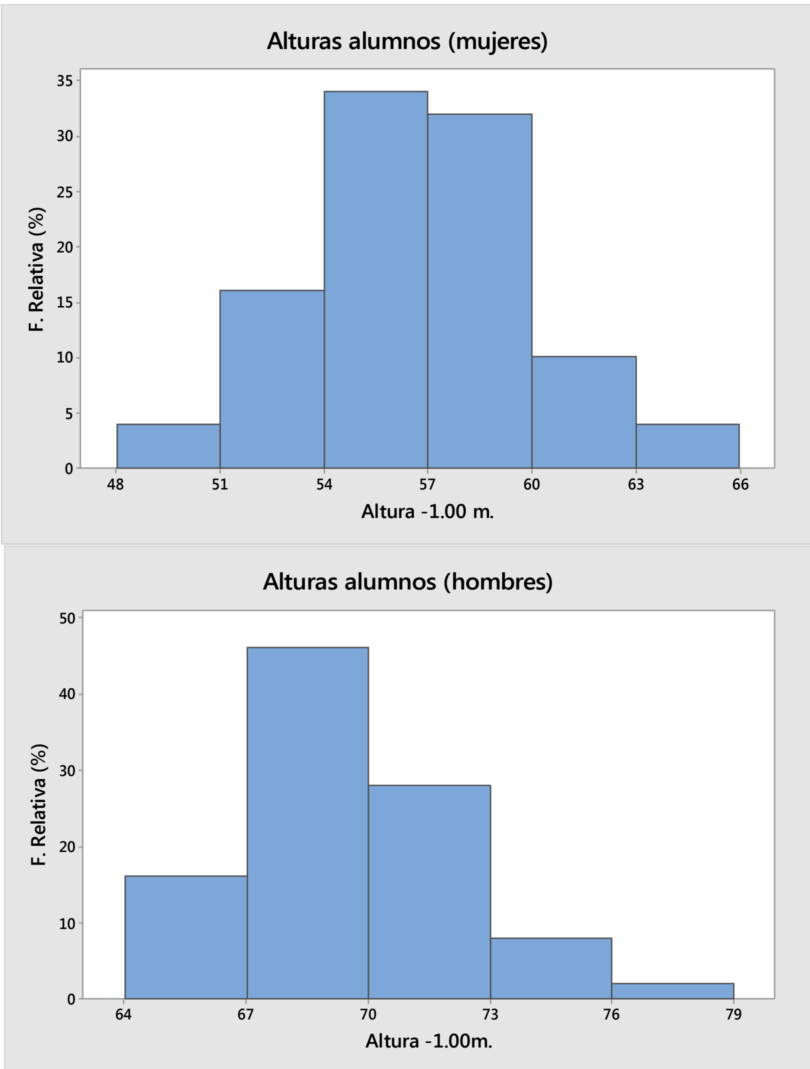
Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres
159	167	159	169	154	167
162	173	164	168	151	172
154	167	152	172	158	166
159	168	158	168	155	168
158	169	157	169	157	174
160	171	151	167	153	169
159	166	156	169	156	166
155	175	162	172	157	167
155	168	155	171	155	164
155	168	160	170	153	176
163	169	151	168	149	170
154	171	157	170	155	172
149	169	155	166	158	167
153	167	154	170	160	166
154	171	154	168		
157	170	157	164		
155	168	157	171		

12 Estos valores se han hallado con datos generados aleatoriamente por mi parte, y que no expongo en detalle (aunque se pueden deducir de la fig. 3) por no constituir el punto principal del ejercicio; se supone que el alumno cuenta con datos verdaderos a partir de los cuales ha generado el histograma

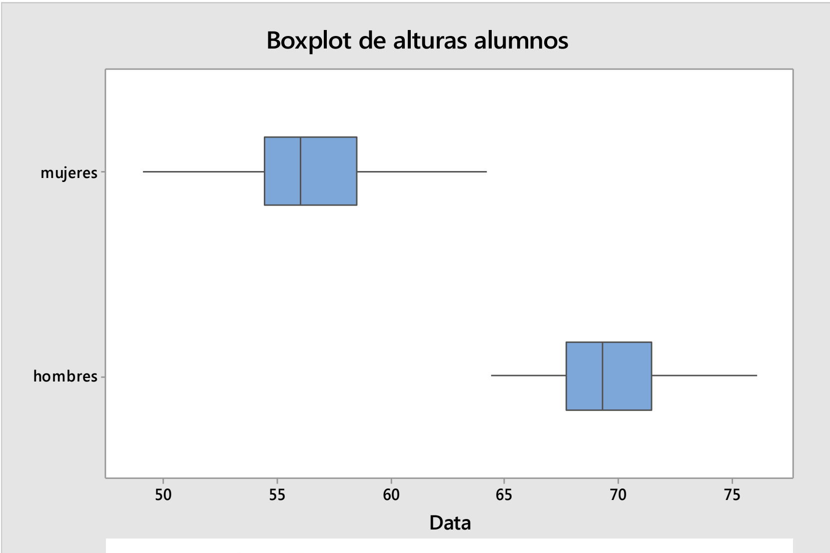
3°) ¿Cuál crees que pueda ser la causa de que esta distribución adopte un forma con dos “picos”? Dado que los datos que habéis recolectado tienen asociado el nombre de cada alumno, segregad los datos por sexos y haced un histograma por cada distribución. ¿De qué medida de centralización de la distribución original están más cerca la media de cada una de las distribuciones actualmente confeccionadas?

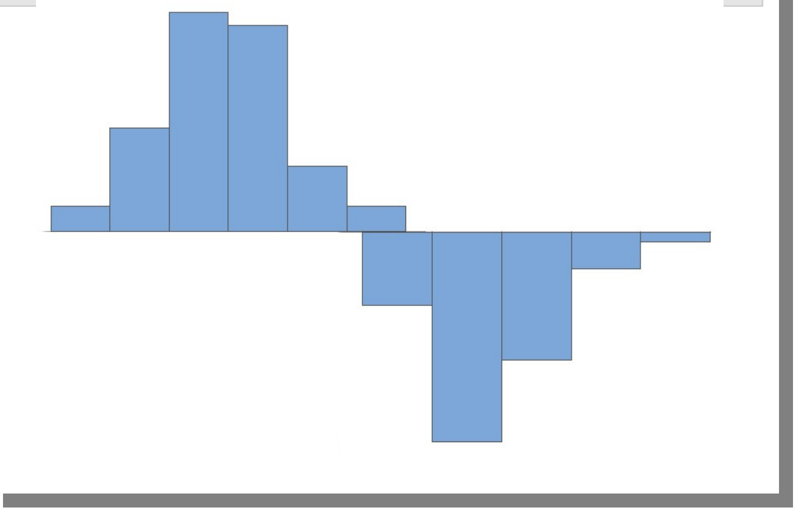
En este momento se explicaría al alumno que la distribución de los alumnos no segregada por sexos constituiría una mezcla de *datos no homogéneos*, y por lo tanto sus medidas de resumen no dan información relevante, ni sobre la población total ni sobre cada una de las subpoblaciones (hombres y mujeres). De manera que siempre es preciso analizar los datos disponibles e hipotetizar las razones por las cuales fenómenos naturales que sabemos que muestran variaciones que siguen una distribución normal no parecen presentar tal distribución.

Fig. 4 y 5 : Hist. f. rel. de todo un curso de ESO. Arriba Mujeres: Abajo, Hombres



Las medias de las distribuciones reflejadas más arriba son, respectivamente, $\overline{X}_m$= 56,51 + 100= 156,51 cm. $\overline{X}_h$= 69,60 + 100= 169,60 cm.; valores mucho más cercanos a las modas halladas anteriormente que a la media y mediana de la distribución original.





4º)Con el mobiliario del aula retirado, formad grupos por sexos de forma que, a partir de 145, estéis en un intervalo de 5cm. (es decir, si mides 165 cm. y eres hombre, vas en el grupo de hombres cuya altura es de 165 a 169 cm). Una vez formados los grupos, colocadlos ordenados de mayor a menor clase en dirección al eje más largo del aula. Finalamente, dentro de cada grupo y siguiendo el eje perpendicular al anterior, colocáos por orden ascendente de manera que el de menor altura dentro de cada grupo quede más cerca del eje principal (el que está formado por la sucesión de grupos). ¿Qué gráfico estadístico hemos formado? Realizad los diagramas de caja de todas las agrupacioones de datos ¿Encuentras alguna diferencia esencial entre los diagramas de datos segregados y la de los conjuntos?

Evidentemente, dos histogramas:

Fig. 6: Histograma f. rel. alturas de Clase de ESO. Mujeres

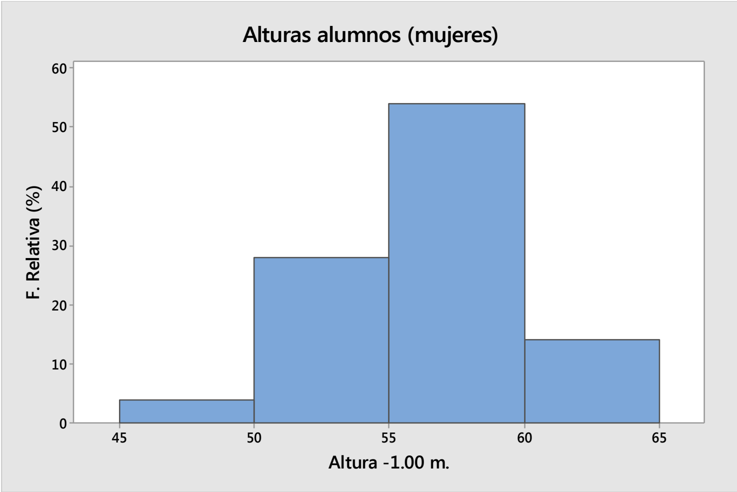
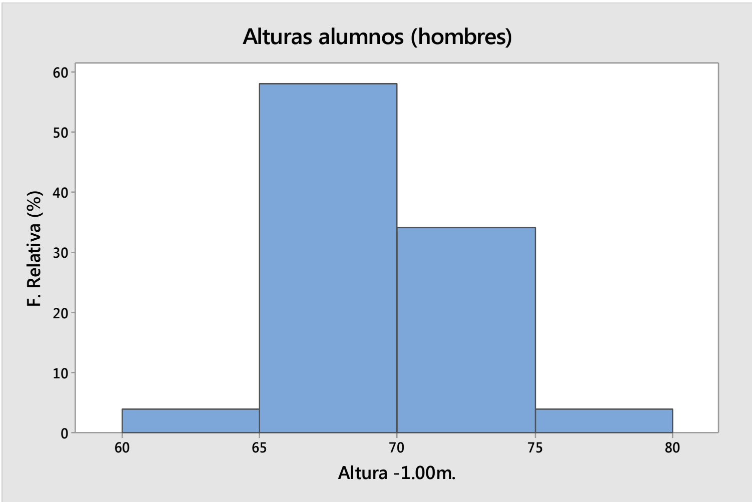
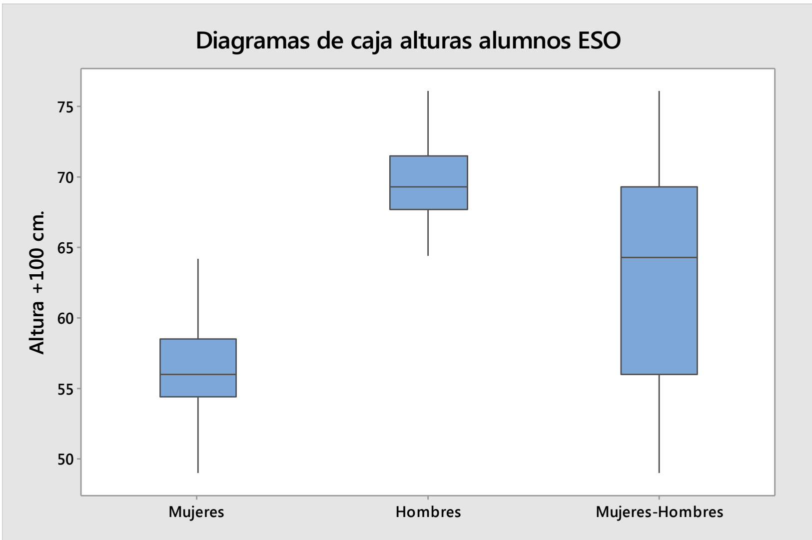


Fig. 7: Histograma f. rel. alturas de Clase de ESO. Hombres



Los diagramas de caja nos informan, además, de que las desviaciones típicas de los datos desagregados son mucho menores que la de los conjuntos.



De hijos y percentiles

No hay que olvidar que la mediana, el primer y el tercer cuartiles son respectivamente también los percentiles 50, 25 y 75, respectivamente. Además, el vocablo “percentil” es casi de uso común. La situación más cotidiana en la que un ciudadano normal se encuentra cuando oye hablar de percentiles es la consulta del médico: muy específicamente, la del pediatra. Efectivamente, en cada fase del programa cronológico de visitas a tal especialista que los padres deben observar, el médico, tras pesar y medir al vástago, suele agregar que en altura se encuentra entre tal y cual percentil (por ejemplo, entre el 50 y el 85; aunque puede que se decida a ser más preciso). Generalmente el médico explica amablemente a la pareja, en la medida en que supone que va a ser entendido, lo que esta medida significa. Que no tiene mucho éxito en la tarea lo demuestra la siguiente ilustración “motivacional”[13]:

www.mipediatraonline.com

http://www.mipediatraonline.com/wp-content/uploads/-1/1//images_peketippercentiles.png

“La mitad de los niños sanos está por debajo de la media” sólo si la media y la mediana coinciden: lo cual podría ser el caso (aproximado), pero demuestra que el artista que pergeñó la anterior postal no ha entendido bien el concepto. Siendo un tema que sirve muy bien para fijar conceptos, en los que la comprobación en la cotidianeidad del alumno de los objetos matemáticos implicados es muy inmediata, plantearemos las siguientes actividades:

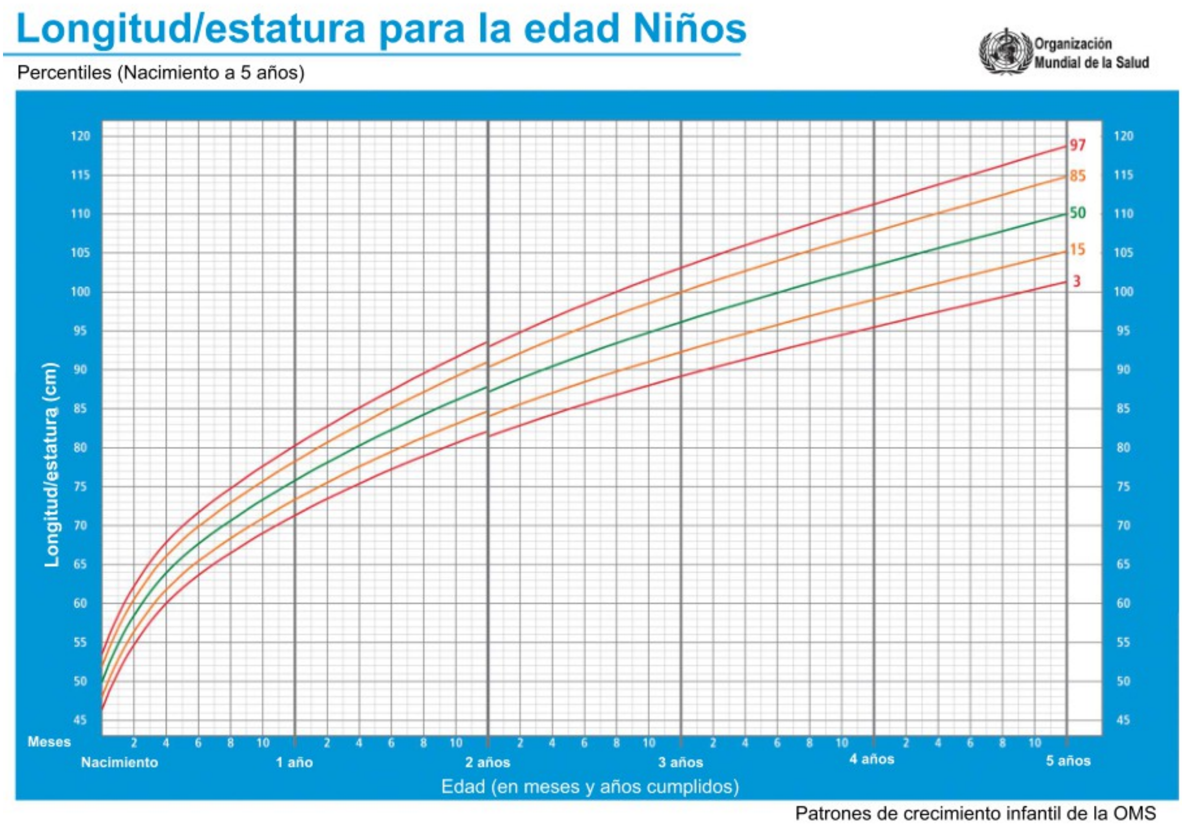
13 J. Garrido García, *Tablas de percentiles de la OMS* [artículo en línea], 3 Enero 2010 [consultado Junio 2015] en Mi Pediatra Online, (www.mipediatraonline.com) Disponible en: http://www.mipediatraonline.com/wp-content/uploads/-1/1//images_peketippercentiles.png

ACTIVIDADES

1º) Observa la siguiente ilustración [vid. supra]. *¿Qué conceptos de estadística se mencionan? ¿Se hace un uso de correcto de estos conceptos? ¿Cómo redactarías de forma correcta la frase o frases que hayas creido que no están bien expresadas?*

Se mencionan los percentiles y la media: medidas que, por cierto, no son homogéneas respecto de su objeto. No se hace un uso estricto del concepto de media, cuando se dice que "la mitad de los niños sanos están por debajo de la media". La medida que cumple la propiedad que se expresa es la mediana, no la media. Luego, la frase correctamente expresada debería rezar como sigue: "La mitad de los niños sanos están por debajo de la mediana".

2º) Observa la gráfica de percentiles de longitud/estatura de niños (varones) hasta los cinco años de la Organización Mundial de la Salud[14].



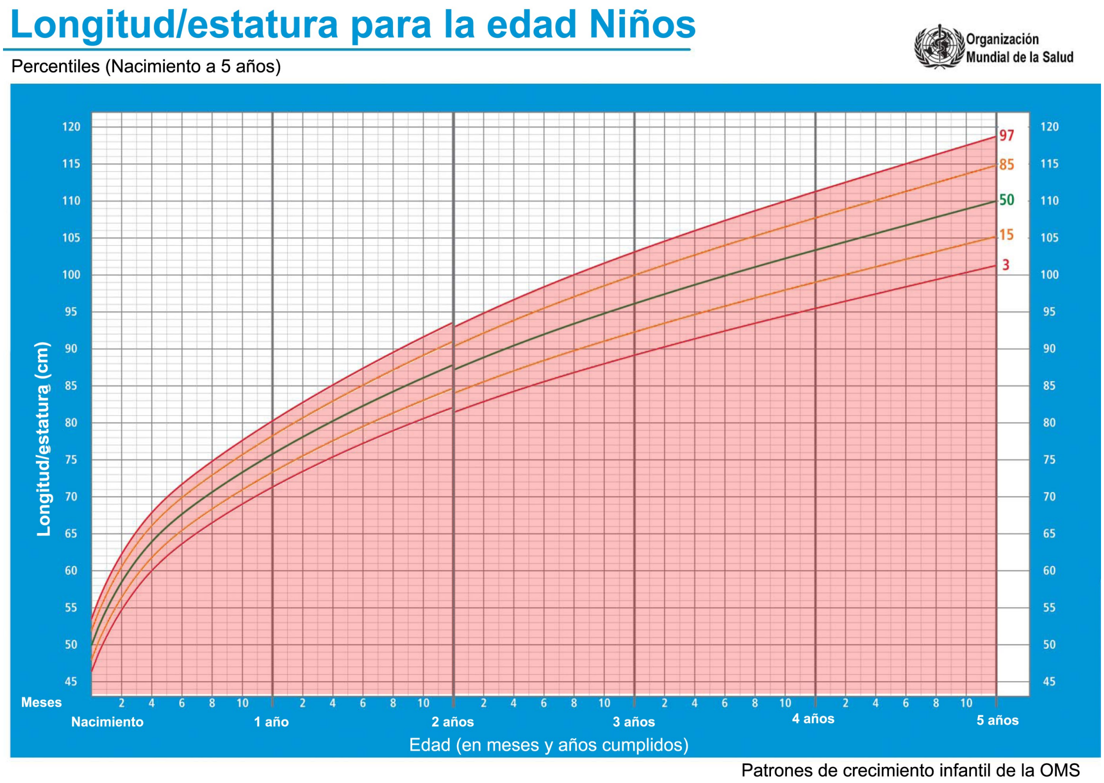
http://www.who.int/entity/childgrowth/standards/cht_lhfa_ninos_p_0_5.pdf?ua=1 www.who.int

a)Como puedes observar, las abscisas miden la edad y las ordenadas la longitud/estatura. Cada curva representa una función distinta, identificada por un número. ¿Qué categoría crees que identifica este número?¿Por qué otro nombre se conoce la curva central, identificada con el número "50"?

14 El siguiente y los demás gráficos relacionados están tomados de Organización Mundial de la Salud. Departamento de Nutrición, *Patrones de crecimiento infantil* [recurso en línea] 2015, [consultado Junio 2015], en http://www.who.int , OMS. Disponibles en: http://www.who.int/childgrowth/standards/es/

Una parte bastante olvidada en relación con las competencias básicas a desarrollar en el alumno de ESO tiene que ver con la capacidad, no tanto de deducir o de razonar, sino de conjeturar y decidir, de una manera controlada, consistente y plausible. Generalmente, y sobre todo en matemáticas, las cuestiones que se proponen al alumno facilitan información completa, autocontenida en el enunciado: y este no es el caso, ni de lejos, en la cotidianeidad a la que el alumno se enfrenta y respecto de la cual se supone que se han de desarrollar las competencias. Los responsables de PISA sí que se han apercibido del asunto, y las pruebas que realizan cuentan con la capacidad del alumno para, entre otras cosas, buscar información de forma activa y valorarla, lo cual incluye a menudo plantear hipótesis sobre la misma información (aunque sean de carácter muy básico), y valorarlas a fin de tomar un curso de acción. En este sentido, el planteamiento de la actividad no incluye toda la información; en realidad, la gráfica sin otro contexto que sí misma no indica que las curvas corresponden a percentiles. El alumno, con la guía del profesor, debería ser capaz de emitir hipótesis sobre la naturaleza de la información que se le suministra.

En cierto modo, el alumno cuenta con más información contextual que la simple gráfica; al fin y al cabo, es de suponer que esta actividad se proponga al mismo tiempo que el temario correspondiente a estadística: luego las curvas, *probablemente* —primera hipótesis— reflejen un concepto relacionado. Los valores de las gráficas toman valores entre 97 y 3; los cuales están muy cerca de 100 y 0: por lo tanto, es de suponer —segunda hipótesis— que tengan algo que ver con *porcentajes*. Ahora bien, porcentajes, ¿De qué cosa*?* La gráfica relaciona estaturas con edades: las estatura y edades son atributos de algo: ¿De qué? Si acudimos al título, observamos que se refieren a niños de cero hasta cinco años; luego las cifras corresponden al porcentaje de niños que, respecto del total, tiene alguna carácterística, o algún conjunto de características. ¿Cuáles? Si escogemos una de las funciones cualquiera,por ejemplo la del 97, vemos que recorre valores que muestran dos características: la estatura y la edad. Luego la gráfica en cuestión puede referirse a que el 97% de los niños muestran esas dos características —tercera hipótesis— , es decir, que el 97% de los niños miden tantos centímetros cuando tiene cierto número de meses de edad. Pero llegados a este punto hay algo en nuestra hipótesis que falla, supuesto que a cada porcentaje le corrresponde una sola combinación de valores: sencillamente, porque 97+85+50+15+3= 250≠ 100. Supuesto que cada gráfica es una proporción del total, la suma de todas las proporciones debe dar ese total, y no 2,5 veces un total. ¿Cómo se resuelve esta inconsistencia? Es entonces cuando el alumno puede recordar parte del temario, cuando se estudiaban las tablas y gráficas de frecuencias relativas, y de frecuencias relativas acumuladas. La única solución plausible es que el valor de 97 incluya asimismo los valores de las demás curvas. Por lo tanto, la relación entre la curva y los valores que refleja no es de uno a uno, sino que al 97% le corresponden un conjunto de valores ¿Qué valores? Si recordamos las gráficas de frecuencia acumulada, al 97% le corresponderían tanto los determinados por la curva como todos los menores a ese valor, para una abscisa dada. Estrictamente hablando, la curva 97 es el límite superior del área cuyos puntos, definidos por la combinación de dos valores (estatura y edad), están dentro del porcentaje definido por esa cifra, lo que se puede representar de la manera siguiente:



Lo cual es aplicable a las demás curvas con los valores 85, 50, etc. ¿Cómo se interpreta esto? Cojamos, por ejemplo, la curva marcada por el 3%. Su determinación implica que a un 3% del total de los niños (representados por un punto en la gráfica, es decir, por un par de valores concretos de edad y altura) les corresponde un punto que está <u>por debajo</u> de esa curva. Asimismo, otro 3% de los niños estarán <u>por encima</u> de la curva del 97% ¿Por qué? Porque por debajo de esa curva están el 97% de los niños, ergo, hay un 3% no considerado que está por encima.

Todo esto puede parecerle obvio o inmediato al lector. Sin embargo, como ya dijimos, la operación de hipotetizar sobre una información incompleta, relacionar la información que se posee y coordinarla hasta hallar el significado de lo que se pregunta es un ejercicio de construcción cognitiva que a menudo se da por supuesto en el alumno y que no tiene por qué tener lugar, si solo se le presentan problemas autocontenidos y definiciones precisas. No solo porque en el ámbito de la realidad rara vez va a encontrar problemas formulados de esta manera, sino porque, cuando los elementos matemáticos se presentan como completamente definidos de principio, sin ningún nexo con situaciones reales, no contribuyen a generar competencias, si como tales entendemos un conjunto de estrategias y herramientas destinados a resolver problemas que conciernan directamente al alumno, y al adulto en que se convertirá.

Por eso mismo, si consideramos el contexto de los datos las explicaciones se vuelven más intuitivas. Un valor alto de ordenada significa un niño más alto, para una edad determinada; un valor pequeño, un niño bajo para su edad. Si se observa la realidad circundante, nos daremos cuenta de que, cuanto más extremos son los valores de estatura de un individuo, menor es su número. Los individuos especialmente bajos o especialmente altos son escasos; la mayor parte de la población posee una altura que se halla a medio camino entre las estaturas del inviduo más bajo y el más alto de una población. Este puede ser el valor medio, *si hubiera el mismo número de individuos altos o bajos respecto a ese valor medio.* Como todavía no hemos establecido este hecho, podemos, sin embargo, realizar la operación siguiente: cojamos los datos de estatura de todos los individuos y ordenémolos correlativamente en una fila: entonces, habrá un valor determinado de altura para el cual, si dividiésemos la fila en dos subfilas dsitintas, el 50% de los individuos estaría en la subfila cuyos miembors muestran estaturas mayores que ese valor, y otro 50% que estaría en la subfila cuyos miembros miden menos que ese valor. Ese valor conseguido mediante tal operación no sería la media, que se calcula sumando las alturas de todos los miembros de la fila y dividiéndo esta suma por en número de miembros, sino la mediana, que tiene en cuenta el orden de los miembros de la población respecto de una determianda propiedad cuantitativa. Si antes de realizar esta operación hemos dividido a los individuos por su edad, cada par de subfilas subsiguiente marcaría un valor de la mediana distinto: lo que podría representarse en una gráfica con los valores de la mediana (medidos en unidades de longitud/estatura) como ejes ordenados y los de edad (medidos en unidades de tiempo) como eje de abscisas: en ese caso, al curva resultante que obtendríamos sría la marcada en la gráfica con el número 50. Análogas operaciones, con los porcentajes dela población respectivamente considerados, nos darían las demás curvas de 15, 85, etc. En términos prácticos para unos progenitores, podríamos decirles que si su vástago está en la curva del 97 está en el percentil 97 (ya que, como hemos visto, la mediana es el percentil 50). Y además, significaría que el 97% de los niños (sobre cuyos datos se hayaconfeccionado la gráfica: la que contemplamos parece ser internacional, no sólo sobre España) miden menos que el suyo para *la edad* que tiene en ese momento.

b) Supuesto que hayas identificado el concepto estadístico al que se refería la pregunta anterior, señala un intervalo de valores de ese concepto que le corresponda a:
1º.-Un niño de 2 años que mide 85 cm.
2º.- Ídem, *3 años y 90 cm.*
3º.- Ídem, *4 años y 110 cm.*
Además de observarse, como parece lógico, que los valores de estatura crecen a lo largo de la gráfica, ¿Qué otro concepto estadístico observas que aumente con el número de años?

Basta con observar la gráfica y concluir que cada niño está entre un par de percentiles:

1º- Entre el percentil 50 y el 15, casi en el 15: es decir, que más del15% de los niños de 2 años de edad miden menos de 85 cm. (si obviamos la discontinuidad de la gráfica, sin duda un error de representación)
2º- Entre el percentil 15 y el 3; algo más del 3% de los niños de 3 años son más bajos que este niño.
3º- Entre el 97 y el 85; al menos el 15% (85-15) de los niños de 4 años son más altos que este.

Se puede observar que la distancia vertical entre estaturas correspondientes a cada percentil aumenta en la dirección positiva del eje de abscisas (de izquierda a derecha): lo cual quiere decir que no sólo la media/mediana, sino también la variabilidad (o dispersión) de las estaturas crece a medida que los niños tienen más edad. Desde el punto de vista factual es muy claro: evidentemente todos los niños nacen con unas dimensiones bastante parecidas, dependientes de la anatomía de la madre, y las diferencias entre ellos van aumentando hasta llegar a la edad adulta, donde se da la máxima dispersión de estaturas entre individuos.

c)¿Podrías decir, con la sola información que suministra la gráfica, si la media para cada edad está muy cerca de la mediana? Explica tu respuesta.

Sin saber nada más, podemos decir que la media debe estar muy cerca de la mediana, o ser coincidentes en la práctica, ya que si este no fuera el caso, los puntos de la gráfica correspondientes a curvas de percentiles simétricos respecto de la mediana mostrarían una diferencia de valores de estatura (de ordenada) distintos: si, pongamos por caso, la distribución estuviera sesgada hacia la derecha, entonces el intervalo de estaturas definido entre el percentil 50 yel 85 sería mayor que el correspondiente intervalo entre el 50 y el 15; exactamente lo mismo podría decirse del intervalo entre 50 y 97 y entre 50 y 3. En cambio, se comprueba que la amplitud de esos intervalos es prácticamente la misma, para todas las edades consideradas. Por lo tanto, la distribución es casi o completamente simétrica: en cuyo caso, media y mediana adquieren valores muy similares (para cada edad), casi coincidentes.

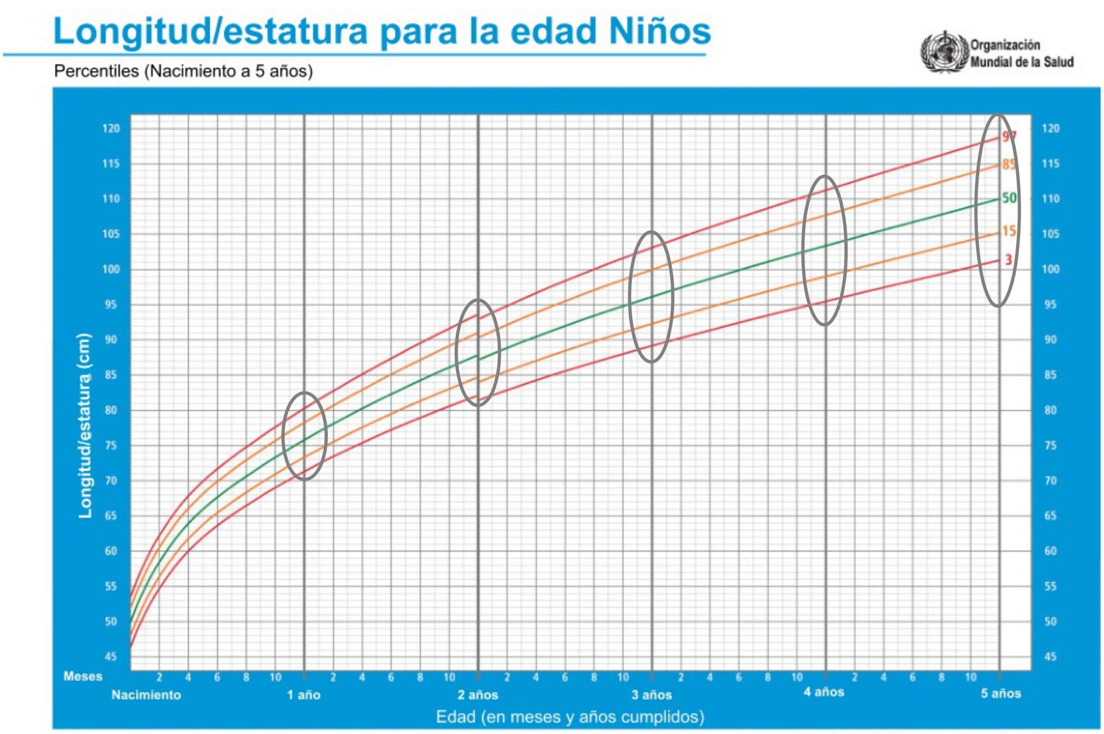
d) Razonando no solo desde las matemáticas, ¿Te parece que esta gráfica ***representa*** *distribuciones muestrales o poblacionales? Dicho en otros términos, ¿Es fruto de una operación de rocgida de datos o de una formalización posterior de los datos? Explica tu respuesta.*

Es este un punto en el que conviene insistir, no solo porque provoca muchos equívocos, sino porque una respuesta correcta por parte del alumno demostraria que se han entendido los conceptos, no solo matemáticos, sino estadísticos, en el aspecto de la toma de datos y de su representación y formalización. Si bien un alumno avispado podría intuir que, dada la procedencia de la gráfica (la OMS), y que su publicación tiene como propósito declarado dar una información general sobre el crecimiento infantil, una información para el uso de cualquiera que la consulte, la gráfica no puede ser una simple distribución muestral producto de uno o varios experimentos (o de una o varias tomas de datos)a fin de contrastar alguna hipótesis científica: al contrario, debe ser la formalización final de una hipótesis, esto es un modelo construido de la realidad general considerada. Quizá no sabría expresarlo en estos términos, pero sería interesante conocer si algún alumno ha llegado por esta vía heurística a la solución.

Pero supongamos que no sabemos ni la procedencia ni el propósito de la gráfica. ¿Podemos todavía deducir, aunque sea indirectamente, que no corresponde a la transcripción gráfica de los datos de un experimento (o de un conjunto de ellos)? Sí, si consideramos la variabilidad intrínseca de los datos: estos nunca nos daría un resultado tan pulcro y regular, tan continuo y suave como el reflejado en la gráfica. La gráfica es, por tanto, una simplificicación de los datos, una modelización, esta vez sí, matemática.

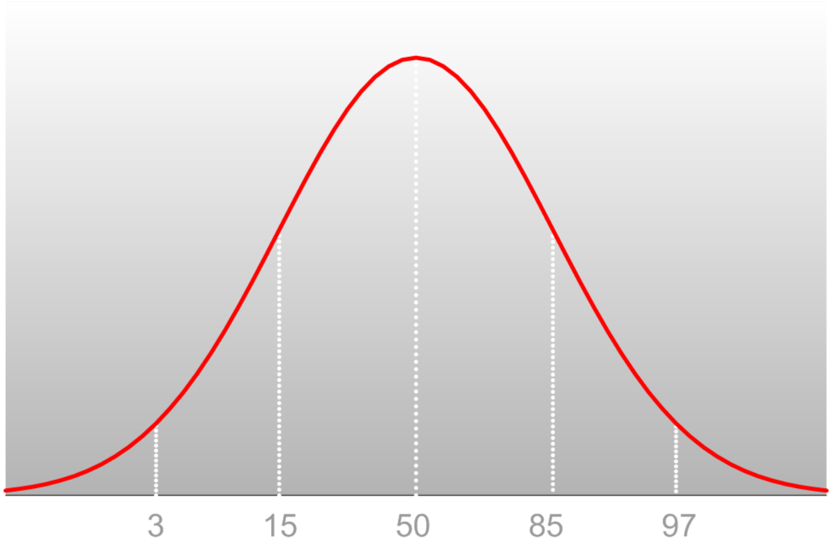
Ahora bien, esta modelización no solo se adopta por razones de sencillez o de estética. Esta una modelización basada en las leyes generales de la estadística, en concreto en la ley de los grandes números, una de cuyas consecuencias determina que en una serie de ensayos o casos, cuando este número es suficientemente grande, la media de sus resultados se acerca cada vez más al valor medio real de la entidad medida. (una demostración de tal ley estaría fuera de lugar en el contexto de enseñanza secundaria; habría que exponerlo por analogía con casos sencillos). De ahí que agregando las estimaciones de una serie suficientemente grande de mediciones se compruebe que nos acercamos cada vez más a un valor ideal, deducido estadísticamente, que corresponde a la media poblacional. Un valor que, por razones prácticas, nunca nos es enteramente conocido, pero al cual podemos aproximarnos con tanta precisión como deseemos. En cuanto hayamos una medida con la precisión que nosparece adecudada, la utilizamos para construir un modelo de la realidad (en este caso, una gráfica que representa un modelo de crecimiento infantil entre los 0 y los 5 años), que no refleja exactamente la realidad misma, sino que no es útil porque: es tratable con una sencillez relativa, descriptiva y matemáticamente; y porque es de carácter general: no depende de las circunstantncias de tal o cual experimento, o serie de experimentos.

e)Por la teoría sabes que un gran número de fenómenos naturales, entre ellos la variabilidad en estatura de la especie humana, responde bastante bien a una aproximación mediante la curva de distribución normal o gaussiana. ¿Cómo se refleja esto en la gráfica? ¿Cómo podríamos saber, deduciéndolo de la misma gráfica, que se ha aplicado el modelo de distribución normal, y no una distribución empírica, es decir, una generalización de la información de los datos que "se parece" a la distribución normal?



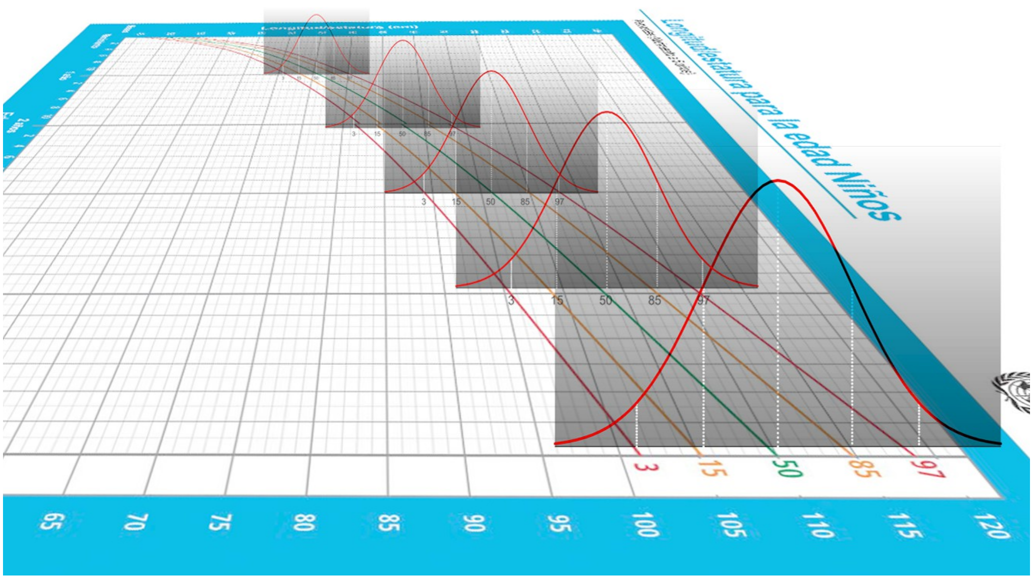
En todos los cortes transversales de las curvas (hemos señalado aquellos que corresponden a edades enteras, pero serviría cualquier corte vertical) se reflejan los percentiles correspondientes. Si la distribución del modelo es normal, cada corte vertical correspondería a esta gráfica:

Fig.8: Distribución normal. Percentiles 3, 15, 50, 85, 97



Donde hay que interpretar la cifra de cada percentil como el área comprendida entre la curva de densidad y el eje horizontal desde $-\infty$ hasta la vertical definida por cada percentil. Si el área total comprendida entre la curva y eje de abscisas es 1 (el 100%, la probabilidad total) el área entre la curva y el je de abscisas hasta el percentil 97 será el 97% de la total; hasta el percentil 85, el 85%, etc. Más abajo se refleja en 3 dimensiones tal y como se vería si representásemos la función de densidad de probabilidad en un hipotético eje z perpendicular al plano de abscisas y ordenadas actual. (El valor en altura, z, no es ni edad ni estatura, sino frecuencia relativa —probabilidad— de niños de esa edad y estatura).

Fig.9 Funciones de densidad de probabilidad para edades enteras



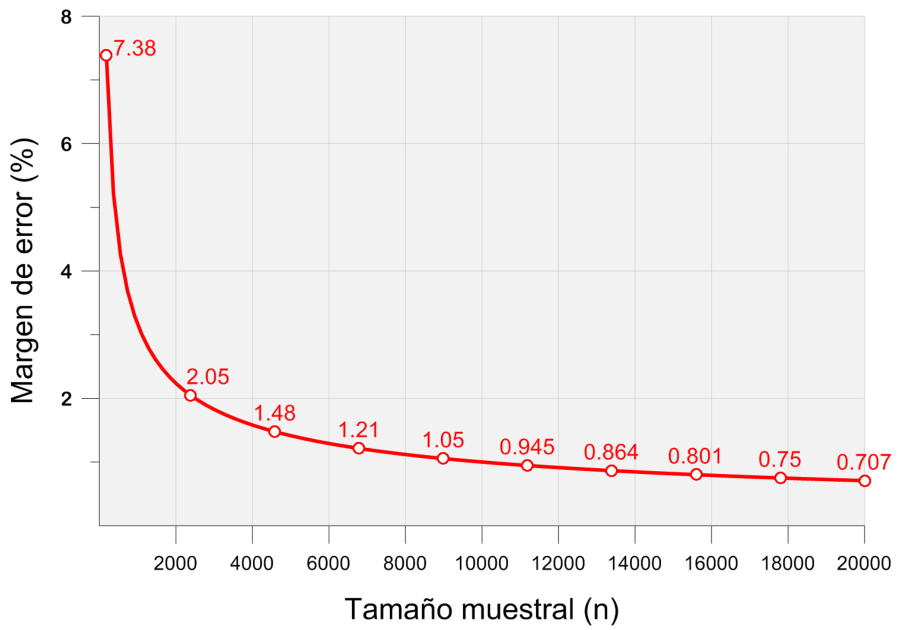
Hay un detalle que nos indica en los percentiles que la distribución es teórica, esto es, matemática: el hecho de escogerse los percentiles 97 y 3 como extremos. Es de suponer que un conjunto de datos, por amplio que sea, aunque abarcara toda la población, tendría un extremo inferior y superior, un rango definido. No es de recibo pensar que hay niños que a los dos años miden 2 centímetros, o dos metros de alto. Por lo tanto, considerando datos, podrían incluirse los extremos abslutos, como percentiles 0 y 100 respectivamente. Pero la distribución normal, por las propiedades analíticas de la función de densidad que le es propia, tiene el eje de las abscisas —el eje donde la probabilidad es 0— como asíntota: es decir, que la función tiende a 0 por sus dos extremos, sin llegar nunca a valer exactamente 0. De ahí que se hayan escogido arbitrariamente los valores de 3 y 97 como extremos de esta representación gráfica, lo mismo que se podrían haber escogido 2 y 98, o 1 y 99; pero no 100 y 0, porque estos se hallan en $+\infty$ y $-\infty$ respectivamente; cosa que no podría suceder en una distribución empírica,que reflejase frecuencias relativas reales.

Se puede hacer notar al alumno, como hecho curioso que, si bien puede fijarse el domino de la función de densidad normal como el de los números reales positivos, dado el absurdo de que alguien tenga una estatura negativa, es mucho más dificultoso encontrar un criterio para dar un valor convencional de las x a partir del cual definir como 0 la función. De tal forma que la probabilidad, según el *modelo matemático* usado, de que un niño mida dos metros, dos kilómetros o dos años luz de estatura, es infintesimal, pero no igual a 0. Sucede lo mismo con los modelos de esperanza de vida, asimismo matemáticos: la probabilidad de alcanzar los 1000 años de edad según estos últimos tampoco es 0, aunque su valor sea del orden de magnitud del radio del electrón, o aun menores. La dificultad es, no obstante, meramente lógica, y afecta, por tanto al objeto matemático constituyente del modelo, no al hecho real: la longitud es un magnitud definida positiva (ya que no tenemos noción de que tenga lugar algún objeto de longitud negativa); en cambio, no hay imposibilidad lógica en que un objeto cualquiera tenga una longitud mayor que cualquier valor real que postulemos, y esto incluye al ser humano también.

El tamaño (de la muestra) importa

Se recordará que en el capítulo dedicado a *Problemas de muestreo*, al hablar de las *encuestas de participación voluntaria (vid. p. 39)*, ya indicamos que el tamaño muestral era esencial (aunque no suficiente) a la hora de estimar la proporción de individuos en los que se da cierta característica dentro de una población dada. Pero ¿En qué sentido hace buena la estimación? Como se puede comprobar en el ejemplo mencionado, en el sentido de *dar precisión* a la misma, o, lo que es lo mismo, en el de disminuir el *margen de error*, que cuantificábamos aproximadamente como $r = 1/\sqrt{n}$ [1]. En el ejercicio mencionado esta mayor precisión no tenía ningún significado, ya que los resultados estaban sesgados de principio por defecto del muestreo. De hecho, el tamaño muestral no tenía tanta influencia como pudiera parecer, puesto que a partir de cierto valor de n, $1/\sqrt{n}$ se estabiliza enormemente, como se puede ver en la gráfica:

Fig.10: Margen de error para una proporción; confianza 1-α = 0.95 [elab. propia]



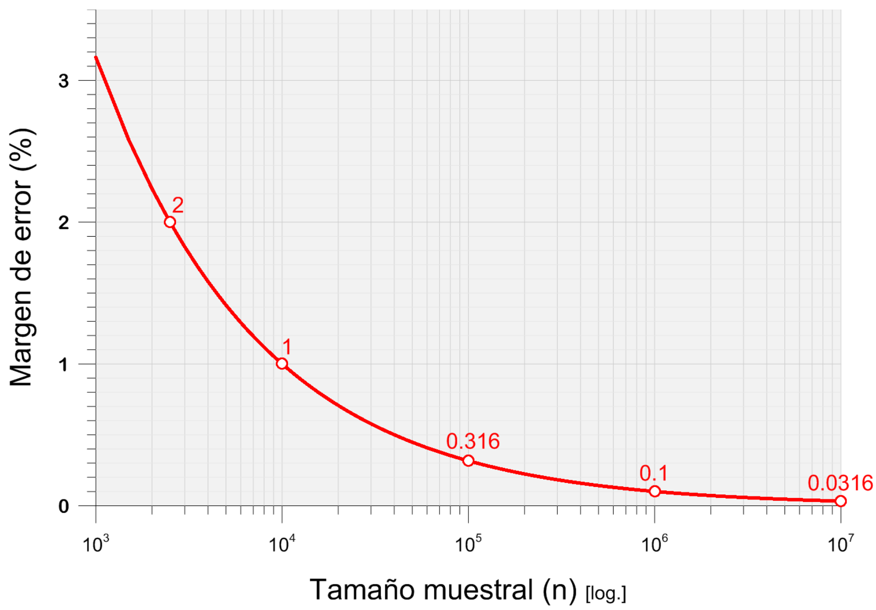
Para los valores del tamaño de la muestra (n) que se suelen utilizar en encuestas de opinión, (n>2000), el margen de error disminuye tan lentamente que muy pronto afinar la estimación resulta prohibitivo en términos de coste (que suele seguir una función lineal dependiente de n). Y sin embargo, a continuación vamos a exponer un par de casos en los que esa pequeña diferencia vuelve a tomar protagonismo, es decir, a ser imprescindible para analizar los datos adecuadamente.

1 *Vid. nota 3, p. 39*

Muestras pequeñas

Teniendo en cuenta la gráfica anteriormente expuesta, ¿Qué sucedería si la proporción de individuos de la población estudiada en los que se presenta la característica que queremos medir fuese muy pequeña? ¿Cómo compararíamos dos muestras de esta población? Si se constata que entre un 0.01% y un 0.02% hay una diferencia del 100%, nos interesa tener un instrumento de mucha precisión (de muy poco error) para evaluar tal diferencia.

Fig.11: Margen de error para una proporción; Rango entre 0 y 3.5% 1-α = 0.95 [el.pr.]



Como se comprueba en la gráfica, es necesario recurrir a la escala logarítmica para encontrar un margen de error de un orden menor que los porcentajes antedichos; y ello a costa de una muestra de tamaño inmenso.

Ahora bien, hagamos por un momento una lectura inversa: el margen de error para la estimación de una proporción depende (o, mejor dicho, se deduce), en efecto, de la desviación típica de la distribución muestral de esa proporción[2]. Esto es, si cuanto más grande es el tamaño n de una muestra menor es el margen de error del estimador de una proporción, esto se debe a que la distribución del estimador de la proporción es tanto menos dispersa cuanto

2 Si suponemos que p es la proporción de elementos de un conjunto que tienen determinada característica de interés, considerando las variable aleatorias $X_1, X_2, \ldots$ donde $X_i=1$ si encontramos un elemento de la característica buscada y $X_i=0$ en caso contrario, entonces la repetición del experimento dará como resultado una suma de variables aleatorias cuya distribución será Bi(n,p), siendo n el número de variables (número de intentos). Por el Teorema de DeMoivre-Laplace (central del límite para binomiales), esa misma suma de variables se distribuirá asintóticamente (n-->∞)siguiendo una dist. normal N (np, np(1-p)); tomando una muestra de esa población de variables de tamaño n, se sabe asimismo (DeMoivre) que la distribución de su media $\overline{X}$ será N(np, p(1-p)). Consideremos $\hat{p}$ como la proporción entre el valor esperado de esa muestra (es decir, el número de veces que se da la característica buscada en el conjunto de variables de la muestra) y su tamaño n, de manera que $\overline{X}=\hat{p}\cdot n$; entonces, se verifica que $\hat{p}$ es un estimador de p. puesto que $E(\hat{p})=E(\overline{X}/n)=E(\overline{X})/n=np/n=p$ y, además,

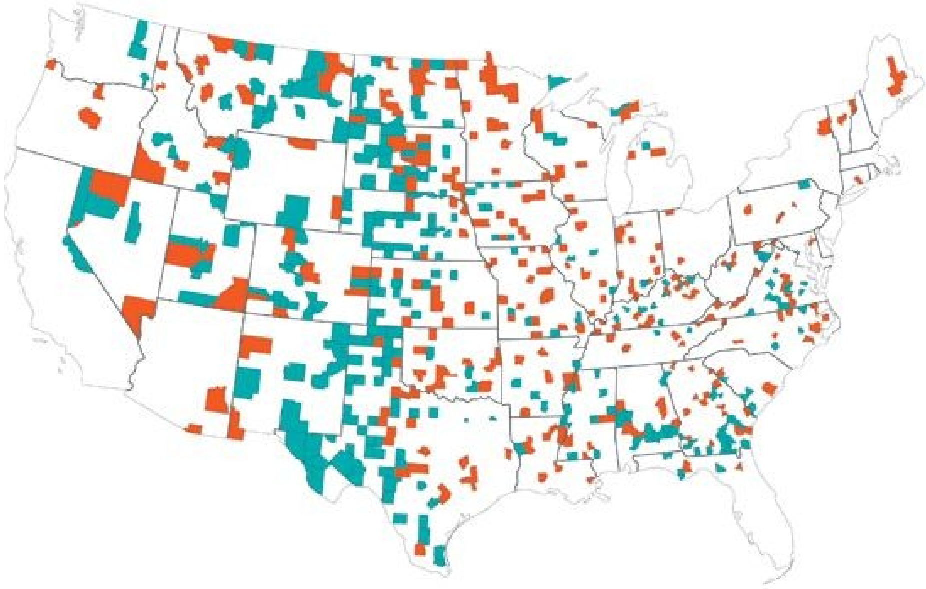
$$\mathrm{Var}(\overline{X})=np(1-p)=\mathrm{Var}\left(\hat{p}\cdot n\right)=(\text{aplicando propiedades de la varianza})\ n^2\cdot\mathrm{Var}(\hat{p}) \longrightarrow \mathrm{Var}(\hat{p})=\frac{\cancel{n}p(1-p)}{n^{\cancel{2}}}=\frac{p(1-p)}{n}$$

de donde la desviación típica de $\hat{p}=\sigma_{\hat{p}}=\sqrt{\frac{p(1-p)}{n}}=\frac{\sqrt{p(1-p)}}{\sqrt{n}}$, que depende, al igual que el margen de error, de $\sqrt{n}$

mayor es el tamaño de la muestra tomada para realizar la estimación.

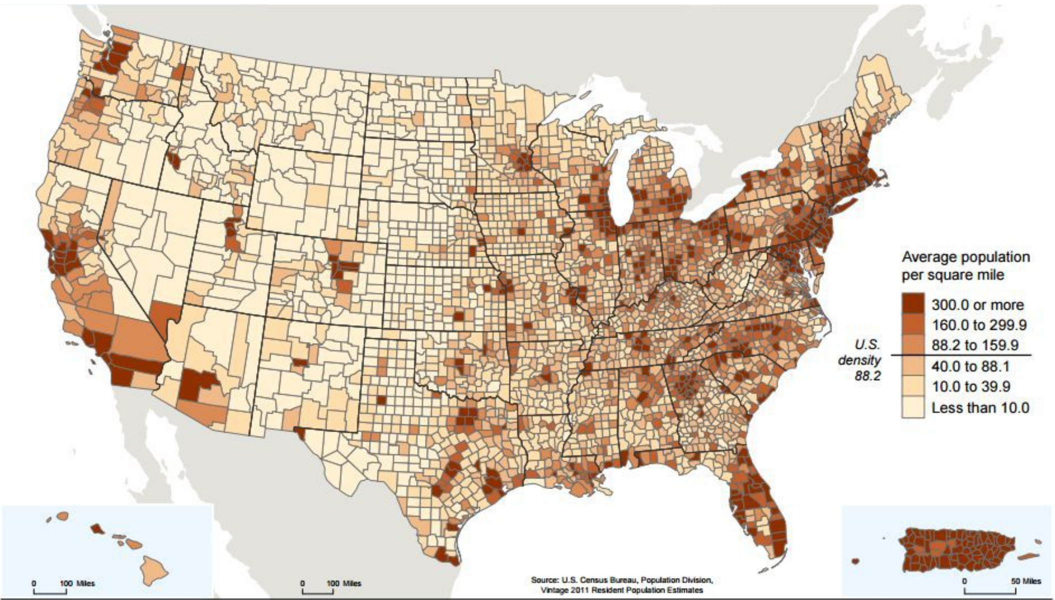
Es decir, que la distribución del estimador de una proporción es tanto *más* dispersa cuanto *menor* es el tamaño de la muestra: a continuación consideraremos un ejemplo de la influencia de estas consideraciones[3]

Distribución geográfica del cáncer de riñon en USA por condados.
Azul → menor incidencia
Rojo → mayor incidencia



D. Schoonmaker *The most dangerous equation.-Fig.2* *Fuente: americanscientist.org*
http://www.americanscientist.org/include/popup_fullImage.aspx?key=5fHzZ5xPSNZBc3MTZWRyGyfhW7w171KE

Densidad de población en USA por condados. Datos de 2011



Fuente: census.gov US Census Bureau
https://www.census.gov/popest/data/maps/2011/County-Density-11.html

3 H.Wainer, *The most dangerous equation* [artículo en línea], American Scientist, Mayo-Junio 2007 [consultado Mayo 2015] Vol.95, nº3 p. 249, en American Scientist (sitio en línea), *americanscientist.org*. Disponible en: http://www.americanscientist.org/issues/num2/ 2007/3/the-most-dangerous-equation/1

He aquí la distribución por condados *(counties)* del cáncer de riñón en Estados Unidos, y más abajo la densidad de población de ese mismo país, también por condados. Si nos fijamos en las áreas señaladas en color azul, donde este tipo de cáncer tiene menor incidencia, podríamos sacar como conclusión, a la vista del gráfico inferior, que menor incidencia corresponde a menor densidad de población, esto es, a menor urbanización: de lo cual se podría inferir que la vida campestre , debido a la baja contaminación, a la alimentación natural, etc. provoca que haya menores tasas de este tipo de enfermedad.

Ahora bien, más tarde nos informan de que las áreas señaladas en color rojo son las que tienen *mayor* incidencia de cancer de riñon. Y casualmente coinciden *también* con zonas de baja densidad poblacional. Si no nos hubieran informado de las áreas señaladas en azul, quizá habríamos llegado a la conclusión de que la zonas campestres, con su mayor índice de pobreza, y relativo atraso cultural, sumada al aislamiento de los centros médicos importantes han causado esta concentración de casos en zonas poco pobladas.

¿Cómo explicar esta aparente paradoja? Sencillamente, basta considerar que la población de cada condado es una *muestra aleatoria*[4] de la población total, y que, por tanto, la incidencia, es decir, la proporción de muertos por esa enfermedad entre un número de habitantes dado es un estimador de la proporción de la población. Entonces, como se expuso más arriba, esta estimación es tanto menos precisa (tiene un margen de error mayor) cuanta *menor sea el tamaño de la muestra* (en este caso, la población de cada condado). Por lo tanto, la **variabilidad o dispersión** de una estimación con n pequeño (condados poco poblados) es **mucho mayor** que la de una estimación con n grande (condados densamente poblados). De ahí que los resultados sean más extremos en ambas direcciones en los condados con menor población, lo que es el caso, y de lo que se puede deducir que, en líneas generales, la distribución real de incidencia del cancer de riñón en USA no esté determinado por ningún factor geográfico. Pero abundemos más en ello en las actividades.

ACTIVIDADES

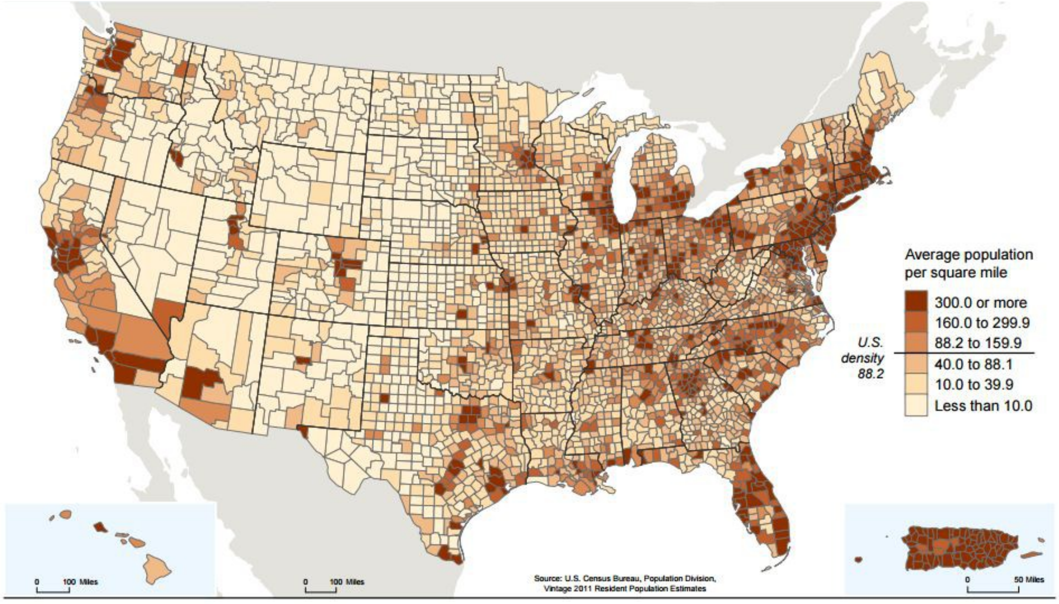
1º) Observa atentamente los dos gráficos siguientes:

· El primero representa la densidad de población de los USA por condado (medidos en habitantes por milla cuadrada)
· El segundo representa una extracción de aquellos condados en los cuales la incidencia de cierta enfermedad (nº muertes/habitante) es menor (áreas coloreadas de azul)

a) ¿Qué característica, a la luz del primer gráfico, puedes decir que tienen la inmensa mayoría de los condados "azules" del segundo?
b) ¿Qué conclusiones pueden sacarse, que relacionen ambas características (baja incidencia de la enfermedad, y la que hallaste en el anterior apartado)
c) Trata de justificar razones plausbles por las que se da la anterior relación (la que existe entre la baja incidencia y la otra característica)

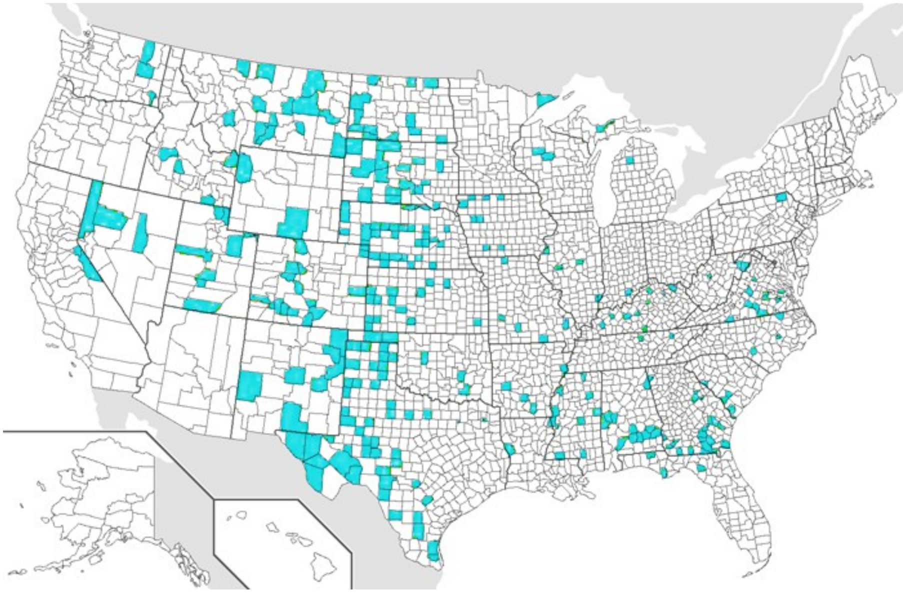
4 Lo cual no es estrictamente cierto: no obstante, como el autor del artículo (vid. nota 3) señala, basta ajustar por edad las poblaciones de los condados respectivos para que la muestra sea asimilable a una aleatoria simple.

Densidad de población en USA por condados. Datos de 2011



Fuente: census.gov US Census Bureau
https://www.census.gov/popest/data/maps/2011/County-Density-11.html

Fig. 12: Distribución geográfica del cáncer de riñon en USA por condados.
Azul → menor incidencia



Las tres preguntas se respondieron en nuestra introducción a las actividades. Sólo son un paso previo antes de la "sorpresa" de la siguiente actividad.

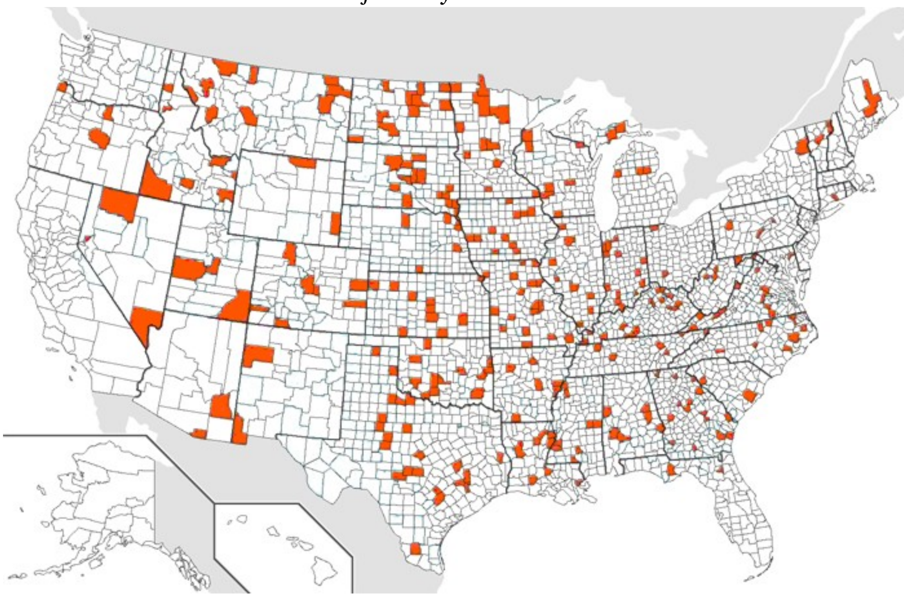
2°) Ahora presta atención al gráfico a continuación:Representa en color rojo el conjunto de los condados que, a diferencia de la anterior actividad, tienen la **mayor** *incidencia de cierta enfermedad.*

a) ¿Qué característica, a la luz del primer gráfico de la anterior actividad, puedes decir que tienen la inmensa mayoría de los condados "rojos" del segundo?

b) ¿A qué sorprendente (teniendo en cuenta la anterior actividad) conclusión puede llegarse, que relacionen ambas características (alta incidencia de la enfermedad, y la *que hallaste en la anterior actividad)*

c) Imagina que no sabes nada de la actividad 1º, ni de su enunciado ni de su respuesta. Trata de esgrimir razones plausibles por las que pudiera darse la aparente relación entre alta incidencia y la característica que se deduce del gráfico de densidad de población acerca de los condados "rojos".

Fig. 13: Distribución geográfica del cáncer de riñon en USA por condados. Rojo→ mayor incidencia



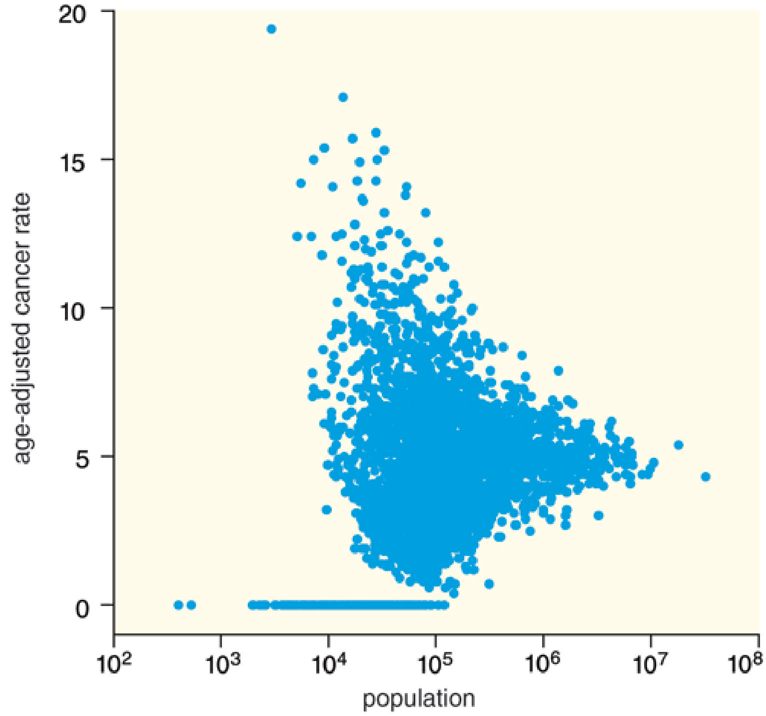
Una de las intenciones de la última pregunta es que el alumno tome consciencia (aunque quizá haya que ponerlo de relieve más claramente) de que ante un conjunto cualquiera de datos se puede especular sobre infinitos modos de relación, y que a menos que se analicen aquellos de una manera exhaustiva, con criterios claros y científicos, es fácil caer en el prejuicio, la trivialidad o el pensamiento analógico. De manera que hay que desconfiar de las tesis apresuradas que suelen manejar los medios; unos medios que, a veces por negligencia y a veces en defensa de su particular agenda política, se pueden aventurar tesis, proclamas, profecías o titulares con poco o ningún apoyo en datos consistentes.

3ª) Ahora pasemos a considerar el siguiente gráfico:

· Es un diagrama de ejes coordenados en cuya abscisa se cuantifica la población de cada condado de USA (no sólo de los "azules" o "rojos" sino de todos), y en la ordenada se representan los valores de incidencia de la enfermada antes tratada, en nº de muertes por cada 100.000 Habitantes. Nota que la escala del eje de abscisas es logarítimica, de modo que a un intervalo de una unidad medida sobre la gráfica le corresponde un intervalo de valores correspondiente a una unidad de potencia de la base expresada (es decir, la distancia entre 2 y 3 medida en la gráfica corresponde a la diferencia de valores entre 10^2 y 10^3)

a)Detecta y compara los valores de incidencia correspondientes a los condados con población de 20.000 habs. con los valores de los condados de más de 1.000.000 habs. ¿Cómo describirías la diferencia que observas?

Incidencia ajustada a la edad de cáncer de riñón en USA por condados según su nº habs.



S. Freese The most dangerous equation.-Fig.3 *Fuente: americanscientist.org*
http://www.americanscientist.org/include/popup_fullImage.aspx?key=5fHzZ5xPSNZ1TOPW7YmZqSYRugVmuJi+

Efectivamente, la diferencia que se observa está en la variabilidad o dispersión del valor de la incidencia: mientras que en las poblaciones de 20.000 habs. los valores se encuentran entre 0 y 17/100.000, en las poblaciones de más de un millón de habitantes la diferencia máxima está entre 3/100.000 y 8/100.000, una diferencia tres veces menor. Si la población aumenta del millón, la diferencia se estrecha aún más.

b) ¿A qué crees que es debida esta diferencia?¿Podría tener algo que ver con la (aparente) paradoja mostrada por los gráficos (mapas) anteriores?

La gráfica parece indicar que, a menor población, mayor variabilidad o dispersión de los datos de incidencia en cada condado.

Y sí, efectivamente, lo que distingue a los datos de incidencia de los condados menos poblados no es que su tasa sea mínima o máxima, sino que adopta *valores extremos* en ambos sentidos. Es decir, lo que distingue tanto a los condados "azules" como "rojos" es la *dispersión* de la distribución de sus tasas de incidencia. Y la gráfica actual indica que esta dispersión está estrechamente (e inversamente) relacionada con el númeo de habitantes.

4°) ¿Podría considerarse la población de cada condado (al menos, aproximadamente) como una muestra aleatoria simple de tamaño n (el número de habitantes) de la población total de los USA? ¿Hay algún concepto estadístico que conozcas cuyos resultados de-

pendan esencialmente del tamaño de la muestra n?

Una vez que los valores están ajustados por edad (después se incidirá en la cuestión), sería bastante plausible, *siempre* que la característica que queramos investigar de la población sea independiente de su situación geográfica. Es decir, siempre que los casos de muerte por la dicha enfermedad estén aleatoriamente distribuidos por toda la geografía de los EEUU. Se indicará al alumno que solo bajo esa hipótesis el contenido de la siguiente actividad puede dar resultados interpretables.

Y sí, efectivamente, el alumno sabe algo(o está apunto de saber en la teoría) acerca de intervalos de confianza para una proporción; se puede contar con su conocimiento de la fórmula simplificada $\varepsilon = \pm 1 / \sqrt{n}$ para determinar el tamaño de una muestra para un margen de error determinado (con un nivel de confianza del 95%). No obstante esta simplificación es pertinente sólo en el caso de que la proporción a estimar no sea ni muy grande ni muy pequeña, y en este caso, como puede deducirse de la gráfica, es del orden de 5/100000; será necesario entonces explicar someramente el concepto de intervalo de confianza para una proporción. Una vez que se haya introducido, este mismo ejemplo es muy útil como ilustración del mismo. Veamos: para el supuesto de que *cada condado* es una muestra aleatoria simple de la población total estadounidense, la proporción recogida por *cada condado* puede considerarse una estimación de la proporción en el total de la población, p.

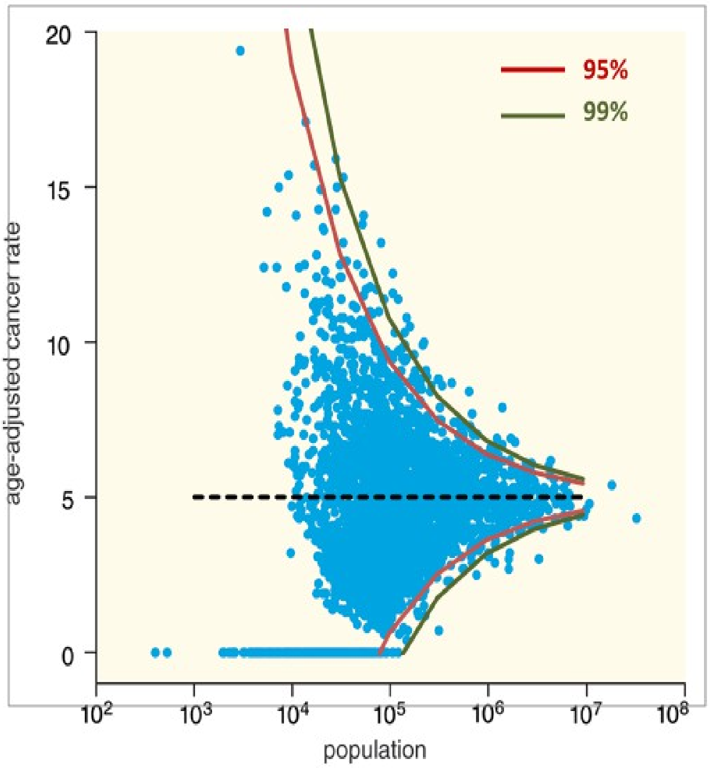
4º) De la gráfica se deduce que el valor de incidencia total en los EEUU de la enfermedad es aproximadamente 5/100000. Bajo el supuesto del problema anterior, calcula el radio de un intervalo de confianza al 95% de los condados de 10.000, 100.000 y 1000.000 de habitantes, suponiendo que cada uno estimara $\hat{p}$ como el anterior valor supuesto según la gráfica (5/100000). ¿Hay alguna "clase" de condados cuyo número de habitantes n permita una estimación adecuada de p?

Aplicando el concepto de intervalo de confianza al 95%, tenemos, para los valores de n del problema:

$$r_n = \sqrt{\frac{p(1-p)}{n}} \cdot z_{\alpha/2} \longrightarrow \begin{cases} n = 10000 \Rightarrow r_n = \sqrt{\frac{0,00005(1-0,00005)}{10000}} \cdot 1,96 = 0,0001386 \\ n = 100000 \Rightarrow r_n = \sqrt{\frac{0,00005(1-0,00005)}{100000}} \cdot 1,96 = 0,0000432 \\ n = 1000000 \Rightarrow r_n = \sqrt{\frac{0,00005(1-0,00005)}{1000000}} \cdot 1,96 = 0,0000138 \end{cases}$$

donde se puede observar que el valor más pequeño del radio r, es decir, la estimación más precisa, supone un ±27% del valor de la p que hemos supuesto (considerada igual a la estimada, es decir, en el mejor de los casos). Por lo tanto, ninguna de estas "clases" de condados nos daría una estimación adecuada de la verdadera proporción; ello se debe a que el valor de incidencia, p, es muy pequeño (o el número de habitantes, no lo suficientemente grande)

Para abundar en el concepto, se podría construir una gráfica que expusiera, alrededor del valor 5/100000, los valores de los radios de los intervalos de confianza resultantes para cada número de habitantes (abscisas) y diversos grados de confianza, quedando:



donde se puede apreciar, no sólo que nuestra hipótesis acerca de los condados como simulaciones de muestras aleatorias simples es muy cercano a la realidad (en el espacio entre las curvas se aprecia perfectamente que se abarcan el 95% y el 99% respectivamente de los valores de proporción de incidencia por cada condado), sino que sólo los condados de cerca de 10.000.000 de habitantes cumplirían con un margen de variación de la estimación aceptable (±10% del valor que se quiere medir, 95% de confianza): sería esta clase de estados los que escogeríamos si quisiéramos hacer una estimación más fina[6], pero la gráfica nos indica que quizá no haya suficientes de los mismos como para que la expresión correspondiente del grado de confianza de un intervalo sea del 95%: esto es, no podemos decir "de cada 100 simulaciones, 95 nos ofrecerán una estimación que estará dentro del intervalo..." porque no hay 100 condados, ni tan siquiera 10, con el número de habitantes requerido.

6ª) Si sabes inglés habrás podido comprobar que las ordenadas está "ajustadas por edad". Aunque no sepas bien en qué consiste este "ajuste", en qué supones que influye la edad en la incidencia de la enfermedad? ¿Cómo afectaría eso a nuestra afirmación de que la

6 Explotaremos las dificultades y aparentes paradojas que resultan de estimar efectos de baja incidencia más adelante: *vid. La falacia del falso positivo*, p. 137

población de cada condado constituye una muestra aleatoria simple del total de la población estadounidense?

Se puede aventurar sin mucho riesgo a equivocarse que, al igual que en España, por ejemplo, los condados menos poblados corresponden a los menos urbanizados; es de suponer que la natalidad en ese entorno es baja, y que, por lo tanto , la población está más envejecida que la media total del país: De esa manera es de suponer que el número de defunciones por habitante en esos entornos, incluidas las de esta clase de cáncer, es mayor que el de un condado con población más joven.

Efectivamente, investigando un poco puede comprobarse fácilmente que las estadísticas sobre incidencia de toda clase de enfermedades presentan casi siempre un rótulo en las ordenadas que especifica que los porcentajes están corregidos o ajustados por edad; de manera que, geográficamente (y sigiuiendo otros criterios de selección de datos, como el género, a veces la ocupación, etc.) las muestras no son verdaderamente aleatorias, dado que las edades de los habitantes no están repartidas de forma equitativa (aleatoria, equidistribuidad) siguiendo las divisiones adiministrativas del territorio.

7º) No obstante la impresión que hayas podido sacar de las actividades anteriores, realmente <u>sí que hay muchas características de la población de un país cuya distribución tiene una razón de ser cultural, socioeconómica, etc. características todas que pueden reflejarse geográficamente</u>. Métodos estadísticos más sofisticados se utilizan para estudiarlos, y gran parte de los mismos estriban en técnicas cuyo fin es, precisamente, eliminar efectos ilusorios como los estudiados en el caso que nos ocupa. Aunque estos métodos no se estudian en educación secundaria, ¿Se te ocurre alguna estrategia que pudiera indicarte si la variabilidad antes observada tiene o no que ver con la situación geográfica?

En esta pregunta, si el alumno no nos sorprende con algún método de su cosecha, se podría indicar a modo de ejemplo el siguiente: supongamos que tenemos los datos de esos condados, a lo largo de, digamos, 20 años, y que el mapa antes expuesto es el resultado de la suma de todos esos datos. Como primer acercamiento, se podrían tomar esos datos, que suponemos están fechados, y agruparlos según clases cronológicas, por ejemplo, en 5 grupos de 4 años cada uno, y volvemos a realizar, en este caso, 5 gráficas. Pues bien, si la distribución de incidencia es verdaderamente aleatoria y no tuviera nada que ver con la zona geográfica, en cada gráfica se observaría, no obstante, una varibilidad similar: lo que sucedería, empero, es que, observados gráfica a gráfica, los datos no conservarían su regularidad; es decir, que lo más probable es que un condado que fuese "azul" en la primera gráfica pasara a ser "rojo" en la siguiente, y otra vez azul en la quinta gráfica, por ejemplo; es decir, <u>si el efecto de los valores extremos depende solo de la variabilidad de los datos, esta es completamente aleatoria en el tiempo</u>. Sólo si observásemos que, regularmente, los condados "azules" son, de forma aproximada, consistentemente "azules" en las 5 gráficas, y lo mismo con los condados "rojos", entonces sí que habría que tomar esta variabilidad en serio como un efecto debido a una causa operante, y no a la variabilidad de la muestra.

Como se puede ver en este gráfico

Los medios de comunicación de masas dependen vitalmente de la imagen (dependencia que podría decirse que ha seguido una curva exponencial a lo largo del tiempo), hasta el punto de que la imagen suele a veces sustituir a la noticia, e incluso engendrarla. ¿Cuántas noticias no son, efectivamente, noticias, sólo porque se dispone de imágenes más o menos impactantes acerca de un suceso que, a fin de cuentas, y por dramático que sea, muchas veces no rebasa lo puramente anecdótico? La imagen no sólo sustituye a la noticia, sino, en muchos más casos, al mismo razonamiento —que suele estar bastante oscurecido de partida por la misma escritura propia del medio, intencionadamente o no— . Y en este fenómeno desempeña un papel muy importante el gráfico estadístico.

Si, como sostienen los teóricos de la redacción periodística, todavía tiene vigencia la clásica organización de la información en forma de pirámide invertida[1] (consistente en disponer la información en orden decreciente de importancia o "interés periodístico" desde el titular hasta el final del artículo, lo cual presupone tácitamente que el lector no pasa de su primer párrafo), el gráfico estadístico, que combina impacto visual e información categórica y/o numérica y que se dispone inmediatamente después del titular, a menudo *es la noticia a todos los efectos*; es por ello que en el gráfico se debe concentrar la información, y también, y muy a menudo, la desinformación, según las intenciones del articulista o la publicación; su importancia es asimilable a la del titular, o, cuando menos, al del párrafo introductorio (*lead* o *intro*), del que se llega a decir lo siguiente[2]:

> *Hemingway, leí una vez, escribió la última página de* Un Adiós a las Armas *dieciséis veces antes de estar satisfecho. Pero son los comienzos los que dan al escritor de periódicos todos los problemas.* [...] *dieciséis intentos de la primera frase o párrafo de una noticia no es nada, como demuestran las cifras de mortalidad de las mismas en los tableros de redacción ejecutiva.* ***Cualquier esfuerzo para conseguir un buen comienzo de una noticia merece la pena, porque el lector la dejará de leer si el escritor fracasa.***[trad. propia]

Sin embargo, el gráfico estadístico no tenía en origen la finalidad de comunicar un determinado mensaje, como sucede en el periodismo. Nació con el honrado y sencillo propósito de resumir y resaltar visualmente la información sobre las variables que, expuestas en tablas, resultaban poco expresivas antes de un análisis detenido. Desde William Playfair (1759-1823), su introductor, pasando por la fusión de estadística, análisis y teoría de la probabilidad, que, pese a Gauss, las funciones generarices de momentos y la curva normal (entre otros nombres y descubrimientos), no tuvo lugar hasta finales del siglo XIX [**20**], y pasando posteriormente por los fundadores de las modernas técnicas de exploración de datos, que lo

1 H. Pö̈ttker, *News and its communicative quality: the inverted pyramid– when and why did it appear?* [artículo en línea], 2003 (pub. en línea 4 Junio 2010) [consultado Junio 2015] Journalism Studies, 4:4, 501-511, en Taylor & Francis Online (*www.tandofline.com*) . Disponible en[previo registro]: http://dx.doi.org/10.1080/1461670032000136596 DOI: 10.1080/1461670032000136596

2 H. Evans, *Essential English for Journalists, Editors and Writers*[libro electrónico], 2º ed, revisada por C. Gillan, Pimlico (Random House Mondadori), London, 2000

desarrollaron como instrumento de análisis propiamente dicho [**45**], hasta la actualidad en la disciplina estadística, esta, la de *resumir y resaltar* ha sido su principal —aunque, como decimos, no única— función dentro de la estadística [**44**]. A partir de la 2º Guerra mundial, los medios de masas, que habían comenzado a conformarse a comienzos de siglo, pasan finalmente a ser accesibles a la mayor parte del espectro social, lo cual provoca en ellos una inmensa avidez de contenidos, y, quizá también, de la respetabilidad que otorga lo cuantitativo, en el sentido de que se percibe como análogo a lo “científico”. Es entonces cuando elementos de la estadística (índices, porcentajes, estimaciones, medias) comienzan a ser de uso regular en la prensa escrita —y en el discurso político, a menudo correlatado con aquélla—, explicando, por ejemplo, el éxito comercial de la obra de Darrel Huff, *Cómo mentir con estadísticas* [**28**](Su primera edición es de 1954), a día de hoy aún el mayor best-seller relacionado con la estadística de todos los tiempos —lo cual dice mucho de la percepción ciudadana de la disciplina, provocada por esos mismos medios a los que sólo un poco más tarde se incorporaría la televisión, y (ya no tan) últimamente, los formatos digitales en línea. Casi desde el primer momento, el periodismo, la publicidad (y la política) pasaron de analizar los datos para buscar titulares a escoger los datos y representaciones gráficas que dieran la razón a titulares concebidos de antemano.

Se nos dirá, ¿Y esto que tiene que ver con la docencia de matemáticas? Pues bien, nos permitimos recordar una vez más que uno de los objetivos que marca el currículo de la ESO es el siguiente (*vid.* *supra*):

1. Utilizar el vocabulario adecuado para la descripción de situaciones relacionadas con el azar y la estadística, analizando e interpretando informaciones que aparecen en los medios de comunicación.

Y además reiteramos la afirmación que acompañaba a la cita anterior en el capítulo correspondiente, a saber: que en el sentido de desarrollar competencias básicas generales de todo ciudadano (no se olvide que la ESO es *obligatoria* por esa sola razón) el análisis de la información —estadística descriptiva gráfica en el caso del presente capítulo— que los medios proporcionan (y que es inmenso) puede tener, *debería* tener un peso relativo en el temario de matemáticas proporcional a la capacidad de juicio y al número y la calidad de instrumentos para la decisión y la acción que otorga.

Noticias ciclotímicas

Empezamos con un ejemplo bastante divertido, sobre (suponemos) los males de abusar de la subcontratación que padecen ciertas publicaciones. Otra cosa no podemos pensar cuando ven la luz artículos de este tipo[3], ante los cuales sólo podemos aventurar que la gráfica y el titular han salido de redacciones distintas, o que los miebros del equipo de grafismo no se hablan con los de la redacción, o que no existe nada que se parezca a un redactor jefe o ejecutivo en la publicación en cuestión. Desarrollaremos todo lo que da de sí la noticia en cuestión en las actividades correspondientes.

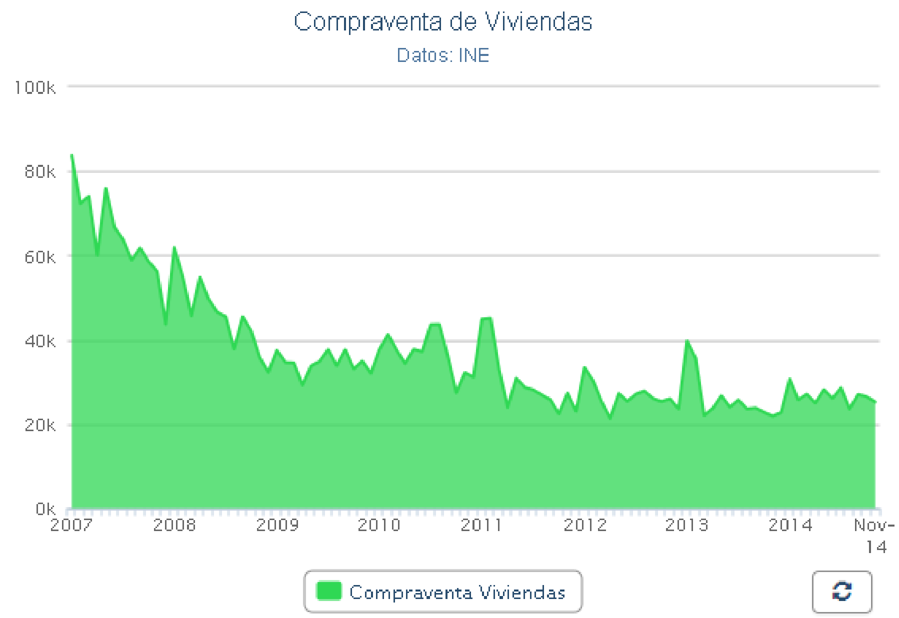
3 Europa Press, *La Venta de viviendas sube el 14% en noviembre y encadena tres meses al alza* [artículo en línea], 12 Enero 2015 [consultado Junio 2015]. El Huffington Post (*www.huffingtonpost.es*) Disponible en: http://www.huffingtonpost.es/2015/01/12/compraventa-viviendas_n_6454622.html

ACTIVIDADES

La venta de viviendas sube el 14% en noviembre y encadena tres meses al alza

EL HUFFINGTON POST / EUROPA PRESS

Publicado: 12/01/2015 12:13 CET | Actualizado: 12/01/2015 12:13 CET



La compraventa de viviendas aumentó un 14% el pasado mes de noviembre respecto al mismo mes de 2013, hasta un total de 25.200 operaciones, debido al fuerte repunte de las operaciones sobre inmuebles de segunda mano, según ha informado este lunes el Instituto Nacional de Estadística (INE).

Este avance interanual es el tercero consecutivo tras los de septiembre y octubre, cuando la compraventa de viviendas creció un 13,7% y un 16%, respectivamente, después de la ligera caída del 1,1% que se experimentó en agosto, mes en el que se marcó el mínimo del año con apenas 23.500 operaciones.

Los datos difundidos por el INE corresponden a compraventas inscritas en los registros de la propiedad procedentes de escrituras públicas realizadas en meses anteriores al de referencia.

Atendiendo sólo a los datos mensuales (noviembre de 2014 sobre octubre del mismo año), la compraventa de viviendas bajó un 4,8%, su mayor descenso en este mes de los últimos cinco años.

www.thehuffingtonpost.es

1ª)No leas atentamente esta noticia todavía; simplemente lee el titular y luego observa la gráfica, teniendo en cuenta que las ordenadas representan "miles de ventas de pisos". a)¿Te parece que el titular y el gráfico están hablando del mismo asunto? Si tenemos en cuenta que cada marca del eje de abscisas corresponde a un mes, ¿Te parece que la ven-

ta de pisos ha subido en los últimos tres meses, como afirma el titular?

Si nos fijamos sólo en el término “subida” y en el gráfico, efectivamente, parece que hablan de cosas completamente distintas. Si después seguimos leyendo “subida del 14% en noviembre y encadena tres meses al alza” y después observamos la gráfica, sencillamente ya no sabemos qué pensar: las tres últimas marcas señalan una bajada, no un alza.

b) Imagina que sólo tuvieras a tu disposición la gráfica, sabiendo no obstante cuáles son las variables reflejadas. ¿Cuál sería el titular que escogerías?¿Se parece en algo al titular que realmente tiene la noticia?

Podría titularse algo así como “Las ventas de pisos en España han bajado un 70% en los últimos siete años”. Siendo muy benévolo, podría interpretarse la secuencia que comienza en Enero de 2014 como un cambio de tendencia hacia una lenta recuperación y titular “Ligero repunte en las ventas de pisos en España durante 2014” Por supuesto que no me fijaría en el ínfimo tramo de los últimos tres meses: más bien lo tomaría como una fluctuación local esencialmente azarosa (aleatoria).

2º)Ahora sí, puedes leer atentamente la noticia.
a) ¿Te aclara algo las cosas las dos primeras líneas? ¿Dicen lo mismo que el titular? ¿El gráfico propuesto es la mejor forma de poner de relieve esas dos líneas?

La acotación “aumentó un 14% el pasado mes de noviembre <u>respecto al mismo mes de 2013</u>” ya sí que parece coincidir con lo que expresa la gráfica. No obstante, nos haría falta aplicar escuadra y cartabón en la pantalla y tener una calculadora a mano para hallar el <u>porcentaje</u>, que no aparece reflejado directamente en la gráfica. El responsable del artículo habría hecho mejor en presentar las tablas de valores en bruto y no haberse molestado dibujando.

b)Teniendo en cuenta que tanto la gráfica como las dos primeras líneas del artículo desmienten el titular, ¿Cuál crees tú que ha sido la motivación del titulador?

Su intención parece ser la de dar a entender a toda costa que las ventas de pisos suben en España, ya que el titular, de la forma que está redactado, no es que sea incompleto, sino que es completamente erróneo.

3º)Fíjate ahora en el último párrafo.
a)¿A qué tramo de la gráfica se refiere en primer lugar?¿Podrías medirlo sobre ella? ¿Crees que este dato es igual de relevante que el titular de la noticia?

Se refiere al pequeño segmento entre las marcas de Octubre y Noviembre de 2014, aunque, como dijimos antes, al estar reflejados los datos absolutos, tendríamos que hacer un cálculo aparte para hallar el dato concreto (4,8%) que expresa la noticia.
Este párrafo desmiente por completo el titular de la noticia, y esta vez no gráficamente,

sino explícitamente. En realidad, el dato relevante es el que el titular de la noticia parece querer expresar —la variación **interanual—** pero no es eso lo que hace, tal y como está redactado.

Magnitudes absolutas (y porcentajes irrelevantes)

Según el FMI[4] China es el país más rico (PIB PPA[5]) del mundo. Incluso aunque el dato tenga significado y utilidad para muchos tipos de análisis, ¿Significa eso que los chinos son los ciudadanos más ricos del mundo? Para conocer algo sobre ese particular (con muchas reservas, dada la siempre desigual distribución de la renta dentro de cada país) una primera aproximación sería averiguar el PIB per cápita, sencillamente dividiendo por el número dehabitantes del país considerado. Sencillo, ¿Verdad? Pues bien, en la mayor parte de los medios de información aún parecen no haberse percatado de la cuestión: no son raras, ni mucho menos, las noticias que siguen manejando magnitudes absolutas e inferir de ellas características de la sociedad o la economía de un territorio. Cuestión para la cual una magnitud absoluta no sirve de nada. En efecto, ¿En qué comunidades autónomas sube más el paro en cifras absolutas? Evidentemente, en Madrid, Cataluña y Andalucía, siempre: pero por el mero hecho de que son las comunidades con mayor población (empleada). El dato que nos daría una verdadera información del comportamiento del paro según cada territorio sería la tasa de cambio por habitante, como *parece* obvio. Aunque si fuera tan obvio es de suponer que no habríamos encontrado el siguiente material:

ACTIVIDADES

1º) Observa atentamente la siguiente gráfica[6]:

a) ¿Encuentras algo incorrecto en ella?

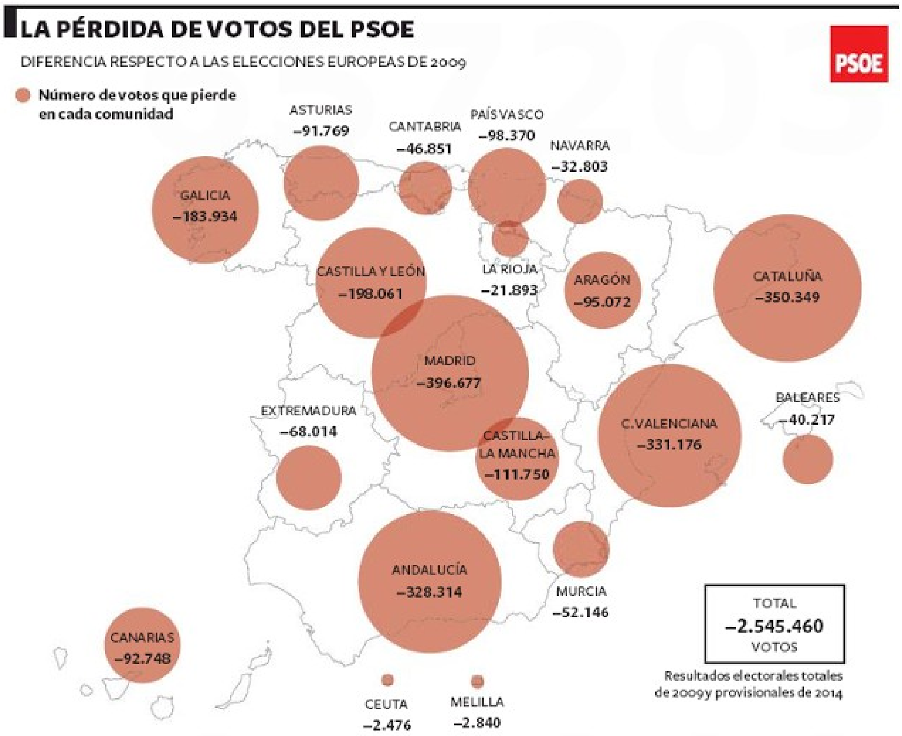
Aquí lo que más salta a la vista son los círculos cuya área expresa la pérdida de votos. Pero así como más adelante veremos noticias donde el error (o la desinformación) se encuentra en la escala del objeto elegido para la representación de la magnitud noticiable, este no es el caso: las pérdidas de votos son realmente proporcionales al área de los círculos. El ejemplo propuesto es conveniente en parte por esta razón: quizá el alumno, acostumbrado a otros enunciados parecidos a los que aquí se proponen, caiga en el error de buscar defectos en la representación por analogía. En esta ocasión tendrá que profundizar un poco más en el significado, no en la representación. En realidad, el gráfico no sostiene ninguna falsedad: lo que sucede es que la información que ofrece es irrelevante.

b) ¿A qué puede ser debido que la pérdida de votos sea mayor en las comunidades donde así lo expresa la gráfica? ¿Puede que el hecho obedezca solo a causas políticas? Redactad por grupos un breve párrafo con vuestra opinión sobre el asunto.

4 Según datos del FMI *International Monetary Found.* disponibles en [consultado Junio 2015]: http://www.imf.org/external/data.htm

5 Producto Interior Bruto ajustado a Paridad de Poder Adquisitivo.

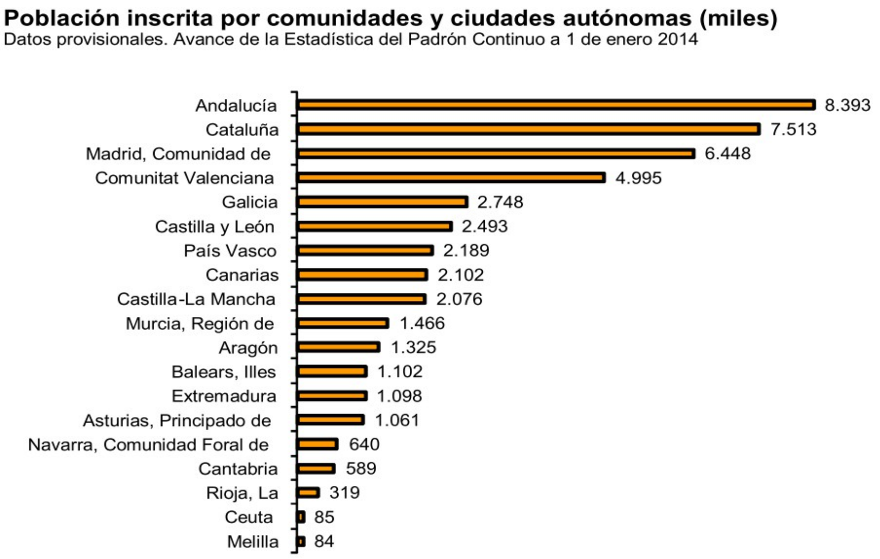
6 Josu Mezo, *Se pierden más votos...¡Donde se tiene más votos!*[artículo en línea] 28 Mayo 2014 [consultado Junio 2015] en www.malaprensa.com . Disponible en: http://www.malaprensa.com/2014/05/se-pierden-mas-votos-donde-se-tienen.html



Fuente: www.malaprensa.com

La pregunta peca de cierta deslealtad, pero se realiza para después obtener más fácilmente un efecto sorpresa. Hay muchas probabilidades de que los alumnos aventuren hipótesis políticas de lo más variopinto, si no caen el cuenta del *quid* de la cuestión. Si nadie da con la respuesta correcta, podemos pasar a la siguiente actividad:

2º) Ahora observa el gráfico siguiente[7]*.*

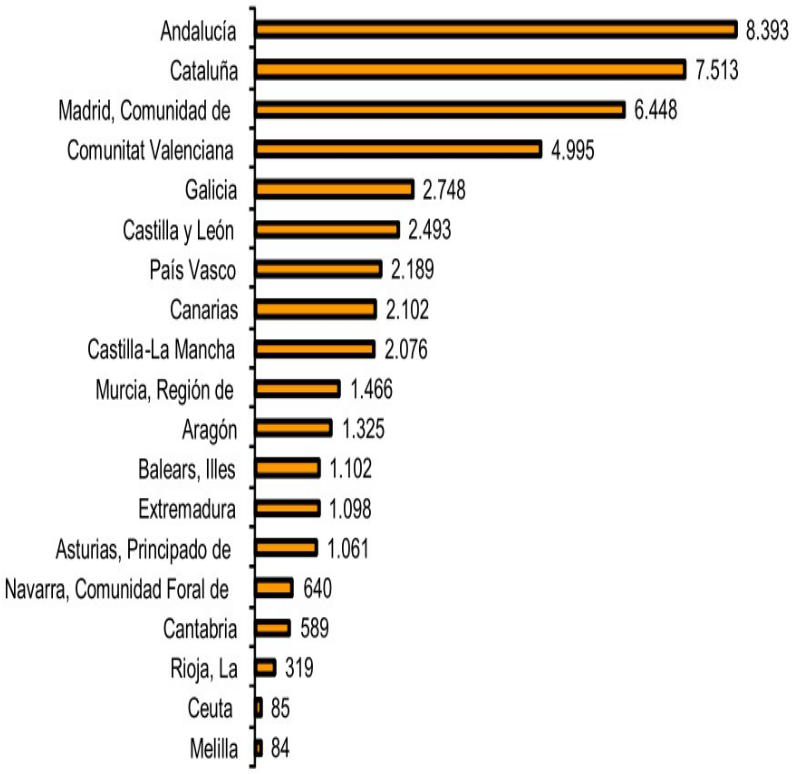


Fuente:www.ine.es

7 Instituto Nacional de Estadística (INE), *Avance de la Estadística del Padrón Continuo a 1 de enero de 2014. Datos provisionales* [recurso en línea], 22 Abril 2014 [consultado Junio 2015] en www.ine.es Disponible en: http://www.ine.es/prensa/np838.pdf

a)¿Encuentras alguna semejanza entre la gráfica de la primera actividad y ésta? Ordena en sentido descendente a las comunidades autónomas según el número de votos perdidospor el PSOE y después compara esa lista con la presente gráfica.

- Madrid
- Cataluña
- C. Valenciana
- Andalucía
- Castilla y León
- Galicia
- Castilla La Mancha
- País Vasco
- Aragón
- Canarias
- Asturias
- Extremadura
- Murcia
- Cantabria
- Baleares
- Navarra
- La Rioja
- Melilla
- Ceuta



Si bien se percibe un cierto "baile" de CCAAs entre uno y otro listado (lo que responderá , probablemente. al fenómeno que en realidad debe tener algún significado político-social) se constata que en términos generales la correspondencia entre ambas se observa bastante bien. (Si la última gráfica relejara al número de votantes del PSOE en las elecciones anteriores, la correspondencia sería más fuerte). Lo cual quiere decir que mediar pérdida de votos en terminos absolutos es solo una manera enrevesada y no muy exacta de medir qué CCAAs tienen más o menos habitantes. Para conocer la tasa de pérdida de votantes es preciso convertir nueos datos absilutos de voto en datos relativos al número de votantes (o al número de votantes del PSOE).

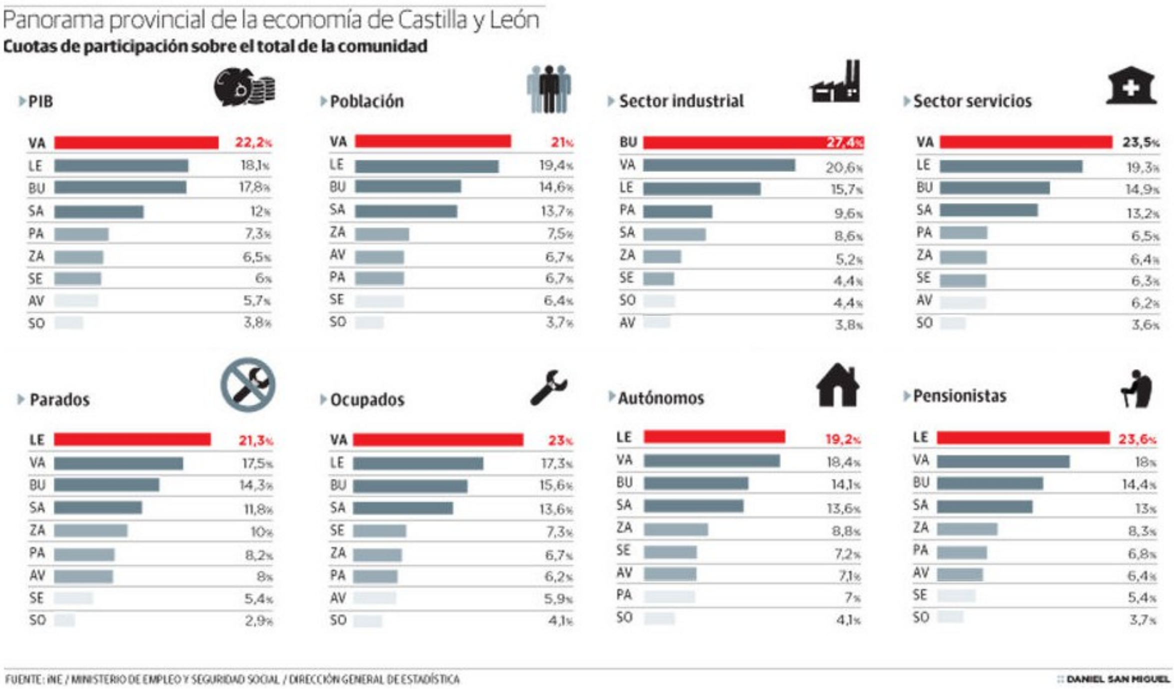
3°) Observa este conjunto de gráficas[8]:

a) Hemos visto que, en muchos casos, si se quiere establecer una comparación significativa entre varias clases, es conveniente escoger variables que expresen magnitudes relativas, no absolutas: ¿Se emplean magnitudes relativas en este caso? ¿Se puede decir, por tanto, que el gráfico establece comparaciones significativas (o informativas)?

Una vez más tratamos de poner en guardia al alumno contra el pensamiento analógico o el empirismo fácil que domina buena parte de la educación secundaria. Un alumno

8 A. Blanco Escalona, *Burgos, Palencia y Valladolid encabezan la riqueza por habitante por su peso industrial* [arículo en línea], 4 Junio 2015 [consultado Junio 2015], en El Norte de Castilla (ed. digital *www.elnortedecastilla.es*) Disponible en: http://www.elnortedecastilla.es/economia/201506/04/burgos-palencia-valladolid-encabezan-20150531201822.html
El hallazgo y sugerencia de la inclusión de este artículo son debidos a Alfonso Gordaliza.

poco atento sin duda sólo se fijaría en que en las gráficas aparecen porcentajes y con-



Fuente: www.elnortedecastilla.es

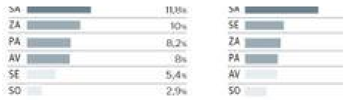
cluiría que las proporciones expresada son relevantes: puede que lo sean para según qué estimaciones, pero son muy pobres desde el punto de vista informativo, ya que lo que nos interesa no es comparar la situación relativa de cada provincia con las demás según estas categorías, sino comparar la situación relativa de los habitantes de cada provincia según cada categoría. No nos interesa saber qué provincia, tomada como objeto, es más rica que las demás; nos interesa saber de qué provincia son los habitantes más ricos de la comunidad. Asi, según las gráficas Valladolid es la provincia con mayor PIB (mayor riqueza) pero ¿No es también la que más habitantes tiene? ¿Son los habitantes de Valladolid más ricos que los del resto de provincias de Castilla y León?

4º) Los gráficos anteriores acompañaban la siguiente noticia (vid. nota [8]). Tras leerla detenidamente, reconsidera si las gráficas son realmente informativas y explica en un párrafo las razones de tu conclusión.

Burgos, Palencia y Valladolid encabezan la riqueza por habitante por su peso industrial

Concentran el 58% del sector y aportan al PIB de Castilla y León cinco puntos más que población

ÁNGEL BLANCO ESCALONA | VALLADOLID
@angelblancoes
Me gusta 13
4 junio 2015 08:54



La provincia más rica de las nueve que componen Castilla y León es Valladolid;pero solo si se tiene en cuenta el valor absoluto de su riqueza, es decir, el volumen de su Producto Interior Bruto (PIB). Si se toma en consideración el número de habitantes, entonces la que posee una economía más potente es Burgos; mientras que también Palencia aventaja a la provincia vallisoletana. Las tres (que son, junto con León, las más industriales de la comunidad autónoma) tienen un peso económico superior al de su población en el conjunto de la región.

MÁS INFORMACIÓN

- **Las administraciones anuncian en la región el doble de obra pública que en 2014**

La reciente publicación por parte del INE del PIB por provincias en el año 2013 permite no solo conocer al detalle la estructura productiva de cada una de ellas, sino también situarlas en un contexto más o menos singular. Por ejemplo: los 368.701 burgaleses censados, que suponen el 14,6% de los habitantes de Castilla y León, produjeron en el año de referencia 9.686,9 millones de euros, lo que representa el 17,9% de la riqueza regional.

Parece que este es el caso, como en el primero que propusimos, en que la noticia y la gráfica se contradicen, con la diferencia de que aquí el articulista es consciente del asunto e intenta rectificar las gráficas por escrito. ¿Por qué no se modifican éstas desde un principio? La descoordinación entre departamentos parece ser cosa general en la prensa española. El artículo continúa más tarde practicando el mismo ejercicio de reforma de las gráficas, pero a nosotros nos basta con que el alumno repare en el primer párrafo y sepa transmitir su significado con sus propias palabras

Escalas tramposas (o sencillamente torpes)

Aunque ampliamente desvelados desde casi los primeros tiempos de los medios de masas (*vid.* **[27][28]**)y bastante risibles por su obviedad (se puede consultar *www.malaprensa.com*, lugar de referencia en español de compilación de esta clase de manipulaciones, entre otras falacias periodísticas), siguen gozando de magnífica salud en medios escritos y audiovisuales. De entre la gran cantidad de casos disponibles pueden distinguirse los que afectan a gráficas cartesianas, o a otra clse de pictogramas (diagramas de barras, en forma de tarta, icónicos, etc.) Comencemos con los primeros, con dos variables numéricas en ejes coordenados.

En la estadística como disciplina científica las escalas de los ejes donde se miden los valores de cada variable se manipulan frecuentemente por razones metodológicas, es decir, en vistas a encontrar patrones que impliquen alguna clase de orden en los datos (comprimiendo los diagramas de nube de puntos puede distinguirse mejor un determinado modelo de regresión), o por razones expositivas, esto es, en orden a expresar las relaciones entre los datos de forma más contundente (una diferencia relativamente pequeña, pero importante, se magnifica para darle resalte). No obstante, estos propósitos puede pervertirse cuando se involucran intereses no científicos (comerciales, partidistas, o de otro tipo) o cuando el redactor no tiene un conocimiento suficiente de la materia que está tratando. La escala puede modificarse por aplicación arbitraria de la unidad de media el alguno de os ejes (o en los dos) o por truncamiento de algún eje. Aunque es corriente econtrar ambas estrategias combinadas. Veamos ejemplos en las actividades: el inmediatamente inferior se volvió casi célebre.

ACTIVIDADES

1º) Observa atentamente el siguiente gráfico[9]:

a) ¿Encuentras algo en él que te resulte chocante? Razona tu respuesta.

Captura del gráfico con la evolución del paro registrado emitido en el Telediario 1 de TVE

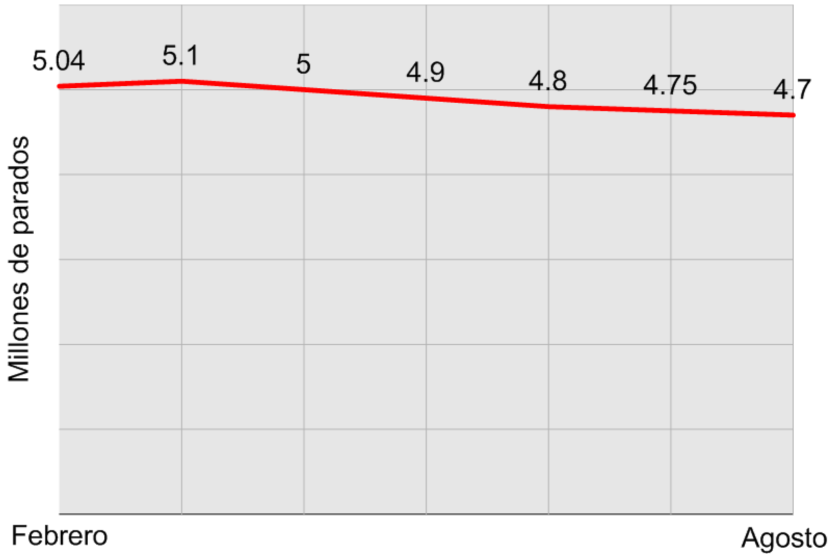
Fuente: www.eldiario.es

Desde luego: si las cifras son ciertas, el inexistente eje de ordenadas está claramente truncado, sin que se refleje tal hecho en ninguna parte, de manera que el valor más bajo de las y puede ser, aproximadamente, 4.500.000; y no sólo eso, la escala de lo que queda de este eje está claramente magnificada para dar a entender una gran diferencia entre los valores del paro de Febrero y Agosto. Un sencillo cálculo arroja que la verdadera variación del paro es una bajada de un 6,77%; y, sin embargo la impresión visual es que alrededor del 90% del paro se ha desvanecido.

b) Si suponemos que la diferencia de valores ordenados en la gráfica es de 500.000, realiza una gráfica aproximada sin el eje truncado y que tenga las escalas proporcionadas de tal manera que encajen, más o menos, en las dimensiones del marco de la foto anterior.

9 [Redacción], *El paro desparece (pero solo en TVE)* [artículo en línea], 3 Sept. 2013 [consultado Junio 2015] en El Rastreador, [bitácora alojada en] www.eldiario.es (periódico digital) Disponible en: http://www.eldiario.es/rastreador/paro-desaparece-solo-TVE_6_171542857.html

Fig.14: Evolución del paro entre Febrero y Agosto de 2013



c) Vuelve a realizar los apartados a) y b) para la siguiente gráfica[10]*:*

Fuente: www.huffingtonpost.es

Un ejemplo bastante más obvio que el anterior, pero que tuvo cierto predicamento en los medios como ejemplo de cómo no se debe hacer un gráfico de barras. Si bien es pre-

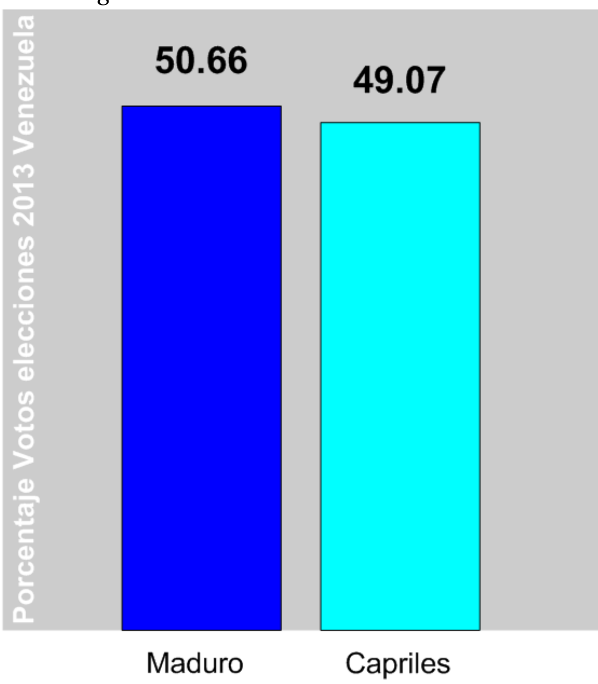
10 [Redacción] *Elecciones Venezuela 2013: El gráfico de la televisión oficial que favorece a Maduro (FOTOS)* [artículo en línea]15 Abril 2013 [consultado Junio 2015], en www.huffingtonpost.es (periódico digital) Disponible en: http://www.huffingtonpost.es/2013/04/15/elecciones-venezuela-2013-grafico-television_n_3083484.html

ciso que apuntemeos que no hemos encontrado esta imagen en la fuente primaria, en cuyo lugar aparece, en cambio, la siguiente imagen, mucho más ajustada[11]:

Fuente: www.vtv.gob.ve (Venezolana de Televisión)

En todo caso, la gráfica reformada quedaría de la siguiente forma:

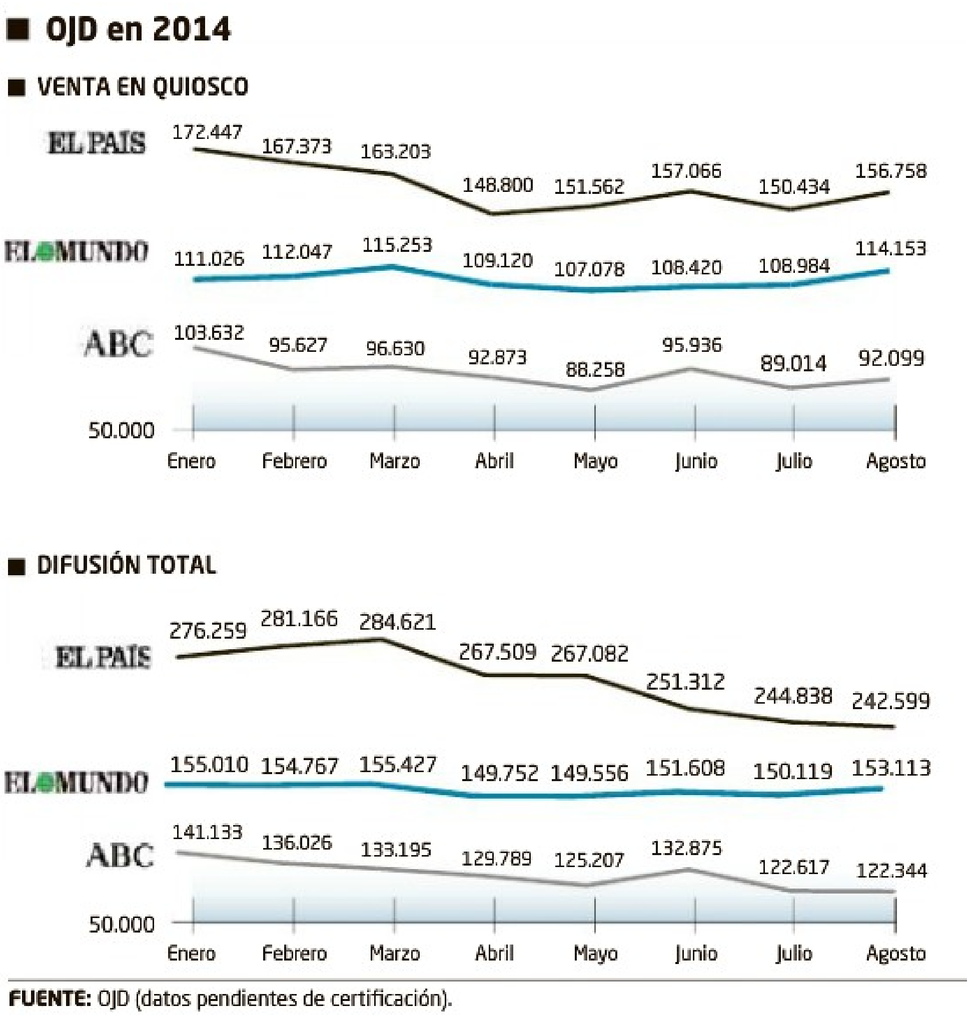
Fig.15: Resultados elecciones Venezuela 2013



11 Venezolana de Televisión, *Ganó el hijo de Chávez: Nicolás Maduro es el Presidente Electo de Venezuela (+Video)* [artículo en línea] en *www.vtv.gob.ve* (sitio en línea de Venezolana de Televisión) Disponible en [consultado Junio 2015]: http://www.vtv.gob.ve/articulos/2013/04/14/gano-el-hijo-de-chavez-nicolas-maduro-es-el-presidente-electo-de-venezuela-7215.html

2°) Observa atentamente los siguientes gráficos:

a)Comprueba si la información numérica concuerda con la información que suministran los gráficos[12]*:*



Fuente:www.malaprensa.com

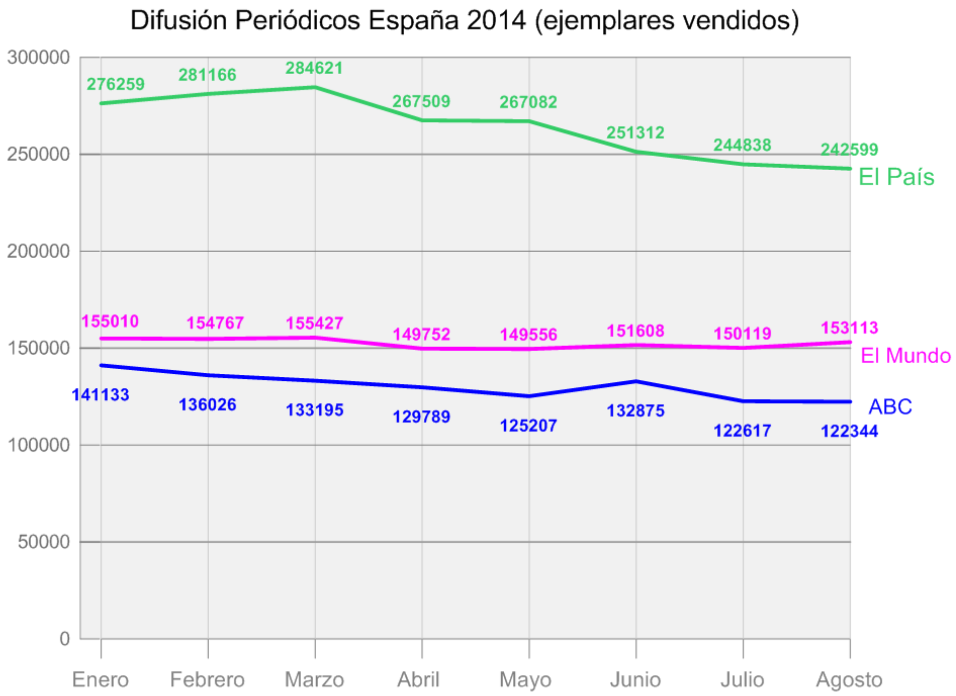
Resulta claro que ha habido por lo menos un truncamiento del eje de ordenadas (inexistente), puesto que las distancias respectivas no expresan adecuadamente las diferencias numéricas: así, en la primera gráfica, la diferencia de ventas en quiosco entre El Mundo y El País es numéricamente superior, y sin embargo la distancia entre las curvas es mucho más pequeña que la distancia entre ABC y el Mundo: exactamente lo mismo se puede decir de la segunda gráfica. De hecho, si se observa con mayor atención, hay otro truncamiento, ya que se puede apreciar que el eje de coordenadas comienza en 50.000 en ambos casos, y que la escala está a todas luces distorsionada, ya que la distancia entre el origen y la primera curva (ABC) no se corresponde tampoco en escala con la distancia entre esta curva y la inmediatamente superior (El Mundo) —por ejemplo, en la segunda gráfica, entre 50.000 y 141.133 hay una distancia similar que entre 141.133 y 155.010—. En conclusión la gráfica solo respeta el orden de las curvas de ventas.

12 Josu Mezo, *Otro Gráfico para la Historia* [artículo en línea], 25 Septiembre 2014 [consultado Junio 2015] en www.malaprensa.com Disponible en: http://www.malaprensa.com/2014/09/otro-grafico-para-la-historia.html

a)Representa correctamente ambas gráficas con un eje coordenado continuo. Sabiendo que la noticia original procedía de El Mundo, ¿Cuál crees que ha podido ser la intención del autor responsable de la gráfica?

Realizamos el ejemplo más claro, el de la segunda gráfica:

Fig. 16



Evidentemente, el autor de la gráfica pretendía disimular la diferencia entre las ventas totales de El Mundo y las de El País, presentándolas como similares, cuando lo cierto es que las ventas de El Mundo oscilan alrededor de un 60% de las de El País.

Escalas erróneas (o intencionadas) en pictogramas e iconos

A continuación expondremos ejemplos de distorsión escalar que afectan a otras formas de expresión gráfica estadística, cuales son los diversos tipos de pictogramas y otras representaciones icónicas[**44**]. Algunos de los casos denotan, más que interés, una falta de entendimiento de propiedades geométricas básicas como longitudes, áreas y volúmenes bastante preocupante por parte de los sedicentes “formadores de opinión”.

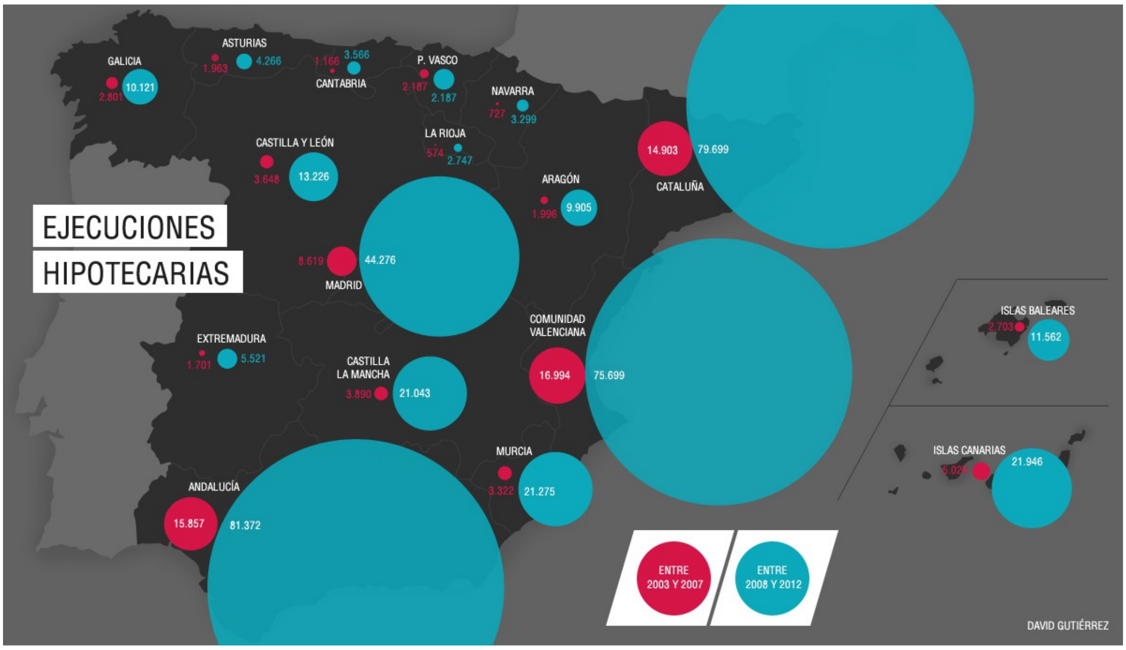
ACTIVIDADES

1º) Observa la siguiente gráfica[13] *y considera si la información visual que se ofrece se co-*

13 A. I. Gracia, *La crisis ha quintuplicado los desahucios en los juzgados en los últimos cuatro años,* [artículo en línea] 15 Marzo 2013 [consultado Junio 2015] El Confidencial (periódico digital www.elconfidencial.com) Disponible en: http://www.elconfidencial.com/espana/2013-03-15/la-crisis-ha-quintuplicado-los-desahucios-en-los-juzgados-en-

rresponde con la numérica. Caso de no ser así, indica cuál es la razón. ¿Se te ocurre alguna manera de hacer esa comprobación gráficamente?

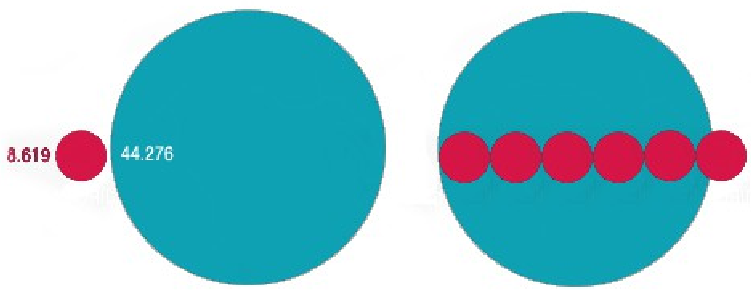
Fuente: *www.elconfidencial.com*



http://www.elconfidencial.com/archivos/ec/2013031484mapaespanahipotecas.jpg

Si se observan las cifras se comprobará que no están en la misma proporción que las áreas de los círculos a los que corresponden, sino a los diámetros (o radios) de los círculos en cuestión. Por lo tanto, los gráficos están construidos en base a los cuadrados de la cifras en cuestión, de manera que no representan la proporción entre los valors numéricos. De hecho, lo que sucede es que se sobreestima ampliamente el aumento de ejecuciones hipotecarias, lo cual puede comprobarse en el siguiente gráfico:

Fig. 17: Comparación de áreas



Donde se aprecia que la relación entre diámetros (radios) es 44276/8619= 5,14

2º)Supongamos que para el caso de Madrid el radio del círculo rojo es la unidad. ¿Cuál debería ser el radio del círculo azul para que la gráfica fuera realmente representativa de la proporción entre las cifras de ejecuciones hipotecarias?

---

los-ultimos-cuatro-anos_198383/

Se puede averiguar la cuestión con un sencillo cálculo que el infografista no quiso o no supo hacer:

$$\pi \cdot 1^2 = \frac{8619}{44276} \cdot \pi \cdot R^2 \longrightarrow R = \sqrt{\frac{44276}{8619}} = 2{,}2665 \neq 5{,}14$$

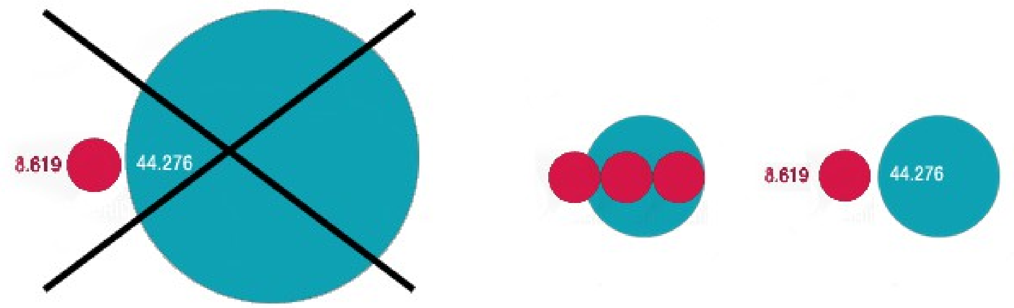


Fig. 19: Nueva comparación de áreas

2º) Observa la siguiente gráfica[14] *y considera si la información visual que se ofrece se corresponde con la numérica. Caso de no ser así, indica cuál es la razón.*

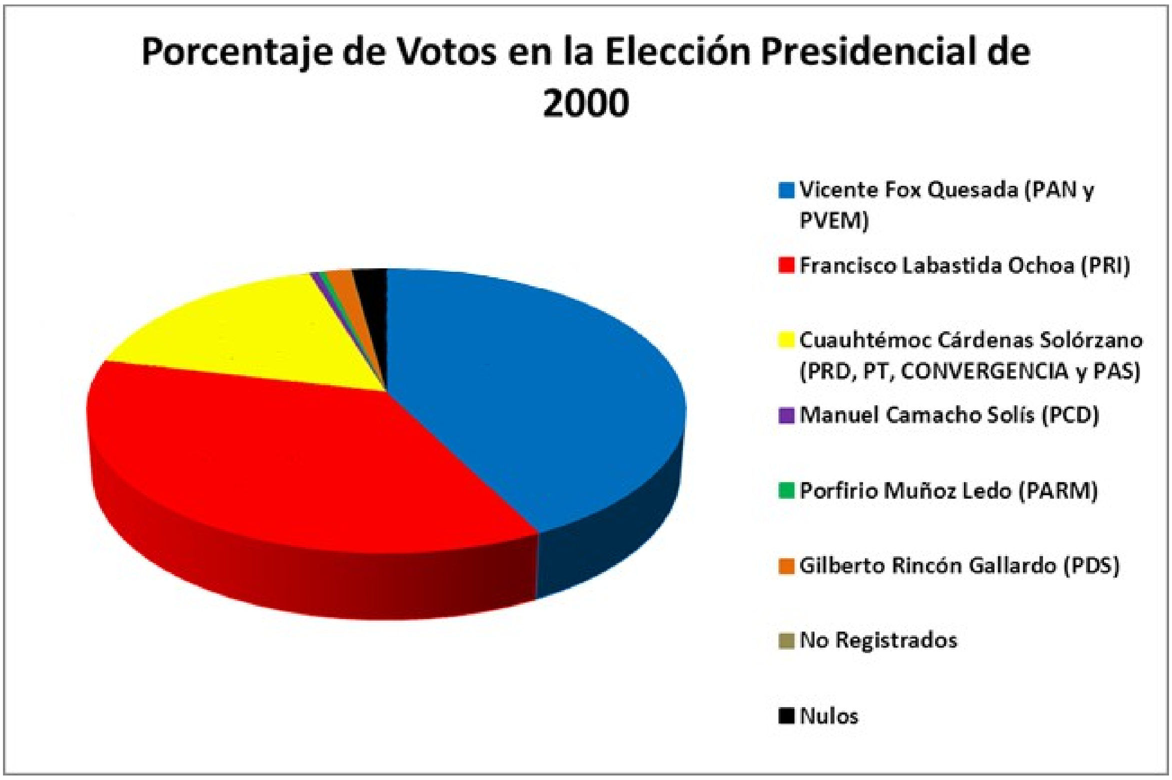
Fuente: www.marca.com

14 L. F. Rojo, *El fichaje más caro de la historia* [artículo en línea], 25 Febrero 2014 [consultado Junio 2015] Marca (edición digital *www.marca.com/*) Disponible en: http://www.marca.com/2014/02/25/futbol/equipos/barcelona/1393326271.html?a=87b03cc89eb1aebd4087a899b669d116&t=1393791124

En este caso aparecen varias distorsiones: primero, se supone que el origen de coordenadas es el centro del diagrama, pero los valores comienzan en un punto arbitrario. Y después, y más importante, no se sabe a qué es proporcional la cantidad reflejada por cada sector, pero no es ni al área ni al radio: parece que sencillamente se ha ido añadiendo un segmento al radio del sector para indicar que las cantidades van aumentando, pero el resultado final (la comparación entre el sector con la cifra 57 y el de 117,7) no tiene ningún sentido. O bien las áreas (o siquiera los radios) deberían ser proporcionales, pero no: tan sólo sirve para enfatizar un titular.

3°) Observa la siguiente gráfica[15]¿Sabrías decir en base a ella qué partido ganó las elecciones presidenciales mexicanas del 2000?

Fig. 20



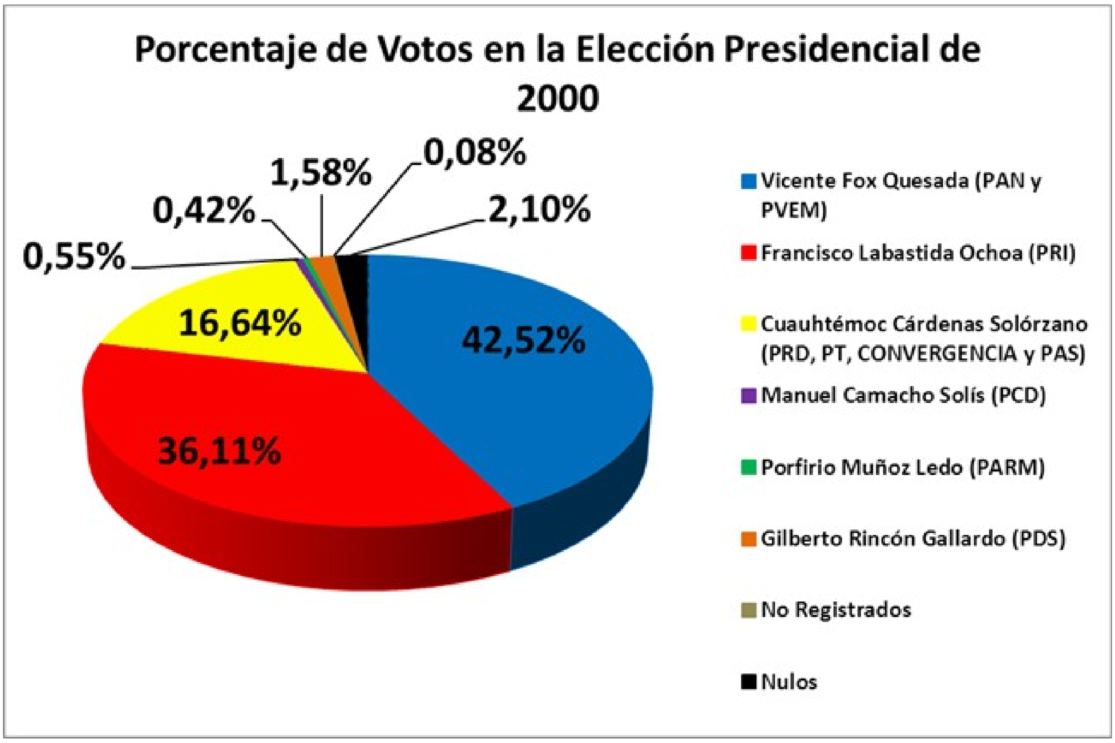
Presentamos aquí al alumno otro de los posibles equívocos provocados por una mala elección de representación gráfica: en este caso, los diagramas en forma de tarta, ya de por sí confusos(puesto que la diferencia angular se percibe peor que la longitudinal), pero en este caso agravado por el uso de la tercera dimensión y la perspectiva. Tufte **[44]** resulta tajante respecto de la utilzación de esta clase de gráficos:

> *Las tablas son preferibles a los gráficos para muchos conjuntos de datos pequeños. Una tabla es casi siempre mejor que un diagrama de tarta; la única cosa peor que un diagrama de tarta es que haya varios, puesto que entonces el espectador es forzado a comparar cantidad distorsionadas dentro y fuera de cada gráfico; dada su baja densidad de datos y su fracaso en ordenar números*

15 M.A.J. Mejía Montaño, *Partidos políticos mexicanos, IFE, tribunal electoral y jurídico, y elecciones presidenciales* [monografía en línea] en *www.monografias.com* Disponible en [consultado Junio 2015]: http://www.monografias.com/trabajos94/partidos-politicos-mexicanos-ife-tribunal-electoral-y-juridico-y-elecciones-presidenciales/partidos-politicos-mexicanos-ife-tribunal-electoral-y-juridico-y-elecciones-presidenciales2.shtml

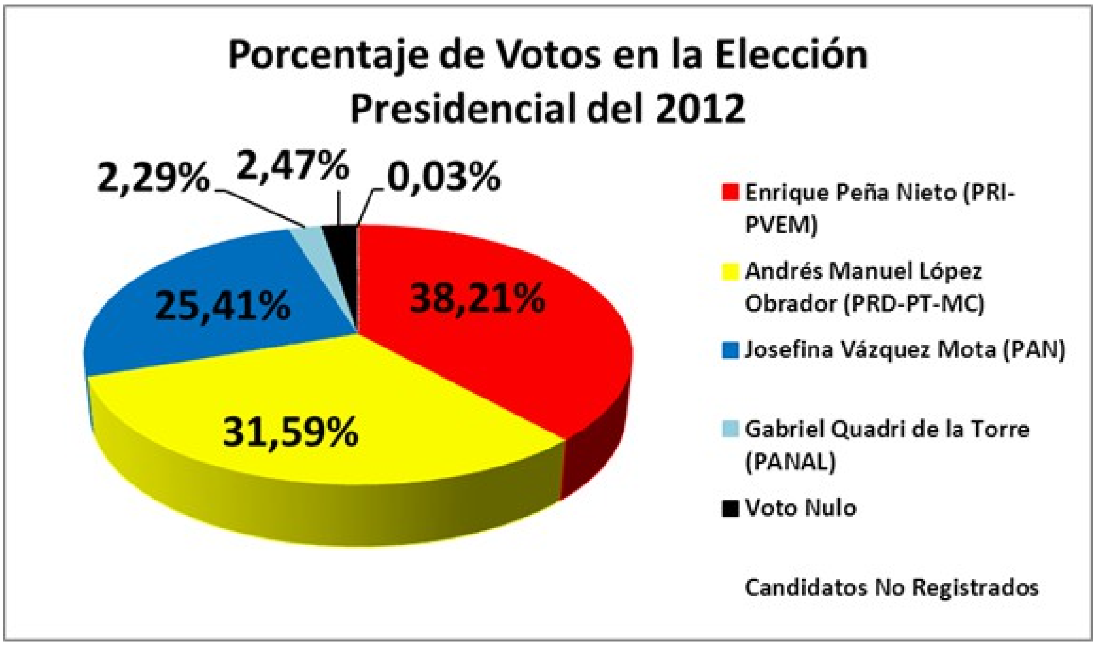
a lo largo de una dimensión visual, los diagramas de tarta nunca deberían utilizarse. [trad. propia]

En nuestro caso, el sector inferior aparece mucho más preponderante de lo que es en realidad, como se comprobará con la gráfica con los porcentajes:



M.A.J. Mejía Montaño, *Partidos políticos mexicanos, IFE, tribunal electoral y jurídico, y elecciones presidenciales*

Dos sectores (rojo y azul) que parecían equivalentes en área, representan porcentajes decisivos, pues dan la victoria (holgada) a Fox frente a Labastida. Lo mismo puede decirse del gráfico siguiente, perteneciente a la misma publicación:

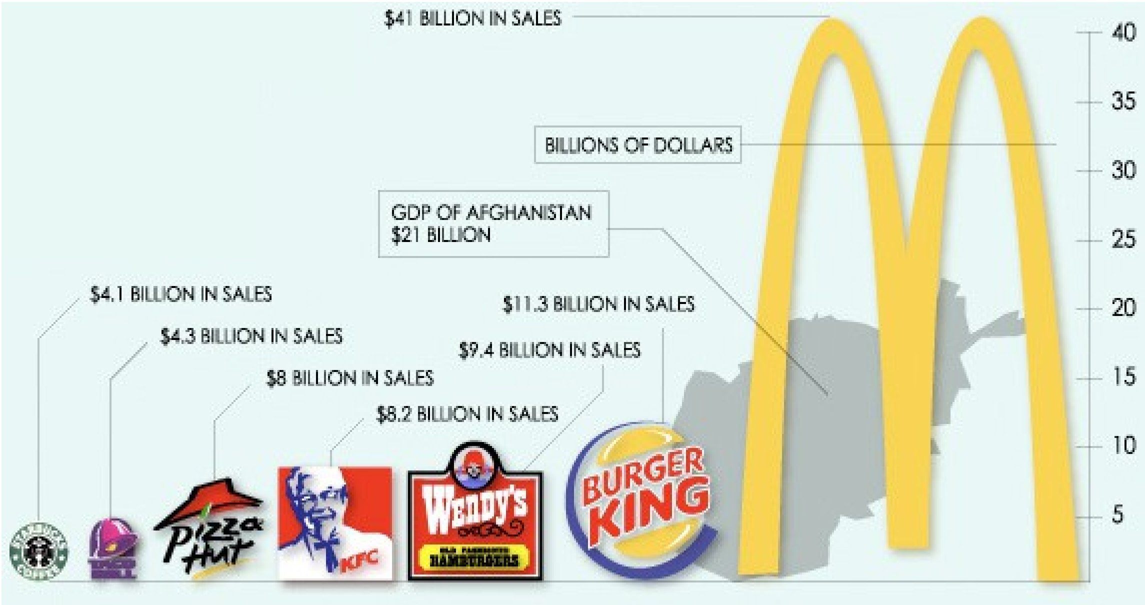


M.A.J. Mejía Montaño, Partidos políticos mexicanos, IFE, tribunal electoral y jurídico, y elecciones presidenciales

Donde se aprecia claramente que el sector amarillo, no sólo por ser el más cercano al espectador, sino por estar mayoritariamente situado en el eje largo de la elipse, aparenta mucha más área que el sector ubicado en su mayor parte en el eje corto: al ocupar mucha más área de dibujo parece que representa un porcentaje mayor que el sector rojo, contra lo que nos dicen las cifras respectivas.

4º) Observa la siguiente gráfica[16] *correspondiente a las ventas de cada marca en miles de millones de dólares(*$ billions) *y considera si la información visual que se ofrece se corresponde con la numérica. Caso de no ser así, indica cuál es la razón.*

Ventas mundiales de Restaurantes de comida rápida



Fuente: www.eagereyes.org

De forma equivalente al caso de las gráficas con círculos, aquí se produce una magnificación de los logotipos que no corresponde a la proporción entre cada conjunto de cifras, al haberse duplicado las dos dimensiones del icono, en lugar de una sola, lo cual habría dado áreas realmente proporcionales, pero habría distorsionado la apriencia de los logotipos (la solución óptima habría sido multiplicar ambas dimensiones por $\sqrt{2}$).Lo que se produce en el gráfico representa un incremento cuadrático del área total, y una impresión visual defectuosa, por tanto. El eje coordenado a la derecha de la imagen no deja lugar a dudas sobre la operación que se ha llevado a cabo.

Así, el icono de McDonalds, que debería guardar una relación de área con el de Burger King de 41/11,3 = 3,63, en realidad se encuentra en una relación de $3{,}63^2$= 13,16

4º) Observa la siguiente gráfica correspondiente a una comparación entre la vieja y la nueva batería para iPad, teniendo la nueva un 70% más de capacidad que la vieja[17]*.*

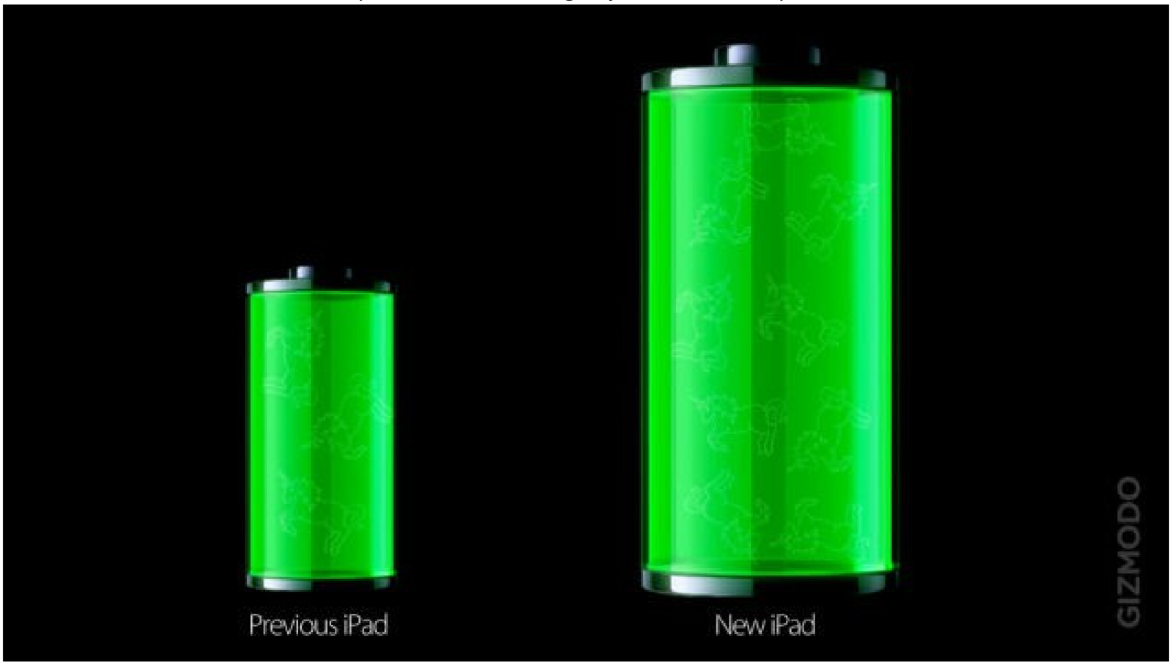
16 R. Kosara, *Linear vs. Quadratic Change* [artículo en línea] 19 Septiembre 2008 [consultado Junio 2015], en *www.eagereyes.org* Disponible en: https://eagereyes.org/blog/2008/linear-vs-quadratic-change

17 J. Díaz, *Holy F*ck, the New iPad Has a Gigantic 70-Percent Larger Battery* [artículo en línea] 15 Marzo 2012 [consultado Junio 2015] en *www.gizmodo.com* Disponible en: http://gizmodo.com/5893738/holy-fck-the-new-ipad-has-a-gigantic-70-percent-larger-battery

Cconsidera si la información visual que se ofrece se corresponde con la numérica. Caso de no ser así, indica cuál es la razón.

El efecto anterior se agrava más aún cuando se comparan dos iconos donde se hallan representadas las tres dimensiones.

Comparación entre antigua y nueva batería para iPad



Fuente: www.gizmodo.com

Aquí la diferencia del 70% debería implicar una relación igual entre los volúmenes de las figuras. En cambio, lo que se observa es que la altura del icono correspondiente a la nueva batería es 1,7 veces la del viejo; por lo tanto el volumen de la nueva es $1{,}7^3$= 4.913 veces el de la antigua; es decir, no un 170% más, sino un 491,3% más de capacidad.

Los maestros también se equivocan

Probablemente hay pocas cosas más atractivas para un alumno de enseñanza secundaria, cuya existencia transcurre en gran medida bajo el escrutinio y posterior juicio de su conducta por parte de sus mayores (padres, profesores, psicólogos, etc.) que comprobar —y señalar con el dedo, usualmente— los errores de estos últimos, algunas veces llegando a resultar impertinentes. No sólo es rebeldía adolescente. En cierta manera, es un alivio para él: por un momento descansa del hecho der ser él quien es evaluado (y por lo general, también, de estar equivocado, al menos en el área de las matemáticas). Por esta razón consideramos que los ejemplos que se expondrán pueden resultarle atractivos, pero también instructivos, en varios aspectos: el primero, tal y como lo formula Chung **[9]** —en uno de los más *verdaderamente* didácticos textos introductorios a la teoría de la probabilidad que conozcamos— de manera muy sencilla[1]:

> *Pascal no intentó explicar su método[...]. Una carta suya posterior parece indicar que volvió a enredarse en este mismo punto al tratar el problema para tres jugadores.* ***El principiante no debería pues descorazonarse tan pronto: incluso a los grandes maestros del pasado no les resultaban sencillos los razonamientos de este tipo.***

Sobre el problema del que se habla se sabrá en pocos párrafos. Ahora indicaremos en qué otro sentido puede ser instructivo, dado que se propone una cuestión de fondo interesante por medio de cada uno de los ejemplos que consideraremos a continuación. Una cuestión que puede formularse —y se ha formulado— así: ¿Por qué la teoría de la probabilidad tuvo un desarrollo relativamente tardío en relación a otros campos matemáticos? Se han propuesto varias respuestas a este interrogante (Gigerenzer *et al.* hacen un buen resumen al comienzo de **[20]**), mas la que nos interesa en el texto presente es uno de sus posibles corolarios, es decir, la que tiene que ver con los aspectos cognitivos del razonamiento probabilístico (estadístico); en otras palabras, la que se expresa mediante la pregunta: ¿Es de alguna manera el pensamiento aplicado a la probabilidad distinto al de otros campos matemáticos? (en orden a responder otra más restringida suscitada por el presente capítulo: ¿Porqué inteligencias eminentes, a veces destacados matemáticos, se han equivocado a lo largo de la historia en cuestiones de probabilidad y estadística?) Las fuentes históricas parecen decantarse todas hacia la vieja, y aún no extinta, controversia entre necesidad (o determinismo) y azar (o casualidad), señalando que tanto los geómetras del mundo antiguo, como los pioneros de las nuevas álgebra y análisis de los siglos XVI y XVII, obtenían sus conclusiones basándose en un modo de pensar determinista, heredado de la geometría sintética, donde los resultados deducidos de teoremas y axiomas (con los cuales se describían los fenómenos de los que la ciencia se ocupaba) eran *necesariamente exactos*, un modo de pensar ajeno a la existencia de lo aleatorio, que no dejaba espacio para la variabilidad (que no está menos sometido, de todas formas, a sus propias leyes) **[23]** :

1 Luego de tratar dos de los problemas que más adelante expondremos; si bien parece no estar muy al tanto —o no muy interesado— en el detalle histórico de algunos asuntos, como la atribución correcta de la resolución del problema expuesto a Pascal, en lugar de a Fermat.

> *A lo largo de la Edad de la Razón, el azar fue considerado superstición del vulgo. Azar, superstición, vulgaridad e irrazón eran de la misma clase.* ***El hombre racional, apartando sus ojos de tales cosas, podía cubrir el caos con un velo de leyes inexorable.*** *El mundo, se decía, puede a veces parecer azaroso, pero sólo porque no conocemos el inevitable funcionamiento de sus fuerzas internas. En cuanto a la probabilidad (a las que los matemáticos llamaban la doctrina del azar) no eran más que la defectuosa pero necesaria herramienta de la gente que sabía demasiado poco.*[...]
> *Pero la erosión del determinismo no es la creación de desroden e ignorancia, sino al contrario […]*
> ***Hay una paradoja aparente: Cuanto mayor la indeterminación, mayor el control****. Esto es obvio en las ciencias físicas. La física cuántica da por supuesto que la naturaleza es en la base irreductiblemente estocástica. Pero precisamente ese descubrimiento ha mejorado inconmensurablemente nuestra capacidad de interferir con y de alterar el curso de la naturaleza.*[...]
> ***El determinismo fue subvertido por las leyes del azar****.* [traducción propia]

y también en **[30]** :

> [...]*Parece que nos vemos arrastrados a la conclusión de que el tardío surgimiento del cálculo de probabilidades es debido a algún factor más fundamental: la misma noción de azar en sí mismo, la idea de ley natural, la posibilidad de que una proposición puede ser cierta y falsa en determinadas proporciones fijas, todos esos conceptos son hoy en día en tan gran medida parte de nuestras rutinas comunes de pensamiento que quizá olvidamos que no lo eran para nuestros antepasados*[...]
> [...] ***Parece haberle costado a la humanidad varios cientos de años acostumbrarse a un mundo en el que algunos eventos sucedían sin causa; o en el que, al menos, extensas áreas de sucesos eran determinados por una causalidad tan remota que podían ser representados adecuadamente por un modelo no causal****.* [traducción propia]

Ambos textos atribuyen el fenómeno a causas históricas que determinan los modos de pensamiento: pertenecen al estudio de la hoy denominada “historia de las mentalidades”. No obstante, señaladas fuentes de la psicología (cognitiva) y de la didáctica de la matemática, como ya hemos visto **[29]** indican que las dificultades de aprendizaje en relación a la probabilidad y estadística son o bien intrínsecas a la psique humana (cognitivas —y no exclusivas del período de las enseñanzas medias o superiores) o bien propias de la misma materia tomada como dominio particular de conocimiento (epistemológicas). Apoyándonos en esa misma tesis, intentaremos, salvando las distancias, mostrar al alumno que sus (posibles) errores no son una noveda histórica, sino que, a veces, los comparte con ilustres figuras de la ciencia .

El Caso Leibniz

Filósofo, matemático, y figura preponderante en ambos dominios, coinventor (según algunos, el verdadero inventor) del cálculo infinitesimal **[31]**, tan conocido e influyente como para que, por proponer un ejemplo ínfimo, todavía hoy su notación para las derivadas en forma de cocientes de diferenciales y la “s larga” para las integrales sea la más usada, comete, no obstante, el siguiente error de 1º de la ESO, cuando entra a considerar el cálculo de probabilidades en sus *Opera Omnia* publicadas en 1768 (tomado de **[22]**, cap.15):

> […] *Por ejemplo, con dos dados,* ***es igualmente probable obtener doce puntos en una tirada que once****; porque ambas pueden conseguirse solamente de una manera*[...]

Por supuesto, ni qué decir tiene que hay dos posibles combinaciones de tiradas de dados (6,5) y (5,6) cuyo resultado es “once”, y sólo una (6,6) con la que se obtiene “doce”: por lo cual la probabilidad de sacar "once" es el doble que la de sacar "doce"; un rasgo típico del principiante en combinatoria y probabilidad, no definir bien el espacio de sucesos, por tener en cuenta solo el resultado final de las tiradas y no saber pasar por alto la "falacia conmutativa" (el término es propio) consistente en el siguiente razonamiento: si 6+5 es equivalente a 5+6, ambas son solo expresiones de una sola y misma operación aritmética, luego deben contarse solo una vez. Si bien es preciso señalar, en descargo del errado, que el cálculo de probabilidades estaba por aquél entonces en pañales, o más bien era un nonato aún (y consistía en poco más que fragmentos de correspondencia entre intelectuales del siglo, Pascal, Fermat, y el caballero de Mère, Huygens, Cavalieri, y más tarde entre los Bernouilli y Leibniz, o entre Newton y Samuel Pepys; El *Ludo Aleae* de Cardano no era muy conocido **[20]** **[43]**) y que la definición clásica de probabilidad no llegaría a formularse hasta la publicación de la *Théorie Analytique des Probabilités* de Laplace (1812). Todhunter (**[43]**, p. 48), no obstante, comenta:

G. W. Leibniz, (1646-1716) C.B.Francke

> *Leibniz nos poporciona así un ejemplo de la propensión al error que parece caracterizar nuestro asunto.* [trad. propia]

El Caso D'Alembert

¿Cómo fue posible que Jean Le Rond D'Alembert, uno de los padres de la física y el análisis matemático (**[25]** pp. 81-84), contribuyente decisivo al cálculo de derivadas parciales y de variaciones, a los principios básicos de sistemas mecánicos, miembro de la Academia de las Ciencias de Francia con solo veinticuatro años, autor del *Traité de Dynamique,* del *Traité de l'équilibre et du mouvement des fluides,* de investigaciones sobre la precesión de los equinoccios y el movimiento de nutación terrestre, de la primera (aunque fallida) demostración del Teorema Fundamental del Álgebra (los autores franceses a veces lo denominan *Terorema de D'Alembert,* pese a ello), de la primera reseña acerca del concepto del límite como clave del cálculo infinitesimal**[32]**, ¿Cómo fue posible, como decimos, que se equivocara de una manera tan, diríamos, pueril, y no una, sino *dos veces* al enfrentarse a nociones básicas del cálculo de probabilidades?

Jean D'Alembert, M. Quentin de la Tour

• **Cara o Cruz:** D'Alembert, además de todo lo reflejado en el párrafo anterior, fue principal colaborador de Diderot en su celebérrima Enciclopedia *(Encyclopédie ou Dictionnaire Raisonné des Sciences, des Arts et des Métiers)*, en la cual escribió el “famoso” (en los anales de la estadística) artículo titulado *Croix ou Pile* (Cara o Cruz), que extractamos a continuación (**[22]**, cap. 12):

> […] *Uno se pregunta cuántas probabilidades tiene de sacar* [al menos] *una cara en dos tiradas sucesivas. La respuesta que se encontrará en todos los autores, siguiendo los principios comunes, es la que sigue: Hay cuatro combinaciones:*
>
> *Primera tirada. Segunda tirada*
>
> *Cara. Cara.*
>
> *Cruz. Cara.*
>
> *Cara. Cruz.*
>
> *Cruz. Cruz.*
>
> *De estas cuatro combinaciones una sola pierde, y tres ganan; las probabilidades son de tres contra una*[1]*a favor de quien apuesta sacar una cara.*[...] *De todas formas, ¿Es esto correcto?* […] *¿No es necesario reducir a una sola las dos combinaciones que muestran una cara en la primera tirada?Porque tan pronto como aparece una sola cara, el juego ha terminado y las segundas tiradas no cuentan para nada. Así que hay solamente tres combinaciones:*
>
> *Cara, primera tirada.*
> *Cruz, primera tirada. Cara, segunda tirada.*
> *Cruz, primera tirada. Cruz, segunda tirada.*
>
> *De manera que las probabilidades son de dos contra una a favor del jugador*[...]

De lo que se olvida D'Alembert es que las combinaciones alternativas que expone no tienen las mismas probabilidades de suceder, es decir, *ya no son equiprobables*, y por lo tanto su suma directa lleva a error, como se verá con más detalle cuando acometamos las actividades[2].

• **La falacia del jugador:** En sus *Opuscules Mathematiques*, D'Alembert vuelve a errar en consideración a uno de los más clásicos y extendidos mitos de los juegos de azar (tomado de **[22]**, cap. 13):

1 Es decir, de tres entre cuatro (3/4). D'Alembert utiliza una convención algo difrente a la formalización actual de probabilidad, y que todavía se usa en el entorno de las apuestas organizadas y los jugadores profesionales.

2 Para no resultar redundantes, y porque su exposición requeriría mucho espacio, no exponemos el más famoso, pero a todas luces equivalente **problema de los puntos**, hallado en la correspondencia entre **Pascal** y **Fermat** acerca de una de las consultas sobre juegos de azar del **caballero de Mère,** y donde se encuentra mención a una objeción por parte de **Roberval** (que sería el maestro equivocado en cuestión: fue catedrático de matemáticas del *College Royal* hasta su muerte) sobre la solución del problema dada por Fermat (pero expuesta por Pascal) que es exactamente la misma que la de **D'Alembert** sobre la presunta finalización prematura del juego (vid. **[22]**, cap. 4, p. 30). Por cierto, más adelante, **Pascal mismo se equivoca** cuando trata de generalizar el problema para tres jugadores.

> […] *Si un evento ocurre varias veces sucesivamente, por ejemplo, en el juego de cara o cruz, "cara" ha sucedido tres veces seguidas, ¿Es igualmente probable que "cara" o "cruz" sucedan a la cuarta tirada?*[...] *Porque después de todo no es posible, es incluso físicamente imposible que "cruz" nunca suceda.* ***Por lo tanto, cuantas más veces sucede "cara", más probable es que "cruz" suceda la próxima vez*** […]

Cualquiera reconocerá aquí la famosa "falacia del jugador", y las, a veces, serias dificultades con las que se enfrenta el enseñante a la hora de romper el extremadamente fuerte prejuicio que la sostiene. De hecho, el mismo D'Alembert insiste en el tópico en su posterior artículo *Doutes et Questions sur le Calcul des Probabilités*([**22**], cap. 13):

> *El cálculo de probabilidades se basa en la asunción de que todas las diversas combinaciones del mismo efecto son igualmente probables. Por ejemplo, si lanzamos una moneda al aire 100 veces consecutivas,* ***se asume que es igualmente probable que "cara"ocurra cien veces a que las "caras" y las "cruces" estén mezcladas*** […] *Los dos casos sin duda son,* ***matemáticamente****, igualmente posibles. Pero esta no es la verdadera dificultad, y los matemáticos mediocres de los que hablaba antes han acometido inútiles esfuerzos en escribir disertaciones larguísimas para probar esta igual posibilidad. Pero uno se debe preguntar si de los dos casos, que son* ***matemáticamente*** *igual de probables, lo son también* ***físicamente****, teniendo en cuenta el orden natural de las cosas.*

Del párrafo en cuestión pueden deducirse más cosas que el mero desliz de su autor. En primer lugar, la consideración de que a la fecha en que fueron escritos ambos artículos (mediados del s. XVII) la física y las matemáticas se hallaban tan interpenetradas que no cabía apenas hacer distinción entre ellas: aquí D'Alembert recurre a argumentos físicos para rebatir cuestiones matemáticas, una forma de presentar la argumentión que un siglo después ya sería inaceptable, pero que en tal momento no daba lugar a interpretación: tan imbricada estaba la idea de que las matemáticas debían seguir a la realidad física. En segundo lugar, da una demostración indirecta de lo importante que es, a fin de cuentas, la formalización precisa de los elementos constituyentes de teorías matemáticas: en la discusión se entremezclan sin distinción las perspectivas incompatibles de la probabilidad objetiva (frecuentista) y subjetiva(bayesiana), y la confusión entre probabilidad y esperanza matemática.

Es de notar el tono polémico utilizado por D'Alembert: los dos artículos no están escritos tanto para demostrar sus tesis como para rebatir algunas conclusiones de la incipiente teoría de las probabilidades, y, por ende, negar su carácter de teoría realmente científica: abundando en lo expuesto en nuestra introducción, los textos reflejan la tremenda resistencia que las mentalidades, incluso aquellas tan indudablemente dotadas y científicamente orientadas como las de D'Alembert, oponían a la existencia de una ciencia de lo aleatorio, que implicase enunciados no deterministas sobre la naturaleza de las cosas, y que, además de explicar en parte el retraso de la formalización de la teoría de la probabilidad, puede ser debida a las distintas capacidades intelectuales involucradas en ambos aspectos de la reflexión matemática (vid. p. 104), diferencia que debería ser contemplada a la hora de plantear la enseñanza de las materias de probabilidad y estadística de una forma propia y particular, no exactamente igual que otras áreas del currículo matemático en enseñanza secundaria.

ACTIVIDADES

1º)Realiza una búsqueda rápida en la red sobre Gottfried Wilhelm von Leibniz [http://es.wikipedia.org/wiki/Gottfried_Leibniz bastaría para nuestros fines]. *Haz un breve resumen de sus logros matemáticos (no hace falta que entiendas del todo los conceptos implicados).*

2º)La autoría del texto siguiente corresponde al matemático que has estudiado en la actividad 1º)[vid. p. 104]*:*

> ***Con dos dados, ES IGUALMENTE PROBABLE OBTENER DOCE PUNTOS EN UNA TIRADA QUE ONCE; porque ambas puntuaciones pueden conseguirse solamente de una manera; sin embargo es TRES VECES más probable sacar siete; porque ello puede suceder de tres maneras: un seis y un uno, un cinco y un dos, y un cuatro y un tres; y cada combinación es tan probable como la otra.***

Teniendo el texto a la vista, contesta a la siguiente pregunta: ¿Es igualmente probable sacar 12 puntos en dos tiradas de dados que 11? ¿Por qué? Recuerda que tú eres un alumno de secundaria y Leibniz fue "El último Genio Universal".

Hay seis resultados posibles de lanzar un dado, que configuran el llamado Espacio de Sucesos (con 6 elementos)$\Omega=\{1,2,3,4,5,6\}$;los sucesos posibles resultantes de realizar dos lanzamientos de dados serán el resultado de multiplicar por 6 cada posibilidad anterior: Si saco un 1 en la primera tirada, en la segunda tirada tengo seis posibilidades distintas que combinar con ese resultado: 11, 12, 13... Igualmente sucede con 2, 3, 4, 5, y 6: luego el espacio de sucesos en caso de dos tiradas tendrá $6\cdot 6=6^2=36$ elementos:

$$\Omega=\begin{Bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{Bmatrix}$$

De los cuales sólo uno, (6,6), da como resultado 6+6=12; en cambio, dos de los elementos, 56 y 65 dan como resultado 5+6=6+5=11; De tal manera que P(12)=P[(6,6)]=1/36; mientras que P(11)= P[(5,6)]+P[(6,5)]= (1/36)+(1/36)=1/18 →P(11)=2·P(12)

Se podría especular sobre las razones por las que se equivocó Leibniz (y que, seguramente, serán parecidas a las razones por las que se equivocaría un alumno): o bien consideró el asunto sólo desde un punto de vista aritmético y llegó a la conclusión de que si 5+6=6+5, es decir, que si por la propiedad conmutativa de la suma de naturales la posición de los sumandos es intercambiable, entonces **ambas expresiones representan una sola operación***;* o bien que, en lugar de considerar dos tiradas sucesivas en el tiempo, creyó que los dados se tiraban al mismo tiempo, y esto hacía irrelevante el orden en que fueran sumados los dados: efectivamente, cuando uno representa los eventos como sucesivos en el tiempo, parece claro que, por cada resultado de la primera tirada (y me guardo el dado en el bolsillo) hay otras seis posibilidades distintas en la segunda tirada.

Para Leibniz en este caso está claro que los dos dados son idénticos: pero idénticos no significa indistinguibles, no significa *que sean el mismo dado*: luego importa que uno de ellos saque un 5 y el otro un 6, y viceversa. Esto se entiende mejor si, aun suponiendo que los dados se tiran al mismo tiempo, **uno de ellos es rojo y el otro, azul.** Es evidente que las probabilidades serían las mismas aunque los dados fueran del mismo color (y no hubiera manera de saber cuál es cual tras una tirada): pero, habiéndolos pintado, queda claro que el hecho de que el dado rojo saque un 6 y el azul un 5 **no es el mismo suceso** que el hecho de que el dado azul saque un 6 y el rojo un 5. En cambio, queda claro que solo hay un suceso (una posible combinación) en el cual el dado rojo **y** el azul sacan ambos un 6.

3º)Imagina que el pasaje anterior de Leibniz aparece en el libro de texto (o en un solucionario que hayas conseguido obtener) ¿Te preguntarías si lo que dice es cierto? ¿Buscarías algún otro texto para contrastarlo? ¿O tratarías de resolverlo tú mismo?

La pregunta es, por supuesto, algo retórica y un tanto injusta con el alumno. Lo mismo un estudiante avanzado de matemáticas, sobre un tema del que no esté muy seguro sin duda tendería a hacer caso a la autoridad de un libro de texto, aunque solo fuera por razones de economía de tiempo o de esfuerzo. No obstante, lo que se busca con esta pregunta no es avergonzar al alumno, sino indicarle que la autoridad (en matemáticas o en cualquier otro ámbito) no tiene la razón por ser la autoridad; sucede al contrario, la autoridad es autoridad *porque* (casi siempre) *tiene razón.* Y no solo eso: en matemáticas, tiene razón porque lo *demuestra*: no sólo porque exhibe un catálogo de aciertos más o menos extenso.

4º)Supongamos que el solucionario está equivocado. ¿Lo tirarías inmediatamente a la basura por inútil? Asimismo, ¿Crees que el cálculo infinitesimal quedaría derogado si se descubriera que Leibniz se equivocó en el texto anterior? Busca en la red noticias de alguien que no se haya equivocado nunca (exceptuando el Papa de Roma en materia de doctrina cristiana)

Asimismo con esta pregunta se busca derogar, en parte, el mito de la infalibilidad, tanto del profesor, de los libros de texto, como de la ciencia en general. El error acompaña cualquier actividad humana, incluida la ciencia. Pero la ciencia está formada por enunciados que, o bien son perfectamente demostrables en base axiomas, como la matemática, o bien son contrastables en base a determinados fenómenos bien localizados y definidos (al menos, en último extremo). Debido a lo cual la ciencia, incluida la matemática[3], está en constante estado de evolución y constatemente sometida a crítica: siempre es mejorable, o perfectible si se quiere, porque, a diferencia del conocimiento ordinario, cuenta con criterios de verdad bien definidos (o, en todo caso, infinitamente más precisos que cualquier otra clase de discurso). Empleando estos criterios de verdad puede distinguirse entre enunciados verdaderos, erróneos, sin senido o irrelevantes; en todo caso, nunca es la autoridad quien los sostiene. De esta manera la ciencia puede soportar ocasionalemnte el error, porque es capaz de detectarlo y de corregirse a

3 La preponderancia del idioma inglés en textos científicos obliga casi a realizar esta distinción, natural en ese idioma, entre *mathematics* y *science,* que no existe en castellano y que puede ser objeto de discusión.

sí misma continua y consecuentemente siguiendo criterios válidos. En cambio, a la autoridad externa, injustificada, le basta equivocarse una vez para verse derogada: tras lo cual solo dispone de otros medios, no intelectuales, para hacers e respetar.

5º)Realiza una búsqueda rápida en la red sobre Jean Le Ronde D'Alembert [*http://es.wikipedia.org/wiki/Jean_le_Rond_d%27Alembert* bastaría para nuestros fines]. *Haz un breve resumen de sus logros matemáticos (no hace falta que entiendas del todo los conceptos implicados).*

6º)La autoría del texto siguiente corresponde al matemático que has estudiado en la actividad 5º)[vid. p. 106]*:*

[...] *Se pregunta cuántas probabilidades se tiene de sacar una cara en dos tiradas de monedas sucesivas. La respuesta que se encontrará en todos los autores, siguiendo los principios comunes, es la que sigue: Hay cuatro combinaciones:*

Primera tirada. Segunda tirada

Cara. Cara.

Cruz. Cara.

Cara. Cruz.

Cruz. Cruz.

De estas cuatro combinaciones una sola pierde, y tres ganan; las probabilidades son de tres entre cuatro a favor del jugador [...] *De todas formas, ¿Es esto correcto?* [...] *¿No es necesario reducir a una sola las dos combinaciones que muestran una cara en la primera tirada? Porque tan pronto como aparece una sola cara, el juego ha terminado y las segundas tiradas no cuentan para nada. Así que hay solamente tres combinaciones:*

Cara, primera tirada.
Cruz, primera tirada. Cara, segunda tirada.
Cruz, primera tirada. Cruz, segunda tirada.

De manera que las probabilidades son de dos entre tres a favor del jugador [...]

Teniendo el texto a la vista, contesta a la siguiente pregunta: ¿Cuál es la probabilidad de sacar al menos una cara en dos tiradas de una moneda? ¿Por qué? Analiza la objeción planteada por D'Alembert: ¿Es correcta? Recuerda que tú eres un alumno de secundaria y D'Alembert es una celebridad matemática.

Efectivamente, el método reseñado por D'Alembert al comienzo del texto es correcto, y está bien explicado. Su objeción es, en cambio, errónea, porque los tres sucesos que se mencionan **no son igualmente probables**, de manera que sus probabilidades respectivas no pueden sumarse la manera en que D'Alembert lo hace; comprobémoslo:

· Es evidente que la probabilidad del primer suceso, “cara en la primera tirada” es la misma que la de sacar sencillamente “cara”: luego su probabilidad es $P(C1)=P(C)=1/2$
· En cambio, la probabilidad del suceso segundo ya no es 1/2; porque el segundo suceso está compuesto de la ocurrencia de dos sucesos simples, uno en cada tirada, “sacar cruz en la primera tirada” y “sacar cara en la segunda tirada”, cada uno de los cuales tiene la misma probabilidad, 1/2; no obstante, su combinación, asumiendo que son independientes (es decir, que el resultado de una tirada no influye en el resultado de la otra, como resulta evidente) tiene de probabilidad $P(Cr1 \cap C2)= P(Cr1)\cdot P(C2)= (1/2)\cdot(1/2)=1/4$; la probabilidad conjunta de los tres sucesos que se indican es, por tanto, $(1/2)+(1/4)+(1/4)=$ **3/4**, no 2/3.

También es posible considerarlo de otra manera: si observamos el espacio de sucesos definido correctamente por D'Alembert al comienzo del texto,

Primera tirada. Segunda tirada

Cara. Cara.

Cruz. Cara

Cara. Cruz.

Cruz. Cruz.

podemos concluir que cada suceso expresado por cada una de las líneas tiene de probabilidad una entre cuatro, ¼. Ahora bien, el suceso “sacar cara en la primera tirada” tal y como se expone en la objeción, **no es un sólo suceso de esta tabla**, sino que incluye dos, a saber: “Cara. Cara” y “Cara. Cruz”: luego su probabilidad es de $(1/4)+(1/4)=1/2$; luego la probabilidad conjunta de los tres sucesos que D'Alemebert señala en la objeción es, al igual que antes, $(1/2)+(1/4)+(1/4)=$ **3/4**, no 2/3.

7º)La autoría del texto siguiente corresponde al matemático que has estudiado en la actividad 6º)[vid. p. 107]*:*

[...] ***Si un evento ocurre varias veces sucesivamente, por ejemplo, en el juego de cara o cruz, “cara” ha sucedido tres veces seguidas, ¿Es igualmente probable que “cara” o “cruz” sucedan a la cuarta tirada?***[...] ***Porque después de todo no es posible, es incluso físicamente imposible que “cruz” nunca suceda. Por lo tanto, cuantas más veces sucede “cara”, más probable es que “cruz” suceda la próxima vez*** [...]

Teniendo el texto a la vista, contesta a la siguiente pregunta: ¿La probabilidad de que, tras conseguir sacar cuatro caras seguidas en tiradas consecutivas de una moneda, en la siguiente tirada salga cruz, es mayor que la de que salga otra vez cara? ¿Cuál es la probabilidad exacta? ¿Por qué?

Es claro que la probabilidad de sacar cara en una tirada cualquiera, no importa el número de veces que se haya lanzado, sigue siendo de 1/2; siguiendo el ejemplo de D'Alembert, la combinación CCCC= Cara, Cara, Cara, Cara es exactamente igual de probable que la combinación CCCX= Cara, Cara, Cara, Cruz **en ese orden preciso**. D'Alembert parece confundir sucesos cuando elabora su tesis: CCCX es un solo suceso (compuesto); en cambio, **que salga cruz alguna de las veces**, que es sobre lo que el autor especula, **no es un solo suceso**, sino cuatro, a saber: XCCC, CXCC, CCXC, y CCCX. siendo así, efectivamente, es más probable que salga una (sola) cruz en cuatro tiradas a que no salga en ninguna. Pero no es así como el fenómeno que se discute está definido.

Se hace necesario aquí explicar el concepto de sucesos independientes. Cuando lanzamos una moneda, la recogemos y la volvemos a lanzar, lo que haya resultado en la primera tirada no tiene **ninguna influencia** sobre las características físicas de la moneda, ni sobre las condiciones del entorno en que se practicó el lanzamiento. De manera que lo que resultara en el primer lanzamiento no tiene **ninguna influencia** sobre lo que resulte en el segundo lanzamiento. Ambos sucesos son, por tanto, independientes.

Se puede expresar lo mismo informalmente diciendo que la moneda "no tiene memoria". Es evidente que, aunque yo consiguiera, lanzando una moneda el número de veces necesario, obtener una "racha" de10 caras, por ejemplo, yo no podría entonces guardarme esa moneda en el bolsillo a la espera de encontrarme con algún incauto en un futuro con quien apostar por que en un lanzamiento salga cruz, basándome en el hecho de que es mucho (no se sabe cuánto) más probable que en la siguiente tirada de esa moneda salga cruz: evidentemente, las probabilidades de obtener cara o cruz siguen siendo las mismas.

En esto estriba la **falacia del jugador**: en creer que, tras una racha de resultados iguales en un juego de azar (diez veces negro o diez veces impar en una ruleta, por ejemplo), la siguiente tirada tiene necesariamente más probabilidades de resultar en el valor alternativo (rojo o par, en nuestro caso). El equívoco, el defecto cognitivo estriba en que, en el ejemplo de la ruleta, yo no estoy apostando **al conjunto de sucesos** que se expresa por la frase "**sale al menos una vez rojo en once tiradas"**, que es ciertamente mayor que 1/2 ($1-(1/2)^{11} \approx 0{,}9995$ o 99,95%, concretamente), sino que estoy apostando **a que salga rojo en una sola tirada**, y en este caso no importa la historia peculiar de tiradas que haya tenido la ruleta, que, como la moneda, tampoco tiene memoria.

Como comprobación de lo incompetente que es nuestra intuición, incluso en considerción a la longitud de las "rachas" que efectivamente tinene lugar de caras o cruces en una secuencia aleatoria, se puede dar cuenta de un experimento llevado a cabo por Révész[4] en una clase universitaria; se divide al grupo en dos mitades, y a una de ellas se le encomineda la tarea de arrojar consecutivamente una moneda 200 veces, anotando ordenadamente los resultados. A la otra mitad, se le encomienda que anote los resultados que crea más *razonables de una simulación* sobre el experimento que realiza el otro grupo. Pues bien, los resultados (donde H (head) es "cara" y T (tail) es "cruz") fueron los siguientes:

Primer grupo:

THHHHTTTTHHHHTHHHHHHHHTTTHHTTHHHHHTTTTTTHHTHHTHHHT

TTHTTHHHHTHTTTHTTTHHTTTTHHHHHHTTTHHTTHHHTHHHHHTTTT

4 M.F. Schilling, *The Longest Run of Heads,* en *The College Mathematical Journal* vol. 21, nº3 Mayo 1990

THTTTHHTTHTTHHTTTHHTTTHHTHHTHHTTTTTHHTHHHHHHTHTHTT

HTHTTHHHTTHHTHTHHHHHHHHTTHTTHHHTHHTTHTTTTTTHHHTHHH

Segundo Grupo:

THTHTTTHTTTTTHTHTTTHTTHHHTHHTHTHTHTTTTHHTTHHTTHHHT

HHHTTHHHTTTHHHTHHHHTTTHTHTHHHHTHTTTHHHTHHTHTTTHHTH

HHTHHHHTTHTHHTHHHTTTHTHHHTHHTTTHHHTTTTHHHTHTHHHHTH

DTHHTTTTHTHTHTTHTHHTTHTTTHTTTTHHHHTHTHHHTTHHHHHTHH

Para alguien que conozca la teoría de la probabilidad pero no sepa a qué grupo corresponden cada una de las anteriores secuencias, le bastaría con observar la longitud de las rachas de cada una ellas: la secuencia cuyas rachas son más cortas, corresponde siempre a la estimación "razonable" del alumno. Esto es lo que Kahnemann y Tversky, en su taxonomía de intuiciones y heurísticas erróneas sobre conceptos de probabilidad denominan "Ley de los pequeños números": Intuitivamente los sujetos asumen que la moneda es "más justa" de lo que realmente es, de ahí que las rachas "estimadas" sean más cortas que las verdaderamente aleatorias.

8º) Investiga y explica en términos de definición de espacios de sucesos, qué diferencia existe entre las expresiones "que salga cruz al menos una vez en cuatro tiradas" y "que salga cruz en la cuarta tirada, teniendo en cuenta que ha salido cara en las tres tiradas anteriores". Calcula las probabilidades correspondientes a los sucesos o conjuntos de sucesos expresados por ambas frases.

Dando por expicado lo anterior, "que salga cruz al menos una vez en cuatro tiradas" establece el conjunto de sucesos formado por los elementos {CCCX, CCXC, CXCC, XCCC, CCXX, …}, es decir, por aquellos elementos formados por la combinación de sucesos simples C y X en lo que X aparece entre una y cuatro veces. En cambio, de la expresión "que salga cruz en la cuarta tirada, teniendo en cuenta que ha salido cara en las tres tiradas anteriores", al ser cada resultado de una tirada **independiente** del anterior, el conjunto de sucesos defiinidos por la frase en cuestión es {C,X}, siendo la coletilla que comienza "teniendo en cuenta que..." irrelevante para la construcción del conjunto de sucesos por el que se pregunta.

Evidentemente, la probabilidad del suceso X en el segundo caso es 1/2; en el primer caso, en cambio, es preciso reflexionar más, y darse cuenta de que la expresión "que salga cruz al menos una vez en cuatro tiradas" es exactamente la complementaria a "que salga cara en las cuatro tiradas"; si la primera frase se simboliza por A, la segunda sería -A o A^c (no A ó el subconjunto complementario de A); si se admite que la probabilidad total, es decir, del conjunto de todas las combinaciones posibles en cuatro tiradas es 1 (100%), entonces la probabilidad buscada A será la total menos la del conjunto complementario de A, es decir, menos todos los elementos en que los que A no sucede; este subconjunto está formado por un único elemento, CCCC: su probabilidad, por tanto, será 1 entre el número de combinaciones posibles de sucesos simples C y X, es decir, de todo el espacio de sucesos,

$$\Omega = \left\{ \begin{array}{llll} \text{CCCC,} & \text{CCCX,} & \text{CCXC,} & \text{CCXX,} \\ \text{CXCC,} & \text{CXCX} & \text{CXXC} & \text{CXXX} \\ \text{XCCC} & \text{XCCX} & \text{XCXC} & \text{XCXX} \\ \text{XXCC} & \text{XXCX} & \text{XXXC} & \text{XXXX} \end{array} \right\}$$

es decir, de 1/16; luego la probabilidad del conjunto de todos los elemento que contiene al menos una X será el número de todos ellos, entre el número total de elementos, es decir 15/16; que es igual a 1-(1/16), esto es, la probabilidad total, 1, menos la probabilidad del suceso contrario a A, 1/16

La diferencia desde el punto de vista del jugador, que, no olvidemos, apuesta por que se realice la **segunda** expresión, no la primera, es tan grande como la diferencia entre **1/2** (que es la que realmente le afecta) y 15/16.

Dependencia e independencia de variables

Se ha visto en el capítulo anterior cómo hasta mentes entrenadas en la disciplina de pensamiento matemático han sucumbido a percepciones erróneas acerca de los más básicos conceptos de probabilidad. En concreto, hemos visto cómo D'Alembert cae en la falacia del jugador *(vid. supra, p. 106)*. Resuminedo lo expuesto en el anterior capítulo, podría decirse que el núcleo, el centro de la equivocación de D'Alemebert cuando escribe:

> *Porque después de todo no es posible, es incluso físicamente imposible que "cruz" nunca suceda.* ***Por lo tanto, cuantas más veces sucede "cara", más probable es que "cruz" suceda la próxima vez*** […]

estriba en el hecho de que *pone en relación* los sucesos definidos por cada una de las tiradas: es decir, hace que el resultado de una de las tiradas *dependa* del resultado de las demás tiradas. En lenguaje estricto, si consideramos el resultado de una tirada como una *variable* (con un determinado valor si es "cara" y otro si es "cruz"), entonces, según el que escribió las líneas precedentes estas variables son *dependientes.* Y si alguien como él pudo cometer tal error, ¿Cuántos no serán los malentendidos de este tipo que se ven a diario? Veamos algunos ejemplos:

Y ahora a por la niña: la falacia del hijo (hija) al enésimo intento

El jugador (compulsivo) sencillamente no puede creer que tras una racha de jugadas de ruleta con resultado "rojo" las probabilidades de que aparezca "negro" en la siguiente tirada sean las mismas que las de lo contrario. En la obra de Darrel Huff **[27]** , no tan conocida como la popular *Cómo mentir con estadísticas*, se refiere un caso particular que refleja la fuerza de este prejuicio, o, quizá, superstición:

> *El 18 de Agosto de 1918, en El Casino de Montecarlo, el negro obtuvo un récord histórico apareciendo 26 veces consecutivas. Sin tener en cuenta el límite de apuestas normal en un casino, si un jugador hubiera apostado 4$ y en la primera de las jugadas y se hubiera levantado justamente en la última, se habría embolsado 268 millones de dólares. Lo que en realidad sucedió fue que hubo una avalancha cercana al pánico de jugadores por apostar al rojo, que comenzó más o menos a la altura de la decimoquinta ronda* [...] *los jugadores doblaron y triplicaron su apuesta, convencidos de que la posibilidad de que saliera negro a la vigésima jugada no era ni una entre un millón. Pero al final la improbable racha había enriquecido al casino en varios millones de francos.*[trad. propia]

¿Por qué esto es así? Porque, desde el punto de vista del jugador, *no de la banca*, sólo se puede apostar al rojo o al negro por cada una de las jugadas: no se puede apostar a que *no va a salir negro 26 veces seguidas*, lo que sí sería una apuesta segura (las probabilidades son del orden de $1-1/2^{26} \approx 99{,}99999\%$). La banca, en cambio, *sí que juega siempre al conjunto de todos los resultados*, y cuenta con una ganacia segura de 1/37= 2,7% sobre el total de las apuestas a lo largo de todo el tiempo de funcionamiento de la ruleta en cuestión[1].

Un ejemplo menos dramático —pero mucho más común— se ve encarnado cada vez que una familia, tras el tercer, el cuarto,... el i-ésimo retoño del mismo sexo, confían, al igual que D'Alembert, al igual que los jugadores del Casino de Montecarlo, en que las probabilidades de que el sexo de su i+1-ésimo hijo sea, al fin, distinto.

ACTIVIDADES

1º)Lee atentamente el siguiente artículo[2]:

Una pareja de EU con 12 hijos varones vence una probabilidad de 1 en 4,000

Kateri and Jay Schwandt han tenido de forma natural 12 niños y están esperando un hijo más que también podría ser varón

Por **Jacque Wilson**
Miércoles, 29 de octubre de 2014 a las 12:43

3K Facebook | 125 Twittear | 14 g+1 | 0 Pinterest | 3 Comentarios | 3K Compartir | Email

Las probabilidades de tener un niño son 1 en 2, pero de tener 12 varones es de 1 en 4,000 (Getty Images/Archivo).

Lo más importante

- Kateri y Jay Schwandt han tenido de forma natural 12 hijos, todos varones
- La probabilidad es de 1 en 4,000 y aún esperan el hijo número 13, que sería 1 en 8,000

Temas relacionados

- El banco de esperma para genios
- Roxana y la inseminación intrauterina

(CNN) — Ellos tal vez no venzan a la conocida familia Duggar, que tiene 19 hijos y contando, pero esta pareja en Estados Unidos tiene su propia historia digna de un programa de televisión.

Kateri y Jay Schwandt, de Michigan, tienen 12 hijos y ninguna hija, todos concebidos de forma natural, de acuerdo con *Detroit Free Press*.

Lee: ¿Es posible tener un bebé a los 94 años?

Las probabilidades de que eso ocurra son aproximadamente 1 en 4,000. De hecho, el androide C-3PO de *Star Wars* dice que hay más probabilidades de pasar con éxito a través de un campo de asteroides que igualar la marca de los Schwandt.

Pero eso no es todo. El bebé número 13 ya está en camino y, al menos Kateri, está esperando que sea otro niño.

Las probabilidades de que eso ocurra son de 1 en 8,000.

---

1 En una ruleta clásica, o europea, en la que hay 36 números, 18 rojos y 18 negros, y en la que es el 0, el 37º dígito, (ni par ni impar, ni rojo ni negro, etc.) el que da la ganancia a la banca.

2 J. Wilson, *Una pareja de EU con 12 hijos varones vence una probabilidad de 1 en 4000*[artículo en línea] 29 Octubre 2014 [consultado Mayo 2015] CNN México (edición digital: *www.mexico.cnn.com*) Disponible en:http://mexico.cnn.com/salud/2014/10/29/una-pareja-de-eu-con-12-hijos-varones-vence-una-probabilidad-de-1-en-4000

a) A la vista del artículo, ¿Cuál crees que es la probabilidad de que Kateri Schwandt vuelva a tener otro hijo varón, teniendo en cuenta que ha tenido doce hijos varones anteriormente?
b) ¿Cuál es la opinion del artículo al respecto? ¿Es consistente en sus afirmaciones? Es decir, ¿Hay alguna contradicción en el mismo texto?
c) Observa que en el texto siempre se habla de probabilidades genéricas, que se supone son las misma, cualquiera que sea la pareja, Debatir si serían necesariamente ciertas las afirmaciones que se realizan si, en vez de referirse a una pareja genérica, se dijeran sólo de la pareja protagionista de la noticia en cuestión. Es decir, ¿Es equivalente decir que "las probabilidades de que una pareja tenga 12 hijos varones es de 1 entre 4000" a decir que "las probabilidades de que Kateri y Jay Schwandt tengan 12 hijos varones es de 1 entre 4000"? ¿Son necesariamente incompatibles las dos frases entrecomilladas anteriores, o pueden ser las dos ciertas a la vez?
d) Dando por cierto que la probabilidad de tener 12 hijos varones es de 1 entre 4000, ¿Cuál es la probabilidad de que, si solo hubiera (en EU, por ejemplo) 4000 familias exactamente con 12 hijos, haya al menos una (en EU, claro está) que tenga 12 hijos varones?

a) Esta sería una buena ocasión para determinar hasta qué punto el alumno tiene el prejuicio común, o es influenciable por la autoridad de un medio escrito (o por la particular redacción del enunciado de la actividad). La respuesta correcta es, naturalmente, 1/2 ó 0,5 aprox., y a continuación se trataría de argumentar de forma parecida a como lo hemos hecho un poco más arriba, incidiendo en el punto de que tener un hijo o una hija es a efectos probabilísticos un suceso análogo al hecho de lanzar una moneda y comprobar si ha salido cara o cruz; esto es, que el hecho de tener un hijo varón (o mujer) al decimotercer intento es independiente del resultado de los otros doce nacimientos (o lanzamientos de moneda)

b) Según *parece* desprenderse del texto (que es considerablemente vago y que, además, es una traducción un tanto apresurada del inglés) de los primeros párrafos —de hecho, casi justamente hasta el final— la opinión del articulista es que la probabilidad de tener 12 hijos varones es de 1 entre 4000, lo cual es aproximadamente cierto: en realidad, todo lo que se dice en el artículo es correcto, hasta que llegamos a las dos últimas frases:

> *El bebé número 13 ya está en camino, y, al menos Kateri, está esperando que sea otro niño.* ***Las probabilidades de que eso ocurra son de 1 entre 8000****.*[trad. propia]

Lo que el autor llama "eso" que tiene una probabilidad entre 8000 no es que el próximo nacimiento sea un varón, sino el suceso bien distinto, compuesto, de haber tenido 13 hijos todos varones. Tal como está redactada la noticia, la tesis que se sostiene es falsa.

Fijándonos más atentamente, en el pie de la foto donde aparece una enfermera con un bebé, lo que se dice sí que es estrictamente cierto:

> *Las probabilidades de tener un niño son de una entre dos, pero las de tener 12 varones es de 1 entre 4000.*[trad. propia]

Lo que podría deberse a que el pie de foto y el artículo no son de la misma autoría.

c) No, las dos afirmaciones no son estrictamente equivalentes desde el punto de vista probabilístico, pues bien pudiera ser —como algunas veces se sostiene, siendo aún materia de controversia— que ciertas parejas tuvieran un propensión especial a tener hijos de uno u otro sexo. Por ejemplo, en el siguiente artículo se sugiere que esta propensión se da exclusivamente en los varones y que además, es hereditaria[3]:

> *Las probabilidades de que un hombre sea padre de un niño o de una niña están escritos en sus genes, según aprecian ciertas investigaciones.*
> *El análisis de más de 900 árboles genealógicos desde 1600 han sugerido la existencia de un "gen de la paternidad" que determina el sexo del vástago de un hombre.*[...]
> *Mr Gellatly dijo que el sistema* [supuesto] *mantendría el equilibrio entre sexos aproxima damente como se observa en la realidad* [...][trad. propia]

Sin embargo, esto no haría que necesariamente las dos tesis fueran incompatibles; bastaría con que esta propensión a tener o bien niños o bien niñas estuviera igualmente distribuida, esto es, aleatoriamente distribuida entre todas las parejas con hijos del conjunto considerado (EU, u otro país cualquiera). Se podría además recordar al alumno que si este último caso fuera cierto, uno de los signos indicativos de tal suceso sería que habría más parejas con sólo niños o solo niñas; en términos de juego (y estadística), las "rachas" de niños o niñas serían más amplias[4]. Aunque, naturalmente, ambas se compensarían entre sí para dar el porcentaje real entre sexos, aproximadamente de 50-50%

d) Si bien la combinatoria está algo descuidada en el currículo, debería ser factible explicar este ejercicio: se ve claramente que la probabilidad de que haya al menos alguna familia con 12 hijos varones es igual a la total, 1, menos la probabilidad de que no haya ninguna familia con 12 hijos varones: la probabilidad de tener doce hijos varones es, supuesta independencia de cada nacimiento $(1/2)^{12}$: por tanto, la probabilidad de que no tener 12 hijos varones es de $1-(1/2)^{12}= 0{,}999756$: luego, la probabilidad de este suceso en 4000 intentos (familias) será $(0{,}999756)^{4000}= 0{,}3766 \approx 37{,}7\%$ pero esta probabilidad es la de que no hay ninguna pareja de las 4000 parejas con 12 hijos varones: de donde la probablidad de que haya al menos 1 en 4000 es del 62,3%.

El caso de Sally Clark

Hasta ahora hemos visto el caso de la falacia del jugador, consistente en suponer que dos sucesos independientes son dependientes. Pues bien, vamos a analizar el caso donde sucede, lo contrario: dos variables que *sí dependen* una de la otra son tomadas por independientes: después veremos que a esta primera mala interpretación se le sumará otra, la que viene a ser llamada la **falacia del fiscal** (*prosecutor's fallacy*)[5].

3 F. McCrae, *Boy or a girl?It's in your genes, dad* [artículo en línea] 10 Diciembre 2008 [consultado Mayo 2015], Daily Mail, en *Mail Online* (ed. digital: *www.dailymail.co.uk*) Disponible en: http://www.dailymail.co.uk/sciencetech/article-1093702/Boy-girl-Its-genes-dad.html ; El artículo original del autor, bastante más cauto: C. Gellatly, *Trends in Population Sex Ratios May be Explained by Changes in the Frequencies of Polymorphic Alleles of a Sex Ratio Gene* [artículo en línea pub. 10 Dic. 2008][consultado Mayo 2015]Evol Biol (2009) 36:190–200, DOI: 10.1007/s11692-008-9046-3

4 Según parece, no existen en los tiempos actuales suficientes parejas con un número elevado de hijos (al menos en los países donde los datos son fiables) como para realizar un análisis de esas mencionadas rachas, confirmando o refutando la tesis de la "propensión".

5 W.C. Thompson y E.L. Schumann, *Interpretation of Statistical Evidence in Criminal Trials: The Prosecutor's Fallacy and the DefenseAttorney's Fallacy* [artículo en línea], Law and Human Behavior, Vol. 11, No. 3, I987 [consultado Mayo 2015] Disponible en [previo registro]: http://www.jstor.org/stable/1393631

He aquí un buen resumen del caso de Sally Clark[6]:

> *Sally Clark trabajaba en 1990 como abogada en un estudio en el centro de Londres. Se casó con Steve Clark, también abogado, y se mudaron a Manchester. Allí nació Cristopher, el primer hijo de la pareja. Fue el 22 de septiembre de 1996. Menos de tres meses después, el 13 de diciembre, Sally llamó a una ambulancia en un intento desesperado por salvar la vida de su hijo. No alcanzó. Cuando los paramédicos llegaron a su casa, Cristopher ya estaba muerto. Sally era la única que estaba con el niño en ese momento. Los médicos que revisaron el cuerpo de la criatura no lograron descubrir nada significativo y consideraron que la muerte había sido por causas naturales (hubo incluso alguna evidencia de una infección respiratoria) y ningún signo de que hubiera habido falta de cuidado o atención por parte de la madre.*
>
> *El matrimonio Clark volvió a tener otro niño, Harry, que nació prematuramente a menos de un año de la muerte de Cristopher: el 29 de noviembre de 1997.* […] *menos de dos meses más tarde, el 26 de enero de 1998, Harry murió repentinamente también y, una vez más, Sally era la única persona que estaba con el bebé en su casa en el momento de la muerte.*
>
> *Esta vez, Sally y su marido fueron enviados a prisión, pero mientras él fue absuelto casi inmediatamente, Sally fue acusada del doble homicidio de sus hijos. Aconsejada por sus abogados, Sally nunca contestó ninguna pregunta, pero siempre sostuvo que era inocente.*
>
> *En el momento del juicio, los abogados defensores sostuvieron la hipótesis de que los niños fallecieron de lo que se llama Síndrome de Muerte Súbita del Lactante (SMSL), pero el jurado en 1999 la encontró culpable después de una declaración impactante de un famoso pediatra inglés, coronado como "caballero" por la reina, Sir Roy Meadow.*[...]
>
> *Meadow fue un testigo clave para el fiscal, ya que él sabía que los datos que se conocían en ese momento decían que la probabilidad de que un niño muriera de SMSL era de uno en 8500 (aproximadamente). Por lo tanto, concluyó Meadow, la probabilidad de que dos niños murieran de SMSL en la misma casa debía resultar de la multiplicación de estos dos números:*
>
> *(1/8500) x (1/8500) = 1/72.250.000.*
>
> *Es decir, la probabilidad de que se produjeran dos casos en el mismo núcleo familiar (según Meadow) era de uno en casi 73.000.000. Y agregó: eso solamente podría pasar en Gran Bretaña una vez por siglo. Ese fue el toque final.*
>
> *Sally Clark fue condenada a prisión perpetua.*

Resulta evidente que Meadow y el jurado consideraron que ambas muertes eran *variables independientes*, y que por tanto la probabilidad del suceso compuesto por las dos era el producto de las probabilidades respectivas. Pero no tardó en alzarse una voz denunciando el error: El profesor Greene, presidente de la *Royal Statistical Society* (RSS) escribió una tajante carta a la Cámara de los Lores[7] al siguiente tenor:

> ***El cálculo que conduce a 1 entre 73 millones es inválido. Sólo sería válido si los casos de SIDS*** [SMSL] ***aparecieran de forma independiente dentro de las familias****, un supuesto que necesita ser justificado empíricamente. No sólo tal justificación empírica no fue provista en el juicio, sino que hay razones de peso para suponer que tal supuesto es falso. Puede muy bien haber factores genéticos o ambientales desconocidos que predisponen ciertas familas al SIDS* [SMSL], *de manera*

6 A. Paenza, *Sally Clark* [artículo en línea], 20 Febrero 2013 [consultado Mayo 2015], Página 12 (edición digital: *www.pagina12.com.ar*) Disponible en: http://www.pagina12.com.ar/diario/contratapa/13-214264-2013-02-20.html

7 P. Green, President of the Royal Statistical Society, *Letter from the President to the Lord Chancellor regarding the use of statistical evidence in court cases* [recurso en línea], 23 Enero 2002 [consultado Mayo 2015], en *www.rss.org.uk* (sitio en línea de la Royal Statistical Society). Disponible en: http://www.rss.org.uk/Images/PDF/influencing-change/rss-use-statistical-evidence-court-cases-2002.pdf

que un segundo caso dentro de la familia resulte ser mucho más probable de lo que sería en otra familia, aparentemente similar.[trad. propia]

Entre tanto, el análisis de independencia de casos de SMSL dentro de las familias ya había sido acometido por alguien [7]:

La probabilidad de que dos muertes sucesivas en la misma familia sean de SIDS viene dado por:
P=1/1300·x/1300
donde x es el ***factor de dependencia*** *para la segunda SIDS, estimado más arriba entre 5 y 10*

Lo que da un resultado de, en el caso más desfavorable para la acusada, de 338.000; cifra la cual también puede parecer una cifra muy baja; y es aquí donde se da la *falacia del fiscal*. Citaremos ni más ni menos que a R. A. Fisher, quien expuso el asunto en los términos más rotundos posibles[8]:

El caso cuya probabilidad es de un millón de millones de millones...contra uno sucede una vez en un millón de millones de millones... de ocasiones, por mucho que nos sorprenda que nos pase precisamente a nosotros.[trad. propia]

O en otras palabras: no es en absoluto sorprendente —sucede todos los años —que a *alguien* le toque la lotería de navidad; es sorprendente que nos toque *precisamente a nosotros* (o precisamente a cualquier persona escogida de antemano).Juzgar esta *improbabilidad de inocencia* en lugar de la probabilidad de culpabilidad es lo que le lleva a caer en la falacia. Imaginemos que, por el hecho de que es muy improbable que a Fulano le toque la lotería de Navidad, un fiscal decida encausar a Fulano en base a que la probabilidad de que le haya tocado es tan baja que necesariamente ha debido cometer un fraude. El fiscal considera que la probabilidad de que el *acusado sea inocente es muy baja*, cuando lo que estrictamente significa que exista 1 probabilidad entre 338.000 de dos caso de SMSL en la misma familia es que, de cada 338.000 familias, al menos una al año ha sufrido la tragedia descrita. El mismo Ray Hill indica la cifra de 650.000 segundos nacimientos al año en todo el Reino Unido, lo cual quiere decir que, por promedio, aproximadamente 2 casos como el de Sally Clark por año suceden en UK (*vid.*[7]).

Lo que sucede es que Sally Clark solo es una persona determinada desde el punto de vista del fiscal: pero, en realidad, si está siendo juzgada no es por ser ella, Sally Clark, persona concreta sino *porque* (es decir, posteriormente al hecho de que) en su familia se han dado los dos casos de SMLM. A todos los efectos, Sally Clark es *una persona cualquiera*, y, por lo tanto, a falta de otra evidencia, podría constituir uno de esos 5 casos al año que el profesor Hill estimaba.

ACTIVIDADES

1º)Lee atentamente el siguiente artículo:[aquí se insertaría el extracto correspondiente de A. Paenza de la p. 119]

7 R. Hill, *Multiple sudden infant deaths – coincidence or beyond coincidence?*[recurso en línea] 2004 [consultado Mayo-2015]Pediatric and Perinatal Epidemiology, vol. 18, pp. 320–326. Disponible en: http://www.cse.salford.ac.uk/staff/RHill/ppe_5601.pdf

8 Citado por K.Mather, *Reviews* [artículo en línea] Febrero 1973 [cons. Mayo 2015] Heredity 30: 89-91; doi:10.1038/hdy.1973.11 Disponible en: http://www.nature.com/hdy/journal/v30/n1/pdf/hdy197311a.pdf

a)¿Qué conceptos de probabilidad que hayas estudiado aparecen (aunque no aprezcan denominados explícitamente) en el texto?
b)A la luz de lo que sabes, ¿Te parece correcto el razonamiento del Dr. Meadow? ¿Porqué?
c)¿Crees que el verdadero quid de la cuestión es puramente matemático, o importan otras consideraciones? ¿O quizá sería mejor decir que primero hay que establecer con rigor los elementos del problema y después pasar al cálculo?
d)En tu opinión, ¿Las matemáticas consisten en calcular o resolver problemas bien planteados, para después calcularlos, o más bien en plantear bien ciertos problemas?

a)Es de suponer que el alumno ha sido instruido, aunque sea sólo nominalmente, en los conceptos de variable estadística, de suceso asociado al valor de esa variable (el suceso de que X=3, por ejemplo[9]), independiente y dependiente. Cada fallecimiento de uno de los hijos sería uno de estos sucesos, cada uno con una probabilidad asignada, y su producto es el medio de conocer la probabilidad conjunta de los dos sucesos, la cual, supuesta independencia, es el producto de las probabilidades respectivas.

b)"A la luz de lo que sabes" es una expresión que pretende que, con la ayuda del profesor, el alumno se de cuenta de que en realidad el problema no está bien planteado; de hecho, según el texto seleccionado, el enunciado está <u>incompleto</u>: el Dr. Meadow *<u>supone</u> que las variables son independientes,* pero eso es algo que no se puede deducir del resto del texto.

Probablemente sería útil proponer ejemplos donde la dependencia entre las variables fuera más intuitivamente plausible que en el caso analizado. Por ejemplo, ¿Es pequeña la probabilidad de medir más de 2,00 metros de alto **y** pesar más de 100 kgs.? Si ambos sucesos fueran independientes, y digamos, P(>2m.)=0,05 (5%) y P(>100kg.)=0,05 (5%), entonces la probabilidad conjunta $P(>2m \cap >100\ kg.)= P(>2m)\cdot P(>100kg.)= 0{,}05\cdot 0{,}05=$ 0,0025 (0,25%); pero se sabe que una gran proporción de personas que miden más de 2,00 m. de altura pesan además más de 100 Kg. ya que peso y altura están relacionados, son <u>dependientes.</u> Por lo tanto, $P(>2m \cap >100\ kg.)$ no será igual a $P(>2m)\cdot P(>100kg.)$, sino más bien, muy similar a P(>2m)= 0,05%; en términos de probabilidad condicional, diríamos que la probabilidad de pesar 100 kg. condicionada a medir 2,00m es muy similar a 1; luego, por la definición de probabilidad condicional, $P(>2m \cap >100\ kg.) = P(>100kg.|>2{,}00m)\cdot P(>2{,}00m) \approx 1\cdot P(>2{,}00m)$

c) En este apartado podría decirse que el entendimiento y la definición de cada elemento del problema son esenciales. El verdadero *quid* estriba en si los sucesos "muerte del primer hijo por SMLM en la familia Clark" y "Muerte del segundo hijo en la familia Clark" son realmente independientes, cosa que solo el contexto real de los datos puede suministrarnos. Es este un extremo que recorre los enunciados de todas las actividades propuestas en el presente trabajo y que caracterizan de modo especial a los problemas que involucran incertidumbre al nivel de enseñanza secundaria.

9 Hablaremos indistintamente de "sucesos" y de "variables", asumiendo que podemos llamar al suceso "primer hijo muere" ≡(X=0) y al suceso "segundo hijo muere"≡(Y=0). Asumiendo que vida y muerte son dicotómicos, X=1 e Y=1 pueden simbolizar los sucesos en que cadauno de los hijos vive.

Sobre este punto podría establecerse un debate con el alumno, quien a menudo se pregunta sobre la "utilidad" de los conocimientos que se le imparten. Y entrecomillamos "utilidad" porque, pese a lo mucho que se le fustiga, el alumno suele considerar un concepto de "utilidad" mucho más cercano a lo cognitivo de lo que se supone; suelen ser sus mayores, entre ellos el docente, los que más insisten en los aspectos pragmatistas (no pragmáticos) de la educación. El alumno suele alegrarse de ser capaz de resolver un problema *genuinamente*, no en base al aprovechamiento material más o menos futuro de su nueva capacidad. Pero sobre todo, y en esto los partidarios de PISA tienen razón, le interesan problemas que pueda trasladar a la realidad, (aunque quizá no a su realidad).

d) Aquí nos permitimos, por una vez, ya que el contexto del problema nos lo permite, plantear una discusión con el alumno sobre la misma docencia de las matemáticas. Generalmente el alumno estima —porque no se le ha mostrado ninguna alternativa— que las matemáticas, incluyendo la parte de probabilidad y estadística, son esencialmente resolución de problemas: lo cual, aunque sea una perspectiva tremendamente parcial desde el punto de vista matemático, podría ser legítima, considerando la materia desde los supuestos de la enseñanza secundaria enarbolada por gran parte de la didáctica, entre ellos los responsables del informe PISA, y también por las leyes de educación más recientes. Pero lo que buena parte del alumnado y la sociedad en general entiende por resolución de problemas en realidad debería llamarse estrictamente computación o cálculo; a lo sumo de cálculo simbólico o lógico: lo que puede hacer un *software* específico si se le suministra un *input* bien configurado y determinado.
Ahora bien, si un software puede hacerlo, ¿Por qué razón se estudian matemáticas obligatoriamente toda la ESO, a menos que uno vaya a ser programador? La única razón de ser de la educación matemática obligatoria es que cultive competencias generales básicas, no específicas y profesionales. Pues bien, creemos que esa competencia general básica que otorgan las matemáticas es, en primer lugar, la de plantear adecuadamente los problemas a los que el ciudadano que recibe la enseñanza se enfrente o vaya a enfrentarse en un futuro[10]. A la vez que se intenta plantear el problema, percatarse de que una buena parte de ellos, (o quizá todos, diríamos si en la sala solo hubiera matemáticos y físicos teóricos), solo obtendrán una solución satisfactoria si se trasponen sus objetos, propiedades y contexto a elementos, operaciones y relaciones de carácter matemático. Finalmente, buscar los intrumentos, que el currículo suministra, para resolver el problema de aquella manera planteado.

2º) Posteriormente, se descubrió que la probabilidad de que una familia tenga dos casos de muertes súbitas infantiles consecutivas no eran de 1/73.000.000, sino sólo de 1/338.000. [Aquí se insertaría el texto del Pr. Hill] *¿En qué cambia eso las cosas?¿No sigue siendo la probabilidad tan baja que la acusada es probablemente culpable? Debatid en grupos el asunto y redactad vuestras conclusiones brevemente, explicando las razones por las que habéis llegado a ellas.*

10 Y que también pueden ser de índole estrictamente matemática, si el alumno decide hacer de las mismas el centro de su desarrollo profesional o, por qué no, personal. Piénsese en las aportaciones a la matemática por parte de elementos no profesionales, como Neper, Fermat, Pascal, Descartes, Leibniz, Bayes, Galois, Bolyai..

Siempre es interesante, emulando los trabajos de Khanneman y Tversky [**29**], comprobar cuáles son las ideas del alumno sobre asuntos relacionados con la incertidumbre, entre otras cosas para compararlos con nuestras primeras reacciones ante temáticas similares (caso de que lo recordemos y seamos lo suficientemente honestos con nosotros mismos). El símil explicado anteriormente, en que Fulano es imputado por haberle tocado la lotería, es bastante intuitivo, en nuestra opinión; no obstante, habría que hacer una pequeña distinción. Todos sabemos, por ejemplo, que pueden existir razones por las cuales es ventajoso, desde el punto de vista fiscal (entre otros) comprar décimos de lotería premiados y aducir que se es el agraciado de los premios correspondientes. Si cierto individuo declarara que le ha tocado la lotería de navidad cierto número de veces en cierto período de tiempo, entonces sí que habría razones de peso para pensar que esto no es cierto. ¿Por qué sí en este caso y no en el estudiado?
En el caso Clarke, el hecho de que la probabilidad de que sucedieran dos muertes consecutivas fuera más pequeña hizo posible y plausible que a una (más bien a 2) pareja le sucediera eso mismo en el transcurso de un solo año. Sally Clark sería, en este caso, esa persona cualquiera, o Fulano, a la que le ha tocado la (siniestra) lotería del doble deceso familiar. En cambio, en el caso del individuo a quien hacemos referencia, la probabilidad de que al guien le toque el gordo de la lotería de navidad, digamos, al menos tres veces en 10 años (1 entre 12 billones), es tan pequeña que hace implausible que le pueda suceder, no a una persona concreta, sino incluso *a una persona cualquiera,* dentro de lo posible según el número de décimos comprados al año de lotería de navidad española, y el número de años considerado. Por otra parte, obra la circunstancia de que ninguna otra prueba positiva, salvo la supuesta rareza del acaecimiento de la doble tragedia familiar podía esgrimirse en contra de Sally Clarke; de hecho, esa era la causa de haber sido *seleccionada para ser juzgada,* sin la cual habría debido considerarse una persona *cualquiera*, a efectos de la probabilidad; mientras que en el caso del afortunado triple ganador de la lotería de Navidad existían otras razones de peso para sopechar que practicaba otras operaciones fraudulentas de blanqueo de dinero; esas otras sospechas son las que le conviertene en una persona *concreta*, desde el punto de vista de la probabilidad, lo cual hace implausible la “suerte” de la lotería le haya tocado *precisamente* a él. [11].

11 Se pueden encontrar resultados numéricos sobre este tópico y otros afines en A. Gordaliza, *La lotería que siempre toca* [artículo en línea], 21 Dic. 2012 [consultado Mayo 2015], El País (edición digital, www.elpais.com). Disponible en: http://sociedad.elpais.com/sociedad/2012/12/21/actualidad/1356105096_890961.html

Probabilidades condicionadas

Pocos conceptos relacionados con la materia provocan tal cantidad de equívocos entre los alumnos y los no alumnos: ya sea porque el lenguaje ordinario distingue mal entre antecedente y consecuente de una proposición condicional, ya sea por indefinición (por elipsis de las condiciones implícitas) de un conjunto de sucesos, lo cierto es que existe una acendrada propensión a:

· Confundir la probabilidad *absoluta* (o inicial, o *a priori*) de un suceso A, P(A), y la probabilidad *condicional* (o final, o *a posteriori*) de ese mismo suceso A *condicionado* al acaecimiento de otro suceso B, P(A|B).

· Confundir entre probabilidades condicionadas *simétricas*, o bien entre el antecedente y el consecuente: es decir, tomar P(A|B) (probabilidad de que suceda A condicionado a que suceda B) por P(B|A) (probabilidad de que suceda B condicionada a que suceda A)

· Confundir entre probabilidad de que ocurra un suceso condicionado a la ocurrencia de otro P(A|B) con la probabilidad de que ambos suceso ocurran $P(A\cap B)$

En realidad pueden englobarse ambos a la misma clase de error: fallo en la definición tanto de sucesos simples como de compuestos, es decir, resultantes de las diversas operaciones lógicas posibles entre ellos. Es este uno de los tramos de la matemática donde más claramente puede comprobarse la importancia de una notación clara y una transcripción estricta de las proposiciones del lenguaje ordinario a objetos matemáticos; de hecho, esta última es la verdadera dificultad de los problemas derivados de unas definiciones muy sencillas (aunque ofrezcan diversas, complejas e interesantes interpretaciones[1], desde el punto de vista matemático, estadístico y epistemológico): Efectivamente, si por definición:

$$P(A|B)=\frac{P(A\cap B)}{P(B)}$$

Es decir, la probabilidad del (acaecimiento del) suceso A condicionada al (acaecimiento del) suceso B es la probabilidad del (ac. del)suceso $A\cap B$, o intersección entre A y B, es decir, del suceso compuesto de que acaezcan *conjuntamente* A **y** B, dividido por la probabilidad de que acaezca B (la condición), entonces

$$P(A|B)=\frac{P(A\cap B)}{P(B)} \longrightarrow P(B|A)=\frac{P(A\cap B)}{P(A)} \longrightarrow P(A|B)=\frac{P(B|A)\cdot P(A)}{P(B)}$$

1 No entraremos en el aún abierto y candente debate entre los defensores de distintas interpretaciones de la probabilidad (y distintas escuelas de inferencia estadística) abiertas por el concepto de probabilidad condicional y centradas en las distintas lecturas (objetiva o subjetiva, frecuentista o bayesiana) del Teorema de Bayes (1763) (que en realidad fue generalizado y formalizado por Laplace mucho después, en 1812)**[20]** y que da lugar a interesantes problemas (**[22]** **[17]**, pp. 136-7) y paradojas(**[16]**, cap.1)lamentablemente fuera del currículo de enseñanza secundaria.

puede obtenerse la expresión más sencilla que relaciona las probabilidades simétricas P(A|B) y P(B|A), y que no es nada difícil de aplicar correctamente siempre que *los términos estén bien definidos,* lo cual no suele ser el caso, como veremos a continuación.

Probabilidad condicionada implícita

A menudo se habla —y se escribe— sobre la probabilidad de que suceda determinado evento, sin especificar que esta probabilidad referida está *condicionada* a otro suceso, bien porque la condición se da por sabida dentro del contexto de la situación, bien porque la condición es tan necesaria para que la información que se da sea relevante, que hacerla explícita se considera innecesario. Como suele suceder cuando entran en relación los lenguajes ordinario y matemático, pueden producirse divertidos errores como el que sigue[2]:

Sólo un 13% de las mujeres son directivos en las empresas españolas

Una de cada tres compañías no cuenta con ninguna mujer entre sus altos cargos, y en el caso de cotizadas del Ibex, este ratio apenas llega al 10%

JOSÉ MARÍA CAMARERO | MADRID

Me gusta 5

21 enero 2015 20:57

María Garaña, presidente de Microsoft España, y Luis Enríquez Nistal, consejero delegado de Vocento, firman un acuerdo. / Isabel B Permuy

5

26

Las mujeres siguen sin situarse al frente de las empresas y su lugar en los diversos órganos de dirección de las compañías continúa siendo minoritaria. Así se desprende del "Mujeres en la Alta Dirección en España", en el que se refleja que de los 1.735 directivos de primer nivel de 147 grandes compañías, sólo 234 son mujeres, lo que supone un 13,4% del total.

Evidentemente, lo que el articulista intentaba expresar sin éxito es que "sólo el 13% de los directivos de las empresas españolas son mujeres". Si fuera el caso que señala el autor, sin duda el lector conocería intimamente o sería familia de al menos una alta directiva de una empresa española, lo cual parece bastante dudoso. Pero sigamos en las actividades.

ACTIVIDADES

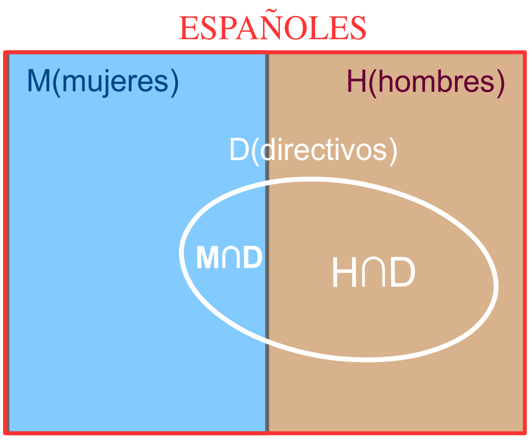
1º) Lee el siguiente extracto de noticia. ¿Encuentras algo sospechoso en la redacción del título? ¿Es equivalente lo que sostiene el titular y el párrafo posterior?¿Cuál de las dos propoosiciones (caso de que no sean equivalentes) crees que es la correcta?

2 J.M. Camarero, *Sólo un 13% de las mujeres son directivos en las empresas españolas* [artículo en línea].Madrid, 21 Enero 2015 [consultado Mayo 2015] El Correo (ed. digital *www.elcorreo.com*) [vía *www.malaprensa.com*] Disponible en: http://www.elcorreo.com/bizkaia/economia/empresas/201501/21/solo-mujeres-directivos-empresas-20150121205156-rc.html

Efectivamente, puede comprobarse que en el primer párrafo (no en la introducción) el redactor entiende que la probabilidad o frecuencia relativa (medida en porcentaje) de “mujeres directivos” debe calcularse entre “directivos”: no es lo mismo decir que el 13% de las mujeres españolas son altas directivas de empresa, a que el 13% de los altos directivos españoles son mujeres. Es interesante insistir frente al alumno en esto, que puede parecer obvio, pero que ayuda sobremanera a fijar conceptos y, sobre todo, a identificar bien los elementos del lenguaje que no se hacen explícitos en el uso ordinario, pero sí en el uso matemático (o lógico). El articulista, por ejemplo, se equivoca al identificar en el titular la probabilidad (en términos de frecuencia) de ser directivo condicionada a ser mujer es decir, $P(D|M)$,con la probabilidad (frecuencia) de ser mujer **condicionada** a ser directivo, es decir, $P(M|D)$; éste último valor es el que es igual al 13%.

2°) Recordarás que un porcentaje es una clase de medida de la frecuencia relativa, en concreto fr·100. En la parte de estadística descriptiva has comenzado representando frecuencias relativas de datos conocidos, y después has pasado a considerar la probabilidad asociada cuando has estudiado la distribución de probabilidad normal. ¿Puedes plantear el enunciado que se sostiene en el párrafo, no en el titular, en términos de probabilidad condicional?

Fig. 21: Esquema conjuntista probabilidad condicionada: Mujeres y Directivos



Además de servir de repaso del concepto frecuentista de la probabilidad (el que se enseña en secundaria), este es el momento de plantear el enunciado correcto en términos matemáticos. Si consideramos el espacio de probabilidad total como el conjunto de la población activa española, que durante el primer trimestre de 2015 (por ejemplo)[3] es de 22.899.400, entonces, por la regla de Laplace, y teniendo en cuenta los valores reflejados en el artículo:

$$P(M\cap D)=\frac{243}{22899400}\quad P(D)=\frac{1735}{22899400}\quad P(M|D)=\frac{P(M\cap D)}{P(D)}=\frac{243/\cancel{22899400}}{1735/\cancel{22899400}}=0{,}14=14\%$$

por cierto, la probabilidad de la que se habla en el titular , es decir, el porcentaje de mujeres españolas que son altos directivos es:

$$P(D|M)=\frac{P(M\cap D)}{P(M)}=\frac{243/22899400}{1/2}=0{,}0000212=0{,}0021\%$$

3 http://www.ine.es/daco/daco42/daco4211/epa0115.pdf

Condición interpretada como causa

El ejemplo anterior puede parecer trivial o inocente, un simple error de redacción de un articulista demasiado preocupado por acortar el titular. Sin embargo, el error se vuelve mucho más usual cuando el planteamiento no es tan evidente. En la vida cotidiana —e incluso en estudios científicos, como veremos— a menudo el equívoco se forja cuando trata de descubrirse la *causa* de un determinado fenómeno, o al menos una *condición* suficiente del mismo, como se ve en este ejemplo ([**24**] **,** p. 112)[3] :

> *Considérese el siguiente ejemplo, del* Management Focus:
>
> > Los resultado de una reciente encuesta a 74 CEO's indica que podría haber una relación entre la posesión de una mascota en la infancia y el éxito de una futura carrera. El 94% de los CEOs todos ello empleados en compañías de Fortune 500, tuvieron un perro, un gato, o ambos, de jóvenes.
>
> […] *Por lo que nosotros sabemos, más del 94% de los niños criados en el ambiente del que proceden los CEOs tiene mascota, en cuyo caso la dependencia podría ser negativa. Puede que el éxito ejecutivo en realidad se halle relacionado con el el hecho de cepillarse los dientes. Probablemente todos los CEOs se lavaban los dientes de pequeños, y podemos imaginarnos que la autodisplina que adquirieron les condujo al éxito en los negocios*[...]

Chistes aparte, el autor señala que no es raro encontrar estas concepciones falsas en determinadas ciencias del hombre:

> *Los psicólogos no son inmunes a estos errores de juicio. Por ejemplo, Branden (1984) escribe:* "No puedo pensar en un sólo problema psicológico […] que no sea rastreable hasta el problema de pobre autoestima". […][pero] *La gente podría tener pobre autoestima* [precisamente] **porque** tiene problemas psicológicos [trad. propia]

Ensayemos la proposición del ejemplo concreto de los CEOs y las mascotas en la actividad siguiente:

ACTIVIDADES

1º) Lee atentamente la siguiente noticia (vid. supra) *¿Te parece plausible lo que se defiende en el texto? Debatidlo por grupos y entregad un párrafo con el resumen de vuestras conclusiones y los argumentos que os han llevado a ellas.*

El propósito de esta actividad es comprobar si el alumno, que ya sabe que está en clase de matemáticas (y que se está impartiendo la materia de probabilidad y estadística) es capaz de sustraerse a la amplia gama de tesis especulativas y/o espúreas que puede concitar el texto, y detectar el elemento lógico y probabilístico de la cuestión. Por ejemplo: puede que un grupo de alumnos emita sus opiniones sobre si el hecho de tener mascota en la niñez forja habilidades sociales beneficiosas para un CEO, o si sucede lo

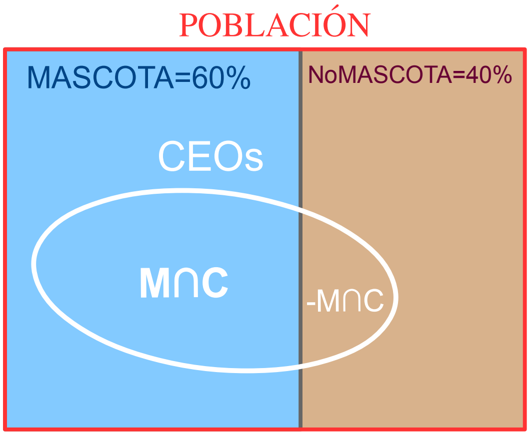
3 Debemos el ejemplo y el planteamineto de la actividad subsiguiente a A. Gordaliza, tutor del presente Trabajo Fin de Máster.

contrario, es decir, que si has tenido mascota es probable que hayas tenido pocos hermanos y por tanto la falta de empatía subsiguiente (o cualquier otra especulación basada en experiencias puntuales o hecho idiosincrásicos culturales, sociales, etc.); y puede que otro grupo, más avispado, detecte el hecho de que el supuesto antecedente está ubicado en el lugar del consecuente, y que si se colocaran en orden inverso ambas partes de la proposición condicional subyacente el enunciado sería más plausible que el del texto: es decir, es más plausible pensar que el hecho de ser CEO (antecedente) implica haber pertenecido a un estrato social en el cual es más probable haber tenido una o más mascotas en la infancia (consecuente), que la proposición simétrica que el texto indica.

2°) ¿Sabrías definir la tesis del texto (o su simétrica) en términos de probabilidad condicional? Puedes suponer, además de las cifras dadas en el texto, que hay un CEO por cada 100 habitantes y que el 60% de los adultos han tenido alguna mascota de niños.

Según el enunciado del problema, P(CEO)=0,01; P(Mascota)=0,6 y P(Mascota|CEO)=0,94 esta última es el verdadero quid de la cuestión. Es conveniente insitir en que no es lo mismo "haber tenido mascota de niño si se es CEO" a "ser CEO si se ha tenido mascota de niño". Esta última se simbolizaría P(CEO|Mascota), e indicaría (sólo indicaría) que hay una relación entre tener mascota de niño y el éxito profesional posterior de adulto. Con los datos a nuestra disposición, tenemos:

Fig. 22: Esquema conjuntista probabilidad condicionada: Macotas y CEOs



$$P(\text{Mascota}|\text{CEO}) = \frac{P(\text{Mascota} \cap \text{CEO})}{P(\text{CEO})} \longrightarrow P(M \cap C) = P(M|C) \cdot P(C) = 0{,}94 \cdot 0{,}01 = 0{,}0094$$

Por lo tanto, la verdadera probabilidad de ser CEO habiendo tenido mascota de niño es:

$$P(\text{CEO}|\text{Mascota}) = \frac{P(\text{Mascota} \cap \text{CEO})}{P(\text{Mascota})} = \frac{0{,}0094}{0{,}6} = 0{,}01567 \approx 1{,}57\%$$

Nótese que la probabilidad P(CEO|Mascota) suponen un aumento del 57% sobre la mera probabilidad de ser CEO (P(CEO)=0,01=1%), lo que podría resultar significativo, si no fuese porque este incremento no se debe, probablemente, a la circunstancia de haber tenido mascota, sino, como se dijo antes, a que los CEOs en su infancia están ubicados en un contexto social, cultural, urbano, etc. que promueven tanto el hecho de tener mascota en la infancia como el de ser CEO en la edad adulta.

El problema de Monty Hall

Exploraremos ahora una de las variantes, probablemente la más famosa, de un problema que puede rastrearse hasta la obra de Bertrand *Calculs des Probabilités* (1889)([**22**], p. 268) El problema sigue siendo tan conocido como controvertido, pese a que su resolución no deja lugar a dudas. El asunto es como sigue, según redacción propia (la original, teatralizada, de Steve Selvin, en [**22**], p. 265-6):

> En un concurso televisivo, se le plantea al concursante optar por una de tres puertas y llevarse el premio que se esconda detrás. Tras una de las puertas hay un coche; tras cada una de las otras dos puertas, una cabra. Una vez hecha la elección, el presentador del concurso, que sabe qué esconde cada puerta, abre una de ellas, donde muestra una cabra. Entonces propone al concursante cambiar su primera elección ¿Cuál es la mejor opción de llevarse el coche: cambiar de decisión o mantenerla?

Probablemente no hay otro enunciado que se pueda proponer en un aula de enseñanza secundaria (o universitaria, si queremos ser justos) que cause más equívocos, llegando incluso a suscitar oposiciones enconadas, acompañadas de reacciones fuertemente emocionales. Gorroochurn**,** dentro de la crónica que realiza sobre el planteamiento original del problema y sus diversos avatares históricos (la propuesta original en sus términos y con el nombre de Monty Hall data, según él, de 1975), señala que incluso algún matematico profesional ha sido burlado por su inocente y sencilla apariencia:

> *Todavía más sorprendente es el hecho de que entre aquellos que creyeron el argumento erróneo había muchos matemáticos, algunos del máximo calibre. Paul Erdős fue uno de ellos. Erdős era indiscutiblemente el matemático más prolífico de todos los tiempos, con publicaciones en un toda una gama de campos que incluían combinatoria, teoría de números, análisis y probabilidad. De todas formas, el problema de Monty Hall demostró ser una trampa para él*[...]. *Hoffmann (1998, p.237) recoge los esfuerzos que la matemática Laura Vázsonyi tuvo que enfrentar para convencer a Erdős de la solución correcta del problema*[...]
>
> > Le dije a Erdős que la respuesta era cambiar de decisión […] pero, para mi sorpresa, dijo: "No, eso es imposible. No debería haber diferencia".[...] ya había experimentado que la gente se excita y suele perder la paciencia ante la respuesta […] tuve que mostrarle el árbol de decisión que empleo en mi curso de Técnicas Cuantitativas de Administración[...] Pero todo fue inútil[...] Una hora después volvió verdaderamente enfadado. "No me has dicho porqué debería cambiar de decisión", dijo "Qué es lo que pasa contigo." […] Vázsonyi había visto esta reacción antes en sus estudiantes, pero jamás pensó que la tuviera el matemático más prolífico del siglo XX.

El planteamiento clásico del problema se relizará seguidamente, en las actividades. Pero antes, expongamos el razonamiento de vos Savant (*vid. supra*), como ejemplo de explicación del mismo más intuitivo que conozcamos:

> *Supongamos que hay un millón de puertas, y tú escoges la número 1. Entonces el presentador, que sabe donde está el premio y que por tanto nunca abrirá la puerta correpondiente, abre todas las puertas excepto la 777.777. Cambiarías tu decisión en favor de esa puerta al momento, ¿Verdad?*

ACTIVIDADES

1º) Supongamos que mañana se estrena una nuevo miniconcurso televisivo llamado "O cabras o cobras", consistente en la presentación al concursante de 3 puertas cerradas como éstas:

Tras dos de ellas hay sendas cabras, y tras la restante una bolsa con 1.000.000 de €. Por ejemplo de la siguiente forma:

Aunque no tienen por qué estar dispuestas de esta forma exactamente. Volvamos a las puertas cerradas:

El presentador nos da a escoger entre una de ellas: lo que haya detrás pasará a pertenecernos (se supone que estamos más interesados en el millón de euros que en cualquiera de las cabras). Supongamos que escogemos la que se ha señalado con el trazo rojo. Entonces, el presentadornos pregunta: ¿Está usted seguro de su decisión? *Nosotros le respondemos que, naturalmente que sí. "¿Por qué clase de individuos sin personalidad nos ha tomado?" pensamos para nuestros adentros. Entonces, el presentador, que sabe en todo momento donde está el millón de euros, abre una puerta, descubriendo a la cabra que está detrás.*¿Y ahora? ¿Todavía está seguro de su decisión? ¿No le gustaría cambiar de elección ahora que aún está a tiempo? *La pregunta, por supuesto, es: ¿Cómo tenemos más probabilidades de ganar el millón de euros; cambiando de puerta o manteniendo nuestra primera elección? Discutidlo en grupos y redactad un breve párrafo donde reflejéis la opción que creéis correcta y las razones que os han llevado a tal convicción (no hace falta una resolución simbólica, basta que lo expreséis informalmente).*

A propósito no hemos indicado la denominación clásica del problema, pues así evitamos (quizá algo ingenuamente) que puedan consultarlo en alguna fuente externa (internet) sirviéndose de la referencia del nombre "Monty Hall. Quedamos así en libertad de, tras el debate en clase, poder dejarlo como ejercicio a entregar al día siguiente: con un poco de suerte, si lo consultan en su casa habremos creado alguna disputa familiar.

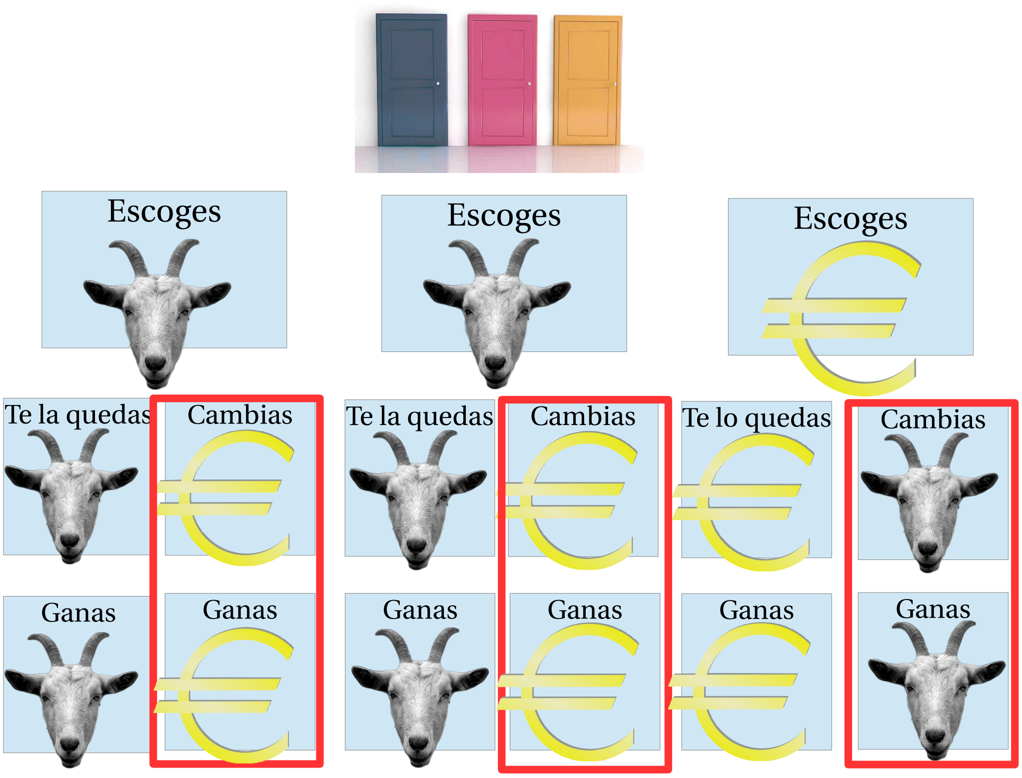
Una vez recogidas las opiniones, sería buena idea que los alumnos expusieran sus puntos de vista y que se debatiera públicamente cada postura, manteniendo el profesor la neutralidad, a ser posible. Finalmente, tras haber creado la suficiente expectación, pasaríamos a la siguiente actividad:

2º) Imagina que un día frotas una lámpara y aparece un genio al que, lamentablemente, no se le pueden pedir deseos, pero que asegura que te puede ayudar para que te toque el gordo de la lotería de Navidad. Tú ya te has comprado un décimo, pero él te dice: "No importa. ¿Ves esta lista **con todos los números de la lotería menos el tuyo***? Pues voy a tachar todos los que* **están en la lista** *y no van a ser el premio gordo." Después de tacharlos todos, deja solamente uno, el 00.000. "Además, me he tomado la libertad de comprarlo. Te propongo un trato; te cambio mi número por el tuyo. Recuerda, no obstante, que el tuyo también podría ser el número del gordo. ¿Qué? ¿Nos los cambiamos?" La pregunta es: ¿Aceptarías el trato del genio? ¿Por qué? ¿Podrías explicar el anterior ejercicio en base al enunciado de éste?*

Desde luego, los alumnos aceptarán el trato, y es de esperar que la mayor parte de los alumnos vean la analogía con el ejercicio de Monty Hall, aunque no es totalmente seguro, dado la fortaleza del prejuicio en contra.

3º) Trata de enunciar el problema en términos de probabilidad condicional y resolverlo cuantitativamente: basta con hallar el valor concreto de la probabilidad de descubrir el millón de euros al final del juego si cambias tu primera decisión.

En estos casos suele ser conveniente una explicación mediante un diagrama de árbol:



Donde se observa claramente que la combinación CAMBIAS y GANAS da como resultado 2 veces el millón de euros y una sola vez la cabra. Por lo tanto, si cambias de decisión tus probabilidades son de 2/3.

Deberíamos comenzar identificando los sucesos y dándoles un nombre adecuado:
Ca1: Elegimos Cabra en el primer intento. (**Ca2**, ídem en el 2º intento)
Eu1: Elegimos el millón de euros en el primer momento. (**Eu2**, ídem en el 2º intento)
Sw: Cambiar la 1º decisión; por tanto, **–Sw**: No cambiar (o persistir) en la 1ª decisión
(Sw procede del inglés *Switch,* y me parece buena elección porque se diferencia mucho tanto de Ca como de Eu). Tratamos de averiguar la probabilidad P(Eu2|Sw).

· Sabemos que P(Ca1)=2/3, P(Eu1)=1/3; por otra parte, P(Sw)=P(-Sw)=1/2. Por la definición de probabilidad condicional, tenemos que:

$$P(Eu2|Sw) = \frac{P(Eu2 \cap Sw)}{P(Sw)}$$

Ahora bien, P(Eu2) es la probabilidad de todos los casos en los que hemos conseguido el millón de euros: entonces $P(Eu2 \cap Sw)$ es la probabilidad de la intersección entre el suceso “Elegimos el millón de euros en el 2º intento” y “Cambiamos la 1ª decisión”, luego es el suceso (compuesto) en el que ambos hechos tiene lugar; Pero, para que cambiando de decisión nos toque el premio, tenemos necesariamente que haber escogido la cabra al principio, suceso cuya probabilidad conocemos: como P(Sw)=1/2, entonces:$P(Eu2 \cap Sw) = P(Ca1 \cap Sw) = P(Ca1) \cdot P(Sw) = 2/3 \cdot 1/2 = 1/3$ y, por tanto,

$$P(Eu2|Sw) = \frac{1/3}{1/2} = \frac{2}{3}$$

La falacia del falso positivo

La mal llamada "paradoja del falso positivo", es una de las variantes más comunes del tipo más general de falacias conocidas como de la "tasa base" o de negligencia de la probabilidad a priori. Todos los subtipos son muy ilustrativos sobre hasta qué punto nuestra intuición natural de la incertidumbre suele resultar incompetente. Tversky y Kahnemann **[29]** nos regalan un ejemplo obtenido —ni más ni menos— que entre estudiantes de medicina de Harvard:

> *Casscell, Schoenberger y Graboys (1978) plantearon a 60 estudiantes y miembros del personal de la Harvard Medical School la siguiente pregunta:*
>
> > Si un test para detectar una enfermedad cuya incidencia es de 1/1000 tiene una probabilidad de emitir un falso positivo del 5%, ¿Qué probabilidades hay de que una persona que ha recibido un resultado positivo realmente tenga la enfermedad, suponiendo que no saben uds. nada sobre los síntomas u otros signos de esa persona?
>
> *La respuesta más generalizada, dada por casi la mitad de los participantes, fue 95%. La respuesta media fue 56%, y solo 11 encuestados dieron la respuesta acertada, es decir, 2%, suponiendo que el test diagnostica sin error a todos aquellos que sí tienen la enfermedad.* [traducción propia]

En nuestro previo análisis sobre dificultades epistemológicas ya analizamos la indencia y las posibles causas de esta dificultad: resta encontrar ejemplos que resulten expresivos para el alumnado que se considera.

El cribado por mamografía

Una de las cuestiones donde la importancia de los falsos positivos pasa a primer plano —generando una controversia que a la fecha de hoy todavía se mantiene— es el análisis preventivo de determinados tipos de cáncer, entre los cuales el de mama parece[1] resultar especialmente problemático. Un buen número de países introdujeron un programa de revisión anual para mujeres de edad generalmente superior a 50 años[2], aunque en algunos países se iniciaron programas también para poblaciones mayores de 40 o de 45. Pero en 2006 comenzaron a aparecer las primeras voces críticas[3] sobre el abuso de las medidas preventivas contra varios tipos de cáncer, el que nos ocupa en particular. Sin embargo, cuando las tesis de Gøtzsche *et al.* co-

1 Lamento la ambigüedad, pero no soy especialista en materia médica.
2 J. Mª Martín Moreno, *El Código Europeo contra el Cáncer. Tercera revisión (2003): insistiendo y avanzando en la prevención del cáncer.* Rev. Esp. Salud Publica (Dic. 2003) [consultado Abril 2015]; 77(6):673-679. Disponible en: http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1135-57272003000600001&lng=es
3 P. C. Gøtzsche, O. J. Hartling, M. Nielsen y J. Brodersen, *La mamografía como método de cribado para detectar el cáncer de mama,* Centro Nórdico Cochrane, 2012 (1ªEd. 2008) [consultado 25 Abril 2015].Traducido por Gloria Córdoba. Disponible en: http://www.cochrane.dk/screening/mamografia-es.pdf

menzaban a constituirse en la posición más aceptada[4], en una fecha muy cercana todavía se pueden ver noticias como ésta[5]:

EL PAÍS

PORTADA | INTERNACIONAL | POL

SOCIEDAD

EDUCACIÓN SALUD CIENCIA MEDIO AMBIENTE IGUALDAD CONSUMO COMUNICACIÓN

ESTÁ PASANDO | La epidemia del ébola | Homofobia | Acoso escolar | Reforma del aborto

El cribado reduce las muertes por cáncer de mama un 31%

- Un estudio holandés revisa el efecto de 20 años de pruebas de detección precoz

EMILIO DE BENITO | Madrid | 21 MAR 2012 - 12:18 CET

Archivado en: Investigación médica Oncología Cáncer mama Cáncer próstata Cáncer mujeres Diagnóstico médico Cáncer Enfermedades Especialidades médicas Medicina Sanidad

0 0 0 0

Enviar Imprimir Guardar

La realización sistemática de pruebas de detección precoz del cáncer de mama (el llamado cribado) ha reducido las muertes por esta patología en un 31%, según un informe que se ha presentado en el congreso europeo dedicado a esta patología que se celebra en Viena. "Comparado con el periodo anterior, de 1986 a 1988, las muertes por cáncer de mama entre mujeres de 55 a 79 años descendieron en 2009 un 31%", ha dicho Jacques Fracheboud, investigador de la Erasmus University Medical Center de Rotterdam (Holanda). "Hemos visto que ha habido un cambio en la tendencia anual. Antes de que comenzaran los programas de cribado, las muertes aumentaban un 0,3% por año, pero después se ha producido un descenso del 1,7% anual. Este cambio también coincide con una reducción apreciable en la proporción de cánceres de mama detectados en un estado avanzado".

Un poco más adelante la noticia se refiere a un estudio[6] donde se puede leer que la noticia está basada en conclusiones “demostradas” por una “microsimulación”.

> *Mrs. Rianne de Gelder […] dijo a la asistencia que ella y sus colegas habían usado una técnica de modelado por ordenador llamada microsimulación […]*

La microsimulación se realiza para comparar la mortalidad entre pacientes que han recibido sólo tratamiento posterior al cáncer; pacientes que han recibido tratamiento más mamografía; y pacientes que no han recibido ninguna de las dos (con datos anteriores a la implan-

4 M. Marzo Castillejo, E. Melús Palazón, B. Bellas Beceiro *et al.*, *Recomendaciones para el cribado del cáncer de mama con mamografía en población de riesgo medio. Actualización PAPPS 2012,* Aten. Primaria (Mayo 2012) [consultado 25 Abril 2015]; 44(6):366---367. Disponible en: http://www.elsevier.es/es-revista-atencion-primaria-27-articulo-recomendaciones-el-cribado-del-cancer-90140925

5 E. de Benito, *El cribado reduce las muertes por cáncer de mama un 30%* [artículo en línea] 21 Marzo 2012, [consultado 25 Abril 2015] Madrid, El País (edición digital) Disponible en: http://sociedad.elpais.com/sociedad/2012/ 03/21/ actualidad/1332328735_675681.html

6 E. Mason, *Breast cancer screening and better treatment both help to save significant numbers of lives* [artículo en línea] 21 Marzo 2012, [consultado 25 Abril 2015] en *EurekAlert!, (eurekalert.org)* American Asscociation for the Advancement of Science, Disponible en: http://www.eurekalert.org/pub_releases/2012-03/eeco-bcs032012.php

tación del programa de cribado). No obstante, y sin conocimientos especializados sobre el tema, se puede objetar que el mero hecho de que haya que recurrir a una simulación indica que no se cuenta con datos ciertos, y , por muy fina que esta sea, sus resultados muy dependientes de la capacidad del modelo utilizado en la simulación para reproducir fielmente el comportamiento del fenómeno en estudio.

> *Mrs. Gelder dijo:* "La efectividad de la mamografía ha sido fuertemente debatida en los últimos dos años. Uno de los argumentos que los críticos esgrimen es que, puesto que en la actualidad los pacientes de cáncer de mama pueden ser tratados de forma muy efectiva con terapia adyuvante[7], los beneficios relativos de la mamografía se vuelven cada vez más pequeños. Nuestro estudio muestra que, incluso cuando se administra terapia adyuvante, el cribado por mamografía (en edades entre 50 y 75) es altamente efectiva en reducir las muertes por cáncer de mama.[...]
> Además, si el cribado pudiera practicarse antes de los 50 años, se podría reducir incluso más la muerte por este tipo de cáncer, incluso cuando los pacientes están recibiendo terapia adyuvante...

Hay varios elementos sospechosos en estas declaraciones, como el restringir la materia de estudio y de juicio exclusivamente a la *mortalidad*, y exclusivamente a la mortalidad *debida únicamente al cáncer específico estudiado:* a primera vista parece que es lo riguroso y exacto, pero Gøtzsche (vid. [3]) nos recuerda el valor del contexto en el análisis de datos:

> *Algunos de los tipos de cáncer y algunos de los cambios celulares tempranos (carcinoma in situ) que se encuentran por el cribado crecen tan lentamente que nunca se convertirían en un verdadero cáncer. Muchos de estos "falsos tipos de cáncer" detectados en mamografías de cribado incluso habrían desaparecido de forma espontánea, si no se hubieran tratado.[...]*
>
> > Si 2000 mujeres son revisadas regularmente durante 10 años, 10 mujeres sanas serán consideradas pacientes con cáncer y recibirán tratamiento innecesariamente. A estas mujeres se les extirpará una parte o toda la mama, y muchas de ellas recibirán radioterapia y, a veces, quimioterapia. El tratamiento de estas mujeres sanas aumenta su riesgo de morir, por ejemplo, por enfermedades del corazón y cáncer.[...]
>
> *[...]las pruebas de detección precoz también dan lugar a sobrediagnóstico y sobretratamiento posterior de mujeres sanas, se extirpará un seno a más mujeres cuando hay exámenes de detección precoz. Además, más mujeres recibirán radioterapia innecesariamente.[...]*
> *Si la radiografía muestra algo que podría ser el cáncer, la mujer debe realizarse exámenes adicionales. En algunos casos, resulta que lo que se ve en la radiografía era benigno, y que por consiguiente, era una falsa alarma.*
>
> > Si 2000 mujeres son revisadas regularmente durante 10 años, alrededor de 200 mujeres sanas experimentarán una falsa alarma. La tensión psicológica hasta que se sabe si existe o no un tipo de cáncer puede ser graves. Muchas mujeres experimentan ansiedad, preocupación, desánimo, problemas para dormir, cambios en las relaciones con la familia, amigos y conocidos, y un cambio en el deseo sexual. Esto puede durar meses y, a largo plazo, algunas mujeres se sienten más vulnerables acerca de la enfermedad y consultan a un médico con más frecuencia.

7 Terapia posterior a un tratamiento contra el cáncer, con el objeto de evitar reincidencias.

[...]la información sobre los beneficios y perjuicios del cribado mamográfico en las invitaciones para participar en cribados y en los sitios web de asociaciones de cáncer y otros grupos de interés es a menudo engañosa.[...]

Los resultados más fiables provienen de ensayos en los que las mujeres han sido elegidas al azar para ser examinadas con la mamografía o no. Alrededor de 600.000 mujeres sanas han participado en estos ensayos (5). La mitad de los ensayos se han llevado a cabo en Suecia. Una revisión de los ensayos suecos de 1993 mostró que el cribado reduce la mortalidad por cáncer de mama en un 29% (6).

Si bien esto parece ser un gran efecto, esto es lo que el 29% significa en realidad. La revision sueca señaló que ***después de 10 años de cribado, esta reducción de la mortalidad era equivalente a salvar una mujer de cada 1000.***

El beneficio de la detección es, por lo tanto muy pequeño. La razón de ello es que, en un período de 10 años, sólo 3 mujeres de cada 1000 contraen cáncer de mama y mueren a causa del mismo. La reducción absoluta de la mortalidad por cáncer de mama fue por lo tanto, sólo del 0,1% (1 de 1000) después de 10 años en los estudios suecos. El cribado durante más de 10 años podría aumentar el beneficio, pero también aumentaría los perjuicios.

La razón por la que sólo describen un período de 10 años es que no hay datos fiables para períodos más largos.[...]

[se practicaron dos revisiones de los datos anteriores, debidas respectivamente a Cochrane y a U.S. Preventive Services Task Force] estas dos revisiones sistemáticas encontraron un efecto en la mortalidad por cáncer de mama que fue solo la mitad del efecto reportado en la primera revisión sueca de 1993. Esto significa que el cribado regular de 2000 mujeres durante 10 años es necesario para salvar a una mujer de morir de cáncer de mama. ***La reducción absoluta de la mortalidad por cáncer de mama fue por lo tanto, sólo el 0,05%.***

Con el cribado mamográfico no reduce el riesgo de morir de cáncer (incluyendo cáncer de mama) (5). Por lo tanto, parece que las mujeres que se realizan una mamografía no viven más que las mujeres que no.

Desde que los ensayos con asignación al azar se llevaron a cabo, ha habido importantes avances en el diagnóstico y tratamiento. Esto significa que el efecto de la detección es menor en la actualidad. De hecho, recientemente estudios rigurosos sugieren que el cribado ya no es eficaz.

Toda la discusión anterior nos parece digna de ser reseñada, en tanto que refleja una problemática real, no hipotética, y en tanto que involucra conceptos y métodos estadísticos accesibles al estudiante de ESO o Bachillerato.

ACTIVIDADES

1º)De la lectura del artículo siguiente se extracta el dato relativo a la incidencia de casos de cáncer de mama al año en España (datos del 2008): http://www.elsevier.es/es-revista-atencion-primaria-27-articulo-recomendaciones-el-cribado-del-cancer-90140925 [8] ***¿Cuál será, para España, y supuesto constante este índice, la probabilidad de contraer cáncer de mama en un período de 10 años?***

Evidentemente, si llamamos C al suceso "contraer cáncer de mama",

$$P\left(C^{10}\right)=1-\left(1-P(C)\right)^{10}=1-\left[1-\frac{61}{100000}\right]^{10}=0{,}00608=0{,}608\%$$

8 *Vid.* 4

2º)En el folleto informativo http://www.cochrane.dk/screening/mamografia-es.pdf [9]*leemos que* "Si 2000 mujeres son revisadas regularmente durante 10 años, alrededor de 200 mujeres sanas experimentarán una falsa alarma"

a)¿Qué sucesos involucra esta declaración?

b)Teniendo en cuenta el enunciado del ejercicio anterior, ¿Cuántos casos cabe esperar de mujeres con cáncer si se revisan 2000 mujeres?

c)¿Qué clase de probabilidad se ha cuantificado, y cuál es su valor? ¿Cuál es la fiabilidad del test empleado, respecto de la posibilidad de falsos positivos?

a)C^{10}= "Padecer cáncer de mama; -C^{10}="No padecer cáncer de mama"; D^{10}="Ser diagnosticada de cáncer de mama al menos una vez en 10 años"

b) $P\left(C^{10}\right)=0{,}0061 \rightarrow 2000 \cdot P\left(C^{10}\right)=2000 \cdot 0{,}0061=12{,}2 \simeq 12 \text{ mujeres}$

c) La probabilidad de que una mujer sea diagnosticada de cáncer (probabilidad *a posteriori*) condicionada a que en realidad no padezca cáncer (probabilidad *a priori*)

$P\left(D^{10} \middle| -C^{10}\right)=\frac{200}{2000-12}=0{,}1006 \simeq 0{,}1=10\%$ luego la fiabilidad del test es $F_p \approx 90\%$

3º)"Si 2000 mujeres son revisadas regularmente durante 10 años, 10 mujeres sanas serán consideradas pacientes con cáncer y recibirán tratamiento innecesariamente" *Esta declaración hallada en el mismo artículo implica que, tras el primer test, las pacientes con un primer positivo son sometidas a otra prueba:*

a)¿Cuál es la fiabilidad de este último test?

b)¿Cuál es la fiabilidad combinada de las dos pruebas en conjunto?

a)Si llamamos D'= “Ser diagnosticado tras el segundo test” y $D^{10} \cap -C^{10}$ =”No padecer cáncer pero haber sido diagnosticada en el primer test”

$P\left(D' \middle| D^{10} \cap -C^{10}\right)=\frac{10}{200}=0{,}05=5\% \rightarrow F'=95\%$

b) $P\left(D' \cap D^{10} \middle| -C^{10}\right)=\frac{10}{2000-12}=0{,}00503 \simeq 0{,}005=0{,}5\% \rightarrow F \cap F'=99{,}5\%$

4ª) El cribado mamográfico es un procedimiento para detectar el cáncer de mama precozmente y así poder tratarlo con más eficacia: Se realizan dos test: si una mujer es diagnosticada según el primer test, se espera un plazo de tiempo de 45 días y se le realiza un segundo test. El primer test tiene una fiablidad (acumulada tras 10 repeticiones) del 90% para falsos positivos, es decir, que diagnostica la presencia del cáncer en mujeres sanas un 10% de las veces que se practica; El segundo test tiene una fiabilidad análoga a la anterior del 95%. Asumiendo que la frecuencia anual de incidencia del cáncer en mujeres españolas es C_m=0,061%, y que el test es perfectamente fiable (100%) en cuanto a falsos negativos,

9 *Vid.* 3

a) ¿Cuál es a probabilidad de que una mujer sana reciba un falso positivo en el primer test, si se revisa anualmente durante 10 años?
b) ¿Cuál es a probabilidad de que una mujer sana reciba un falso positivo tras los dos tests, en las mismas circunstancias?
c) Supongamos que una mujer recibe un resultado positivo en uno de los diez análisis anuales ¿Cuál es la probabilidad de que padezca realmente cáncer de mama?

a) Pongamos nombre a los distintos sucesos:

C^{10}="Contraer cáncer de mama a lo largo de 10 años"
D1="Ser diagnosticada de cáncer de mama en el primer test"
D2="Ser diagnosticada de cáncer de mama en el segundo test"
Se nos pide $P\left(D1 \cap -C^{10}\right)$ y conocemos $P\left(D1 | -C^{10}\right)$; se comenzará calculando $P\left(-C^{10}\right)$

$$P\left(-C^{10}\right) = \left(P(-C)\right)^{10} = \left(1-0{,}00061\right)^{10} = 0{,}99392 \rightarrow P(C^{10}) = 1 - P\left(-C^{10}\right) = 0{,}00608$$

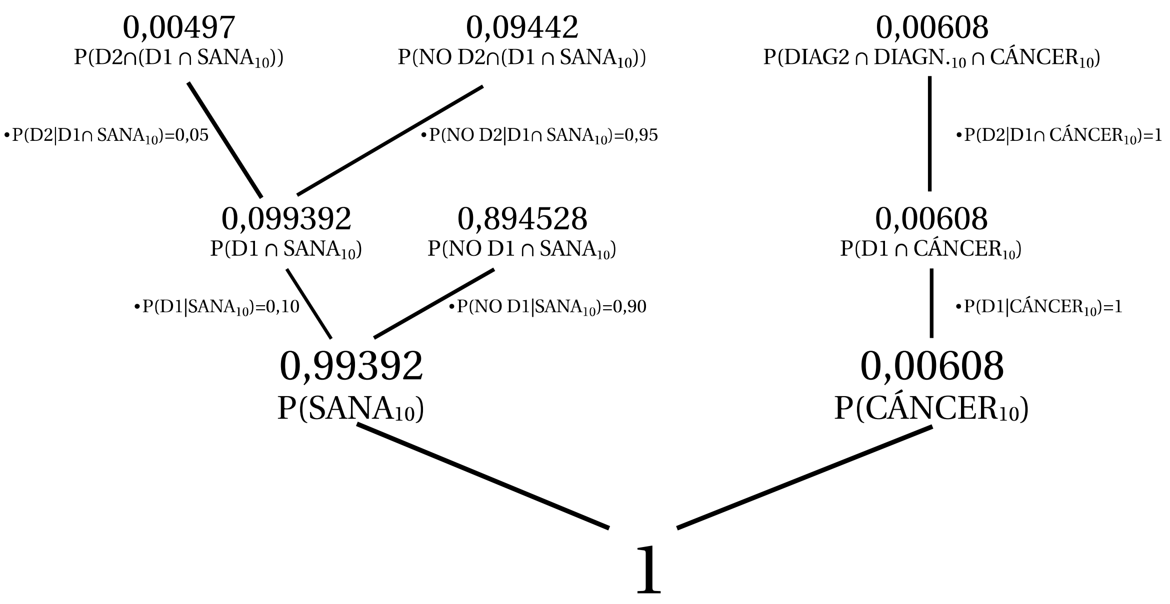
Es decir, la probabilidad de que en un año cualquiera de los 10 hallamos contraído cáncer es la probabilidad total (1) menos la probabilidad de que no lo contraigamos ni una sola vez en esos 10 años.
Decimos que conocemos $P\left(D1 | -C^{10}\right)$ porque —C^{10} implica (o incluye) —C (si estamos sanos durante los diez años, lo estamos cada uno de los años en que se practica el test); sabemos que es igual a 0,1, luego

$$P\left(D1 \cap -C^{10}\right) = P\left(D1 | -C^{10}\right) \cdot P\left(-C^{10}\right) = 0{,}1 \cdot 0{,}99392 = 0{,}099392 \simeq 0{,}1 = 10\%$$

b)Tenemos que averiguar ,$P\left(D2 \cap (D1 \cap -C^{10})\right)$ siendo el suceso D2 = "Ser diagnosticada en el segundo test" y sabiendo que $P\left(D2 | (D1 \cap -C^{10})\right) = 0{,}05$

$$P\left(D2 \cap (D1 \cap -C^{10})\right) = P\left(D2 | (D1 \cap -C^{10})\right) \cdot P\left(D1 \cap -C^{10}\right) = 0{,}05 \cdot 0{,}099392 = 0{,}0049696$$

que es más o menos el 0,5%
Ambos puntos pueden expresarse mediante un diagrama de árbol:



c) Aquí se nos pide también una probabilidad condicional, pero con las condiciones *a priori* y *a posteriori* cambiadas: se nos pide $P\left(C^{10}\middle|D1\right)$ (ó $P(CÁNCER_{10}|D1)$)
Del gráfico podemos inferir que

$$P(D1)=P(D1\cap SANA_{10})+P(D1\cap CÁNCER_{10})=P(D1\cap -C^{10})+P(D1\cap C^{10})=$$
$$=0{,}099392+0{,}00608=0{,}105472$$

y, entonces

$$P\left(C^{10}\middle|D1\right)=\frac{P\left(C^{10}\cap D1\right)}{P(D1)}=\frac{0{,}00608}{0{,}105472}=0{,}057645\simeq 6\%\rightarrow P\left(-C^{10}\middle|D1\right)\simeq 94\%$$

¿Qué sucede, por tanto? ¿No tenía una fiabilidad del 90% el test (en su conjunto)? Sí, exactamente. Pero, pese a que la fiabilidad parezca razonable, el error de medida provocado (el 10%) no es del mismo orden de magnitud que la incidencia de la enfermedad analizada, sino aproximadamente 15 veces mayor: en concreto, 0,1/0,0068 veces mayor. Un buen instrumento de medida debería tener un error de un orden de magnitud como mínimo inferior a aquello que se quiere medir. Lo que sucede, pues, es que la precisión de nuestro instrumento de medida es muy pobre, y se intenta atenuar el hecho con una segunda prueba posterior. Pero en un primer momento, la operación de realizar un test como el descrito (teniendo en cuenta que se repite 10 veces en 10 años) es análoga a la de intentar coger un sólo grano de arena fina entre dos dedos.

Probablemente esta actividad es la más exigente de las que forman parte de nuestra recopilación. Quizá podría haberse simplificado más el enunciado para hacerlo más asequible, pero hemos preferido realizar solo una simplificación mínima, por dos razones: primero, para no desvirtuar completamente el caso real al que hace referencia: es preciso que el alumno vea que con medios algo difíciles, pero no inalcanzables por su parte, puede dar cuenta de un problema muy próximo al problema real. En segundo lugar, creemos conveniente, al menos de forma esporádica, ampliar el horizonte de la enseñanza de la estadística y la probabilidad aunque sea a costa de contradecir —puntualmente— todo lo dicho hasta ahora en relación a la enseñanza secundaria de estas disciplinas. De forma más explícita: creemos que el alumno, en su concepción de las ciencias, o sencillamente en su imaginación, no debe quedarse solo con la idea de que el estrecho margen del currículo abarca toda la probabilidad y estadística. Debe entrever la existencia de un campo vasto, valioso y potente de conocimiento, y, por qué no decirlo, de alto valor estético más allá de lo arbitrario o pedante (y la enseñanza secundaria muy a menudo lo es) que pueda parecerle a veces el temario, aunque quizá no pueda todavía aprovecharlo. Nos consta que este ejercicio será difícil para él: quizá sería mejor que el docente convietiera esta actividad en una clase magistral especial.

Se habla a menudo de la “atención a la diversidad” en relación a ESO y Bachillerato. Sin embargo, la probabilidad y la estadística, las matemáticas de secundaria en su conjunto, también pueden ser fuente de interés para algunos alumnos: al fin y al cabo, alguno de ellos *serán matemáticos*. Por tanto, hurtar completamente a todos los alumnos de, al menos, un vistazo casual de mayor alcance sobre su naturaleza y sus métodos puede perjudicar también a cierta clase de alumnos (que también requieren de “atención”).

La paradoja de Simpson

La estadística es el análisis de datos en un *contexto* [**37**], aunque estos datos sean la mayor parte de las veces cuantitativos. La importancia de analizar la procedencia y los tratamientos matemáticos a los que se ha sometido a los datos, es decir, la importancia de realizar un análisis cualitativo de los datos nunca debe perderse de vista. En suma, la Estadística no es una *caja negra* procedimental que recibe un determinado *input* numérico y que, tras aplicar unas determinadas transformaciones obtiene un determinado *output igualmente* numérico como respuesta. En estadística los números tiene significado, como en las ciencias factuales. De ello que los peligros de aplicar fórmulas o algoritmos ciegamente abunden, como por ejemplo en el caso que nos ocupará a continuación.

De una propiedad aritmética tan sencilla (y fácilmente demostrable)como esta

$$\begin{cases} \dfrac{A}{B} > \dfrac{a}{b} \\ \dfrac{C}{D} > \dfrac{c}{d} \end{cases} \not\Rightarrow \frac{A+C}{B+D} > \frac{a+c}{b+d}$$

se sigue uno de los más sorprendentes y elusivos resultados de la estadística, conocido como la paradoja de Simpson[1], así nombrada por primera vez por Blythe (1973), aunque fue estudiada en primer lugar por Pearson, seguido por Yule . Una exposición de un ejemplo de la paradoja puede ser la que sigue ([**22**], cap. 30) :

> *En una universidad 1 dada 200 de cada 1000 varones, y 150 de cada 1000 mujeres estudian económicas.En la universidad 2, 30 de cada 100 varones y 1000 de cada 4000 mujeres estudian económicas. Luego en la Universidad 1 20% de hombres y 15% de mujeres estudian económicas, y en Univ. 2, el 30% frente al 25%). Sin embargo, cuando las cifras de estudiantes de las dos univrsidades se combinan resulta que 230 de 1100 hombres (20,9%) y 1150 de 5000 mujeres (23%) estudian económicas. Luego, después de todo, ¡Hay un mayor porcentaje de mujeres estudiando económicas que de hombres!*[trad. propia]

Como se puede observar, este caso ejemplifica la propiedad aritmética inicial. Es de notar que Pearson dedica palabras categóricas al problema, en relación con asuntos de mayor alcance afectados por la paradoja en cuestión, esto es, el dilema correlación-causación (*vid. supra*):

> *Nos vemos forzados a la conclusión de que una mezcla de grupos heterogéneos, en el que cada uno de los cuales no muestra en sí mismo ninguna correlación orgánica, mostrará una mayor o menor cantidad de correlación.Esta correlación debería llamarse propiamente espúrea, mas, al ser imposible garantizar la absoluta homogeneidad de ninguna comunidad, nuestros resultados sobre correlación*

1 E.H. Simpson, *The Interpretation of Interaction in Contingency Tables* [artículo en línea], 1951 (consultado Junio 2015) Journal RSS Meth. Vol.13, nª2, pp. 238-241 Disponible en [previo registro]: http://www.jstor.org/stable/2984065?origin=JSTOR-pdf

serán siempre susceptibles de error, cuya magnitud no puede ser predicha. A aquellos que persisten en contemplar la correlación como causa-efecto, el hecho de que esta correlación se puede producir entre dos caracteres incorrelatados, mediante la elección de una mezcla artificial de dos razas estrechamente similares, puede resultar un golpe muy duro. [trad. propia]

Palabras que no solo hablan en contra del uso abusivo de la estadística, sino en favor del constante análisis cualitativo de los datos. que ningún software armado con todo el arsenal matemático disponible podrá realizar.

ACTIVIDADES

1º) Lee el siguiente artículo[2]*:*


Matifutbol

Football agains
Fútbol contra el
Futbol contra

Y ahora, ¿quién tira el penalti?

PRIMERA PARTE

Final de la Copa Mundial de Clubes de la FIFA 2012. Se han clasificado para la final que se celebra en Japón el Sport Club Corinthians paulista, campeón de la **Copa Libertadores de América** y el Chelsea FC de Londres, campeón de la **Liga de Campeones de la UEFA**.

VS.

Es el **minuto 89**, el resultado es de empate a 2 goles, y el árbitro acaba de pitar un **penalty** a favor del Chelsea. El entrenador Roberto Di Matteo tiene que **decidir qué jugador** será el que lance la pena máxima. Para ello consulta a su asistente técnico, **Steve Holland**, sobre qué jugador tiene **mejores estadísticas** en el **lanzamiento de penaltis**.

2 H. López Arroyo, *Y ahora, ¿quién tira el penalti?*[artículo en línea] en *www.matifutbol.com* Disponible en [consultado Junio 2015]: http://www.matifutbol.com/es/penalti.html [material bajo licencia Creative Commons Atribución-CompartirIgual 3.0 Unported http://creativecommons.org/licenses/by-sa/3.0/deed.es_AR] encontrado vía A. Gordaliza.

Steve saca la tabla adjunta, según la cual este año, en la Premier League, Frank Lampard ha lanzado **18** penaltis y ha marcado **12**, mientras que Fernando Torres ha tirado **10** penaltis y sólo ha marcado **6** (un **66.67%** de aciertos de **Lampard** frente al **60%** de **Fernando Torres**). En competiciones internacionales, **Lampard** ha lanzado **9** penaltis y ha marcado **5** goles (**55,58%**) mientras que **Fernando Torres** ha lanzado **2** y ha marcado **1** gol (**50%**).

Estadística de penaltis lanzados	Esta temporada			Otras temporadas		
Premier League	Lanz.	Goles	%	Lanz.	Goles	%
Frank Lampard	18	12	66,67%	12	11	91,67%
Fernando Torres	10	6	60,00%	20	18	90,00%
Competiciones internacionales	Lanz.	Goles	%	Lanz.	Goles	%
Frank Lampard	9	5	55,56%	4	3	75,00%
Fernando Torres	2	1	50,00%	11	8	72,73%

Y si acudimos a las estadísticas de otras temporadas, los números también son siempre favorables a **Frank Lampard** (**91,67% > 90%, 75% > 72,73%**).

Por tanto, **en función únicamente de los datos** facilitados por su ayudante, **Di Matteo** decide que sea **Lampard** quien lance el penalti.

¿Crees que habrá tomado la decisión correcta?

a) Preguntamos lo mismo que en el artículo. ¿Crees que Di Matteo tomó la decisión correcta?¿Por qué?

A primera vista parece que sí, la decisión es correcta, como haremos ver al alumno mostrándole que de las cuatro maneras en que se comparen los porcentajes, la situación siempre es favorable a Lampard (sus porcentajes de acierto son mayores que los de Torres en todos os casos). Por si quedara alguna duda, podrían incluso sumarse los lanzamientos y goles correspondientes a la temporada actual y a las otras, resultando las siguiente tabla:

Estadística de penaltis lanzados	Esta temporada			Otras temporadas		
Premier League	Lanz.	Goles	%	Lanz.	Goles	%
Frank Lampard	18	12	66,67%	12	11	91,67%
Fernando Torres	10	6	60,00%	20	18	90,00%
Competiciones internacionales	Lanz.	Goles	%	Lanz.	Goles	%
Frank Lampard	9	5	55,56%	4	3	75,00%
Fernando Torres	2	1	50,00%	11	8	72,73%
Total	Lanz.	Goles	%	Lanz.	Goles	%
Frank Lampard	27	17	62,96%	16	14	87,50%
Fernando Torres	12	7	58,33%	31	26	83,87%

Donde se puede comprobar que, aun realizando la operación antedicha, los porcentajes de acierto siguen favoreciendo la elección de Lampard sobre Torres.

b) Las tablas que hemos realizado se denominan en estadística "tablas de contingencia". Tan sólo quedaría, a la luz de los datos disponibles, una última compro-bación ¿Cuál sería? Elabora la tabla de contingencia correspondiente.

Nos quedaría como última comprobación sumar todos los lanzamientos y goles de los jugadores respectivos, tanto de la presente y otras temporadas, como de las distintas li-gas, resultando una tabla de contingencia de la siguiente forma:

Estadística de penaltis lanzados	Esta temporada			Otras temporadas			Todas las temporadas		
Premier League	Lanz.	Goles	%	Lanz.	Goles	%	Lanz.	Goles	%
Frank Lampard	18	12	66,67%	12	11	91,67%	30	23	76,67%
Fernando Torres	10	6	60,00%	20	18	90,00%	30	24	80,00%
Competiciones internacionales	Lanz.	Goles	%	Lanz.	Goles	%	Lanz.	Goles	%
Frank Lampard	9	5	55,56%	4	3	75,00%	13	8	61,54%
Fernando Torres	2	1	50,00%	11	8	72,73%	13	9	69,23%
Total	Lanz.	Goles	%	Lanz.	Goles	%	Lanz.	Goles	%
Frank Lampard	27	17	62,96%	16	14	87,50%	43	31	72,09%
Fernando Torres	12	7	58,33%	31	26	83,87%	43	33	76,74%

Como se puede comrpobar, tanto en el resultado total de lanzamientos y goles de todas las temporadas de las ligas consideradas, como cada suma parcial de todas las tempora-das de cada liga por separado, esta vez dan ventaja a Torres frente a Lampard. ¿ Cómo es posible? Sencillamente sucede esto (*vid. supra*):

$$\begin{cases} \dfrac{A}{B} > \dfrac{a}{b} \\ \dfrac{C}{D} > \dfrac{c}{d} \end{cases} \not\Rightarrow \frac{A+C}{B+D} > \frac{a+c}{b+d}$$

Sería preciso explicar al alumno que esta es una propiedad de los números racionales, y que por tanto es una dificultad matemática, más que puramente estadística, de la que hay que ser consciente al analizar los datos e inferir proposiciones sobre los mismos. Asimismo, la comprobación de la homogeneidad de los datos, en última instancia, no tiene una solución *estrictamente matemática*, por lo que siempre es necesario explorar los datos teniendo en cuenta su *contexto* (qué categorías miden, como se han obtenido los datos y elaborado sus resúmenes, etc.)

Hemos dicho que no tiene una solución matemática válida para todos los casos, pero hay modos de prevenirla: desagregando los datos según cualquier clase de categoría de la que tengamos constancia que puedan estar afectando a las variables construyendo tablas de contingencia, o, como en el ejemplo de la actividad propuesta, también *agre-gando los datos* se llega a resultados más válidos que con los desagregados. Ahora bien, es preciso tener en cuenta, como dice Pearson, que la posibilidad de estar obviando una variable importante que afecte a las estudiadas subyace a todo tipo de tablas de contin-gencia, con lo cual la detección de asociaciones por correlación o regresión debe enten-derse, no como inválidas, sino como paso previo de identificación y comprobación del comportamiento general de un conjunto de variables, a la espera de la formulación de un modelo teórico de mayor alcance que el de un experimento particular.

¿Correlación implica Causalidad?

Ya hemos avanzado en el anterior capítulo algunos de los comentarios suscitados por la cuestión acerca de si dos variables correlatadas son también (¡Y en qué orden!) causa-efecto la una de la otra. Como distinción obvia general podemos decir que si sabemos que, de dos variables X e Y, X representa a un fenómeno que es causa del fenómeno representado por Y, entonces X e Y están *también* asociadas (y la correlación mide la más sencilla de esas asociaciones). Pero el enunciado recíproco no es cierto necesariamente, lo cual es la fuente de todos los problemas; se supone que analizamos dos variables porque no sabemos si una es causa de la otra, si una *explica* a la otra. Si lo supiéramos de antemano no realizaríamos nunca ningún análisis de correlación o de regresión. Y de todas formas, el hecho de que dos variables se hallen *asociadas* o correlatadas no se puede, sencillamente, obviar: es indicativo de alguna clase de relación entre los fenómenos representados por esas variables, aun indirecta. ¿Cuál es el estatus lógico y epistemológico de la correlación-regresión, pues?

Aunque voces muy autorizadas se han pronunciado al respecto, no solo dentro del campo de la estadística —pues la cuestión afecta también en buena medida a los fundamentos de conceptos como observación, experimento e incluso causación, bastante vapuleado desde el pasado siglo[1], aunque más recientemente rehabilitado— se pueden exponer ejemplos de mayor valor ilustrativo que las generalizaciones **[19]** :

> *Para niños en edad escolar, la talla del zapato está fuertemente correlatada con las habilidades lectoras. Sin embargo, aprender palabras nuevas no hace que tus pies se hagan más grandes* [la recíproca tampoco es cierta]. *En lugar de ello, hay un tercer factor involucrado: la edad. A medida que los niños crecen, aprenden a leer mejor y los zapatos se les quedan pequeños. (en términos estadísticos, la edad es un* ***factor de confusión*** (confounder*).En el ejemplo, el factor de confusión es fácil de señalar. Pero a menudo, esto no es tan fácil. Además,* ***el cálculo del coeficiente de correlación no previene*** *contra la acción de terceros factores.* ***La correlación mide asociación. Pero asociación no es lo mismo que causación***. [trad. propia]

Según el anterior y otros autores, la existencia de estos factores de confusión o variables ocultas (*lurking variables*) afecta decisivamente a los estudios meramente observacionales, y sólo pueden controlarse —hasta cierto punto— mediante el ajuste de variables *conocidas* que puedan afectar a la relación: en estudios sobre reeacciones a fármacos, por poner un caso, se llevan a cabo ajustes por edad, por género, por nivel de ingresos, por grupo racial,

1 Vapuleado desde un punto de vista de orden superior al que se usa comunmente en estadística, o en las ciencias factuales; por ejemplo, caundo B. Russell dice: *"La ley de la causalidad, como muchas cosas que aceptan los filósofos, es una reliquia de una era pasada que sobrevive, como la monarquía, porque se cree erróneamente que es inocua." (On the notion of cause, 1913)* ataca la noción filosófica de causa; el científico puede, aun de forma operacional, entender que un fenómeno representado por la variable X **causa** otro representado por Y cuando, manteniendo invariantes las demás condiciones del sistema al que pertenecen ambos fenómenos, a las variaciones de valor de X le siguen variaciones del valor de Y.

por entorno cultural, etc. No obstante, no todas las variables en juego tienen por qué ser conocidas. De esta manera los primeros estudios, simplemente observacionales, sobre la posible relación causa-efecto entre el tabaquismo y el cáncer de pulmón no eran del todo concluyentes (**[19]** , p. 20):

> *Algunos epidemiólogos habían encontrado una fuerte asociación entre fumar y padecer cáncer de pulmón: los que más fumaban se veían aquejados de cáncer en mayor proporción que los que fumaban menos, y estos últimos en mayor proporción que los que no fumaban. Los epidemiólogos concluyeron que, efectivamente, la asociación* [correlación] *existía porque fumar* ***es causa*** *de aparición del cáncer de pulmón.Sin embargo, algunos estadísticos – incluyenda a Sir R.A. Fisher– pensaban que la asociación podía ser explicada por algún factor de confusión.*
> [...] *Por ejemplo, supongamos que hay un gen que incrementa el riesgo de padecer cáncer de pulmón. Pues bien, si este gen también fuera responsable de una mayor propensión a fumar, entonces tendríamos un claro factor de confusión. [...]*

Es decir, si el gen en cuestión fuera a la vez causa del hecho mismo de fumar y de padecer cáncer de pulmón, entonces el sólo hecho de fumar no provocaría cáncer de pulmón. Una persona con ese gen podría fumar o no, o dejar de fumar, y ello no influiría en nada en su propensión a padecer cáncer de pulmón.

> *La idea es muy sutil: un gen que causase el cáncer pero que no estuviera relacionada con el hecho de fumar no sería un factor de confusión, y sería irrelevante a la discusión, porque no daría cuenta de los hechos: la asociación entre fumar y padecer cáncer. La "hipótesis constitucional" de Fisher explicaba la asociación en base a factores de confusión genéticos* [no directamente observables]

La tesis de Fisher hubo de ser refutada posteriormente mediante estudios sobre gemelos (individuos de idéntido patrimonio genético), pero lo que sí demuestra hasta qué punto una variable oculta puede, efectivamente, estar oculta. El único remedio para librarse de estas *lurking variables* estriba en realizar experimentos bien diseñados mediante muestreo aleatorio y creación de grupos de control.(**[37]**, p. 88):

> *El experimento comparativo aleatorio* (randomized comparative experiment) *es una de las ideas más importantes en estadística. Está diseñado para permitirnos inferir conclusiones de la forma causa-efecto.[...]*
> *· La aleatorización produce grupos de sujetos que deben ser similares a todos los efectos antes de que apliquemos el tratamiento.*

Es decir, están eliminadas todas las posibles variables ocultas, *conocidas o desconocidas*, por el hecho de estar igualmente distribuidas entre todos los individuos sujetos a experimentación.

> *· El diseño comparativo nos asegura que ninguna influencia fuera de la experimental opera de igual forma en todos los grupos.*
> *· Por todo ello, una diferencia en la variable respuesta necesariamente ha sido debida a efectos del tratamiento.*

Ello no obstante, por razones prácticas y éticas, a menudo no es posible llevar a cabo esta clase de experimentos, y no queda más remedio que tratar de controlar las variables que conozcamos se hallen en juego de un estudio observacional y *ajustar los datos* debidamente a las mismas. La evidencia resultante, de todas formas, siempre será más debil que la obtenida mediante experimentación bien diseñada.

Llegados a este punto debemos hacer una acotación: seguimos nuevamente a Moore **[37]** **[10]**, en el sentido en que nos parece necesario que el alumno, parte del cual recibirá en 2º de Bachillerato docencia sobre los conceptos de intervalo de confianza, p-valor y test de hipótesis, conozca al menos en la enseñanza secundaria obligatoria un planteamiento de los conceptos más básicos de diseño de experimentos (muestreo aleatorio, tanto del grupo de estudio como del de control, elección de las variables, de los criterios y medios de medición de las mismas, comparación de datos, diferencia significativa, etc.) sin los cuales los contenidos curriculares antes reseñados se encuentran faltos de su principal razón de ser (como tantas cosas en la enseñanza media de matemáticas). No hay que olvidar que el germen de tales conceptos se halla en la obra de R.A. Fisher, no por casualidad titulada *Diseño de Experimentos.*

ACTIVIDADES

1º) Lee la siguiente tira cómica[2]:

http://es.xkcd.com/strips/correlacion/ Original en http://xkcd.com/552/

a)Explica en tus propios términos qué quiso decir el personaje que habla en último lugar.

Es probable que el alumno esté acostumbrado a un tipo de humor un poco más, digamos, básico. Pero esto no es necesariamente mala cosa, por que en ese caso se verá obligado a reflexionar sobre lo que se enuncia ¿Por qué no está seguro el personaje de que la asignatura de estadística le haya ayudado (o aprovechado)? Por que, si bien su desempeño en estadística y el hecho de haber recibido una asignatura del tema son variables correladas, el personaje ya no está seguro de que ello implique que la causa de haber mejorado su conocimiento hayan sido las clases que ha recibido. Sólo reconoce que la

2 *Correlación*, en http://es.xkcd.com/ Disponible en [consultado Junio 2015]: http://es.xkcd.com/strips/correlacion/ (Traducción del original: *Correlation*, en http://xkcd.com/ Disponible en [consultado Junio 2015]: http://xkcd.com/552/)Licencia: http://creativecommons.org/licenses/by-nc/2.5/

correlación señala una probabilidad, un "quizá". Pero, ¿Y si ambos sucesos pueden ser explicados por una variable oculta? Si los estudios son universitarios, por ejemplo, la asignatura de estadística ha podido ser elegida por el personaje entre otras muchas opcionales, lo cual quiere decir que el tema le interesaba *antes de recibir la docencia.* ¿No podría ser que fuera *este interés* por el asunto el que ha causado los dos hechos: el haber elegido la asignatura y el haber aumentado sus conocimientos en estadística?

b) ¿Se te ocurre alguna manera estadística de que el personaje se asegure de que las clases de estadística le ayudaron?

Efectivamente, en lugar de observar una mera correlación, podría diseñar y ejecutar un experimento aleatorio comparativo: escogería una muestra aleatoria de alumnos de tamaño adecuado, de los cuales segregaría un grupo de control que no recibiría el curso. Después, compararía los resultados de ambos grupos y trataría de encontrar una diferencia estadísticamente *significativa* (es decir, de la que pudiera decirse que es muy improbable que fuera provocada por el azar) en orden a probar la hipótesis de que el grupo que recibió las clases ha mejorado sus conocimientos en estadística respecto al grupo que no ha recibido el curso.

2º)Lee el siguiente extracto de noticia[3]*:*

EL MUNDO — Edición España · Versión Clásica

SECCIONES | Salud | Sida y hepatitis | Cáncer | Nutrición | Biociencia

PEDIATRÍA Efectos beneficiosos

La lactancia materna prolongada aumenta el cociente intelectual a largo plazo

EFE > Londres

Actualizado: 18/03/2015 **11:00 horas**

f 14514 | 384 | +

La lactancia materna prolongada mejora el rendimiento escolar, aumenta el cociente intelectual en el adulto y se relaciona con unos ingresos altos en el futuro, según un informe publicado hoy en la revista británica The Lancet.

El estudio, que analizó a un grupo de casi 3.500 personas desde su nacimiento hasta que cumplieron los treinta años, muestra la primera evidencia de que la lactancia materna durante más de doce meses tiene un gran impacto en el desarrollo cognitivo.

3 EFE, *La lactancia materna prolongada aumenta el cociente intelectual a largo plazo* [artículo en línea], Londres, 18 Marzo 2015 [consultado Junio 2015], El Mundo. (ed. digital *http://www.elmundo.es*) Disponible en: http://www.elmundo.es/salud/2015/03/18/55094cab22601d5a6c8b458c.html

"El efecto de la lactancia materna prolongada influye en el desarrollo cerebral y la inteligencia de los niños, pero estos efectos persisten también en la edad adulta", señaló Bernardo Lessa Horta, de la Universidad Federal de Pelotas, en Brasil.

Según el informe, un niño que reciba lactancia materna un mínimo de un año conseguirá con treinta años un cociente intelectual mejor, tendrá 0,9 años más de escolaridad y cobrará un salario de 98 euros más al mes que aquellos que no reciban la lactancia materna durante tanto tiempo.

Horta y su equipo analizaron los datos de cerca de 6.000 bebés que nacieron en el año 1982 en el municipio brasileño de Pelotas.

Entre ellos, 3.493 realizaron un test de cociente intelectual a los 30 años.

Los expertos dividieron a los participantes en cinco grupos, basándose en el tiempo en el que habían recibido la lactancia materna y controlaron diez variables sociales y biológicas que pueden contribuir al incremento del cociente intelectual.

Algunos de esos factores son los ingresos familiares, el nivel de escolarización de los padres, la genética, la edad de la madre y si ha fumado durante el embarazo, el peso del bebé y el tipo de parto.

Los autores señalan que la leche materna tiene una composición única en la que destacan los ácidos grasos de cadena larga, "que son esenciales para el desarrollo cerebral".

"Hemos encontrado que la lactancia materna prolongada está ciertamente relacionada con el cociente intelectual en la edad adulta, lo que refleja que la cantidad de leche materna consumida juega un papel importante", dijo Horta.

Según Erik Mortensen, de la Universidad de Copenhague, en Dinamarca, "con la edad, los efectos del desarrollo temprano pueden diluirse debido a factores ambientales posteriores, o mejorar como consecuencia de los logros educativos o profesionales".

Sin embargo, añade Mortensen, este estudio "sugiere que los efectos de la lactancia materna prolongada en el desarrollo cognitivo persisten en el adulto y son importantes para la salud".

Los expertos concluyen que este estudio todavía necesita reforzarse con otras investigaciones futuras que se centren en los efectos a largo plazo derivados de la lactancia materna prolongada.

a) Fíjate únicamente en el titular y compáralo con el título del estudio al que se refiere la noticia, que es el siguiente:

"Asociación entre lactancia e inteligencia, logros educativos e ingresos personales a la edad de 30 años. Un estudio de cohorte de nacimientos prospectivo en Brasil[4]"

¿Qué diferencias en relación a conceptos estadísticos y de diseño experimental puedes señalar?

En primer lugar, el titular indica claramente que la relación entre la duración del período de lactancia y la inteligencia posterior del adulto de 30 años es de *causa-efecto*. Es decir, mantener la lactancia del niño durante más tiempo hace al adulto posterior más inteligente. En cambio, el título de artículo se conforma con una palabra más modesta: "asociación". Es decir, que configurados como variables los sucesos "lactancia prolongada" e "inteligencia " (aunque en el artículo se consideran otras dos más, a saber: logros educativos e ingresos personales), se encuentra que estos se hallan positivamente (y significativamente) asociados (ya sea por correlación un otro modelo de regresión)
En segundo lugar, el título del artículo define la clase de estudio de que se trata: de cohorte prospectivo. En la noticia, en cambio, esto no se llega a explicitar.

b) Una vez completada la lectura del artículo, responde a lo siguiente: ¿El redactor quiere dar a entender que el estudio es un experimento? ¿Se deduce del artículo que el estudio sea realmente *un experimento? ¿Por qué? (Puedes buscar información por la red sobre en qué consiste un "estudio de cohorte prospectivo")*

El artículo, dirigido a un público general, está redactado de manera que se quiere dar respetabilidad científica a lo que se afirma, aunque en ningún momento se llega a decir que el estudio constituya un experimento. Pero una lectura algo detenida del artículo nos informa de que este no puede ser el caso, por varias razones:

1º)No se dice nada sobre el método de elección de la muestra, que podría no ser aleatorio.

2º) Aunque la muestra se divide en cinco grupos distintos entre los cuales realizar la comparación, con lo cual podría decirse que el período de lactancia equivale proporcionalemente a diversos grados de "tratamiento" (o exposición a determinado agente), estos no se han elegido *aleatoriamente*, sino que se han *autoseleccionado:*

Los expertos dividieron a los participantes en cinco grupos, basándose en el tiempo en el que habían recibido la lactancia materna...

4 C.G. Victora, B. Lessa Horta, C. Loret de Mora *et al.*, *Association between breastfeeding and intelligence, educational attainment, and income at 30 years of age: a prospective birth cohort study from Brazil* [artículo en línea], Abril 2015 [consultado Junio 2015], *The Lancet G. H.* Vol. 3, No. 4, e199–e205 en *www.thelancet.com* Disponible en: http://www.thelancet.com/journals/langlo/article/PIIS2214-109X%2815%2970002-1/fulltext DOI: 10.1016/S2214-109X(15)70002-1

Del texto así parece deducirse, y en todo caso parece difícil que las familias de los participantes se hubieran prestado a acortar o extender el período de lactancia a indicación de los autores del experimento.

3º) Los individuos a los que se ha podido realizar un seguimiento, es decir, aquellos que finalmente, 30 años después de su elección realizaron el test de inteligencia, son solo el 3493/6000 del total de la muestra, es decir, el 58,2%. Si tenemos en cuenta que la falta de respuesta, o dato faltante, es uno de los problemas mayores de esta clase de estudios, el hecho de que solo se realizaran test de inteligencia (entre otras mediciones) de un porcentaje tan bajo solo puede explicarse por medio de la negativa de 41,8% a realizar el test. Lo cual significa que la muestra está *doblemente* autoseleccionada.

4º) Finalmente, en el artículo se dice:

> [los expertos] *controlaron diez variables sociales y biológicas que pueden contribuir al incremento del cociente intelectual.*

este análisis por factores de influencia o confusión no sería necesario si el estudio fuera un experimento, donde la aleatoriedad de la muestra garantiza, dentro de márgenes medibles, que tanto estas como otras posibles variables ocultas cualesquiera no están interviniendo en los resultados.

De todo ello se puede deducir que el estudio es observacional, lo que implica que sus conclusiones son mucho más débiles que aquellas que hubieran resultado de un experimento bien diseñado y ejecutado, y especialmente débiles en lo que refiere a la inferencia de causalidad.

Sin embargo, convendría mencionar frente al alumno que por diversas razones, a menudo esta experimentación es impracticable, por lo que hay que conformarse con estudios observacionales de este tipo, cuyos resultados deben cumplir, no obstante, ciertas condiciones generales: la muestra debe tener el tamaño necesario, los datos deben ser bien ajustados según un análisis factorial adecuado (como podría ser este caso, con las diez variables que se dice se han considerado) y los resultados deben exhibir una asociación fuerte (según algún modelo de regresión) que suponga una diferencia estadísticamente significativa entre los valores medidos correspondientes a cada grupo analizado.

c) Ya hemos convenido que se trata de un estudio observacional; no obstante, ¿Son fiables, dentro de las posibilidades de ese tipo de estudio, las conclusiones a las que se llega? ¿Podrías señalar algún punto discutible en la metodología adoptada, que hiciese que incluso esa "asociación" no fuera concluyente? Razona tu respuesta.

Además de las mencionadas más arriba, se podrían objetar una serie de puntos al diseño del estudio, tal y como lo describe el redactor del artículo:

1º) Como se dijo antes, no está claro si ya de principio la muestra se eligió aleatoriamente o era una muestra de conveniencia: para empezar, no sabemos hasta qué punto el municipio de Pelotas, Brasil, es suficientemente representativo o no para hacer inferencias sobre el ser humano en general o sobre el mismo Brasil en particular. Es de suponer que la selección se hizo en maternidades de hospital, que es el lugar donde la recolección de datos y la concertación de la partipación en un estudio es más viable (tampoco tenemos datos sobre la posible tasa de rechazo de participación en el estudio dentro de la misma muestra original), pero por eso mismo es posible que la muestra peque de sesgada desde un momento incluso previo a los que explicamos en la anterior actividad, ya que hay muchas probabilidades de que, considerando las características socioeconómicas de Brasil, mucho nacimientos no tengan lugar en hospitales. O que la elección de hospitales, teniendo en cuenta la extensión de la sanidad privada en ese determinado país, no se haya ponderado como factor. En España, por ejemplo, es más fácil que cooperen con la universidad los hospitales públicos que los privados; por tanto, si la conveniencia marcaba como objetivo a los primeros, puede que la muestra esté sesgada en ambos sentidos, los de la poblaciuón no hospitalizada y la de renta más alta que acude a hospitales privados. Esto podría haberse señalado desde el principio, pero como el estudio tiene pretensiones clínicas, no habría resultado de recibo, por razones obvias, titular el artículo de la forma: *La lactancia materna prolongada aumenta el cociente intelectual a largo plazo de los niños de clase media*[5]*"*

2º)Entre los factores de influencia, se habla de la *"genética".* Aunque no se presente el contenido al alumno, de la lectura del artículo original (*vid. nota*[4]) se deduce que esta expresión se refiere a la extracción por grupo étnico de cada sujeto (esto se puede dejar claro al alumno en un primer momento) Lo cual nos deja la posibilidad de señalar un defecto, no en los factores analizados, sino en los que no lo están; si se mide el cociente intelectual del sujeto, ¿No tiene éste nada que ver con el cociente intelectual de los padres, el cual no aparece entre los factores analizados? ¿Y no podría suceder que este cociente intelectual de los progenitores (y no solo su índice de escolarización y nivel de ingresos, que sí se analizan) influyese en la decisión de prolongar el período de lactancia del bebé? Si así fuera, entonces este hecho, el cociente intelectual de los prgenitores, explicaría ambas variables analizadas; sería la variable oculta que explica a las otras dos, y entonces la lactancia más o menso larga no sería la causa de un mejor o peor cociente intelectual.

3º) Finalmente, se debe hacer notar al alumno que, pese a lo dicho anteriormente, no es lo mismo "diferencia estadísticamente significativa" que "diferencia significativa" a secas. La primera se puede medir con suficiente precisión matemáticamente, y expresa que entre los grupos analizados se da una diferencia, efectivamente, que no es producto del azar, de la dispersión de los datos recogidos; es una diferencia sustantiva al fenómeno, no contingente. Pero la última depende del contexto de la investigación, de los aspectos cualitativos de la variable que se quiere medir. Es evidente que

5 En realidad, el mismo cociente de inteligencia ha sido criticado históricamente por la misma razón: se supone que debe medir una cualidad innata y, curiosamente —sin importar qué método de medición se haya adoptado— siempre aparece correlatado positivamente con los ingresos del individuo (o de la familia, si este es menor de edad). En el caso de este estudio, ¡Incluso se llega a hacer explícito tal extremo! *Algunos de esos factores son* ***los ingresos familiares****...*

en física una diferencia sustantiva, aunque pequeña, es un fenómeno de la misma importancia científica que uno grande: si no fuera así, no existiría la moderna física de partículas y nos conformaríamos con explicar los fenómenos macroscópicos en los que está involucrada la materia, cosa que la mecánica clásica ya hacía bastante bien; pero cuando lo que nos interesa medir es el nivel de ingresos de un individuo o el tiempo que pasa escolarizado, una diferencia estadísticamente significativa (es decir, no aleatoria) de *0,9 años de escolaridad* ***o de 98 € al mes*** entre uno u otro individuo es muy poco importante, *casi despreciable*, máxime si se tiene en cuenta que la primera selección de la muestra (*vid. supra*) ya parecía dejar fuera a las clases menos pudientes del municipio en cuestión[6].

d)¿Se te ocurre alguna otra variable oculta que el estudio no haya considerado, y que pudiera provocar, no solo la asociación (correlación o regresión) entre variables, sino que pudiera ser buen candidato a constituirse como causa de la variable respuesta (nivel de ingresos, logros académicos, cociente intelectual?

Se debería, en este punto, dejar una cierta "carta blanca" al alumno para que exprese sus opiniones y dudas, sin tratar de coartarle, tanto si lo que sugiere es relevante como si no. Por nuestra parte creemos que el factor de mayor efecto es la autoselección de los individuos que realizan el test cuando ya tienen 30 años, con ese porcentaje de falta de respuesta del 41,8%, por lo que debería especificarse los términos en que este análisis se realizó: si había alguna compensación económica, los resultados podrían estar sesgados hacia la parte de la muestra con menos ingresos; si, por el contrario, el sujeto sabía que iba a ser preguntado por sus ingresos, puede que el sesgo fuera contrario.
Podría, asimismo, suponerse que en 1982 la estructura familiar del Brasil era más patriarcal que en la actualidad, y que entonces la madre, al estar fuera del mercado de trabajo, pudiera emplear un mayor tiempo en la lactancia del bebé. De ese modo serían las familias con problemas económicos (que además suelen ser las que tienen mayor número de hijos y, por tanto, menor posibilidad de dedicación a cada hijo particular) las que requerirían más a menudo ingresos extraordinarios, con lo que la madre tendría una menor disponibilidad para largos períodos de lactancia. De esta manera, lo que mediría el estudio no sería la influencia del dicho período, sino de la extracción socioeconómica de los padres lo que realmente influiría en el posterior desempeño económico, académico e intelectual del sujeto.

6 Somos conscientes de que nuestro análisis del estudio en cuestión es, en alta medida, especulativo. Pero el propósito didáctico de la actividad no es tanto desarrollar un ejercicio de metanálisis serio y fundamentado, sino provocar que el alumno especule a su vez, es decir, que se pregunte sobre el cómo y el por qué de los elementos del estudio y que ejerza el pensamiento crítico sobre la información que recibe, tanto si esta es de carácter estadístico o no, tanto si es cuantitativa o cualitativa. Por otra parte, resulta sumamente improbable que el prestigio académico de Victora, Horta, *et al.* salga malparado de un debate producido en un aula de enseñanza secundaria.

Primero explora tus datos

En el capítulo precedente tratábamos de mostrar al alumno mediante ejemplos y actividades cómo una relación netamente estadística como la de asociación entre variables (por simple correlacion, o según otro modelo de regresión), *id est*, una relación de carácter matemático entre valores numéricos estructurados (datos) puede no ser suficiente para inferir una relación causal (cuyo dominio ya no es matemático, sino factual—al menos como hipótesis de trabajo) entre los fenómenos representados por aquellas variables. Se proponían como objeto de discusión, a su vez, las condiciones que hacen plausible la práctica de esta inferencia: el buen diseño de un experimento aleatorio comparativo—o analítico—provee de razones más sólidas de inferencia que un estudio observacional; además, se consideró la posibilidad de paliar (hasta cierto punto) la influencia de variables ocultas mediante diversos procedimientos, a los que es necesario recurrir dado que muchas veces la realización de un experimento *comme il faut* no está a disposición del investigador.

Al mismo tiempo, implícita y explícitamente se defendía una posición didáctica según la cual el análisis del contexto circunstancial de los datos, junto a la exposición por parte del profesor de los principios básicos de la investigación (y de la información) estadística (o científica) podrían integrarse en el currículo matemático de ESO pensado desde un punto de vista competencial, ya que conforman en los alumnos de enseñanza secundaria (que sólo calcularán coeficientes de correlación y rectas de regresión en 2º de Bachillerato, si escogen la opción de Ciencias Sociales) capacidades de interpretación y juicio de la información decisivas, además de una perspectiva más amplia acerca de las mismas matemáticas.

Desde esta posición, por tanto, continuamos con algunos tópicos en relación al análisis exploratorio de datos **[45]**, que sirvan de complemento (y, en cierta manera, fundamento) a la estadística matemática. En el anterior capítulo se analizaba la calidad y cualidad de las variables asociadas: en el presente veremos que la misma *morfología* de los valores de estas variables (los datos), no sólo contribuye a darles significado, sino que su análisis es una práctica que no se debe soslayar.

El Cuarteto de Anscombe

Así llamado desde la publicación del abundantemente citado artículo del autor homónimo[1], escrito, por cierto, en defensa del método gráfico de análisis de datos:

> *Los gráficos puede tener varios propósitos, como: i) ayudarnos a percibir y distinguir aspectos generales de los datos, ii) permitirnos mirar bajo esos aspectos generales y ver qué más hay La mayor parte de las clases de cálculo estadístico descanasan en suposiciones sobre el comportamiento de los datos. Estas suposiciones podrían ser falsas, y entonces los cálculos podrían desorientarnos.*

1 F. J. Anscombe, *Graphs in Statistical Analysis* [articulo en línea] Febrero 1973 [consultado Junio 2015] en *American Statician,* Vol. 27, nº1, pp. 17-21 Disponible en [previo registro]: http://www.jstor.org/stable/2682899?origin=JSTOR-pdf

> *Debemos siempre tratar de comprobar si las suposiciones son razonablemente correctas; y si son falsas, debemos ser capaces de percibir de qué manera son falsas. Los gráficos son muy valiosos para estos propósitos* [trad. propia].

Siguiendo en parte estas consignas propondremos la siguiente actividad, cuyos ejemplos están tomados de ese mismo artículo.

ACTIVIDADES

1º) Halla las medias y desviaciones típicas de cada variable, y la covarianza y coeficiente de correlación de cada par de variables; halla también la recta de regresión (de y sobre x), todos correspondientes a las cuatro siguientes agrupaciones de datos (usando solo la calculadora):

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Los valores son los siguientes, iguales para I, II, III y IV

·Media x: $\overline{x} = 9.0$

·Media y: $\overline{y} = 7.5$

·Desviación típica x: $S_X = 3.317$

·Desviación típica y: $S_Y = 2.032$

·Covarianza xy: $S_{XY} = 5.501$

·Coeficiente de correlación: $r = \dfrac{S_{XY}}{S_X \cdot S_Y} = 0{,}816$

·Ecuación de la recta de regresión: y= 3+0,5x

2°) A la vista de los resultados, ¿Dirías que los cuatro grupos de datos se comportan de manera muy similar? ¿Dirías que proceden del mismo estudio? ¿Están asociadas, según tu criterio, las dos variables en los cuatro casos?

Preguntas más o menos retóricas que pueden hacerse al vuelo: aunque efectivamente, como resúmenes estadísticos están incluidas las medias, las desviaciones típicas y la covarianza (necesarias para calcular el coeficiente de correlación) y la recta de regresión con un r bastante alto. Observadas sin mucho ahínco, podría decirse en términos generales que podrían proceder del mismo estudio y quie, efectivamente, x e y están positivamente asociados según una ley lineal simple.

3°) Representa cada agrupación de datos en un diagrama de puntos, junto con la recta de regresión en cada uno de ellos [aquí podría indicarse al alumno que utilizara una hoja de cálculo]

Fig. 23: Cuarteto de Anscombe I

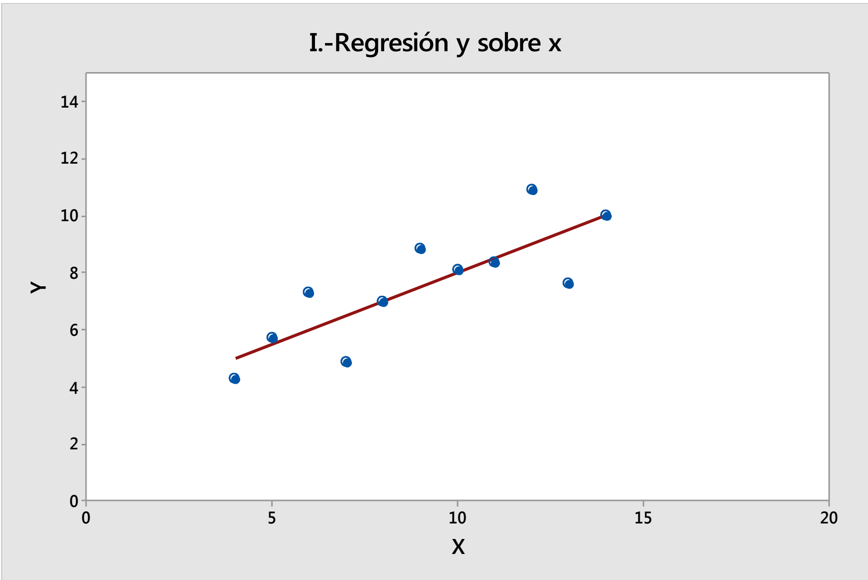


Fig. 23: Cuarteto de Anscombe II

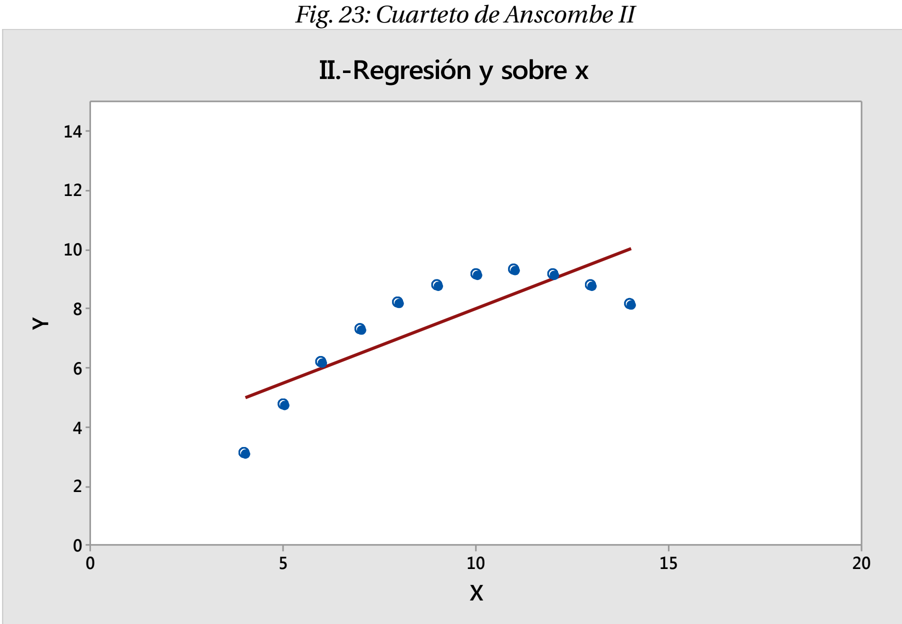
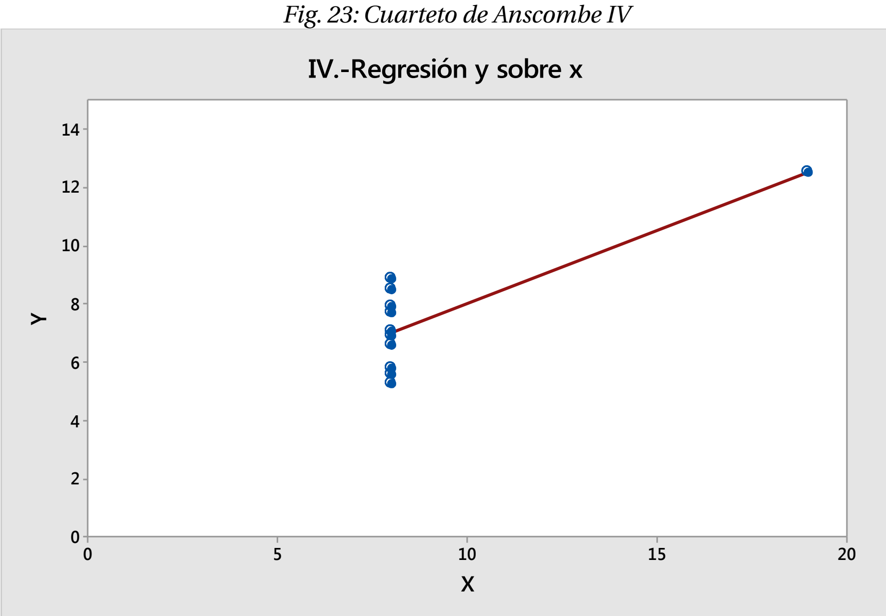


Fig. 23: Cuarteto de Anscombe III

Fig. 23: Cuarteto de Anscombe IV



a) Señala en lenguaje común cuáles son las diferencias que encuentras entre las nubes de puntos de los cuatro grupos ¿Pareen ahora resultados del mismo estudio?

Las del primero tiene la forma usual de una toma de datos entre dos variables relacionadas linealmente, con forma elipsoidal inclinada. La segunda ofrece el aspecto de una curva exactamente cuadrática. La tercera es, una vez más exactamente lineal excepto por uno de los valores, que provoca que la recta de regresión se halle desplazada respecto de la que tendría, muy fuertemente lineal, si no existiera ese valor. Finalmente, la cuarta presenta un aspecto muy extraño para cualquier proceso, solo observamos la y básicamente en un valor de x. Entre esta x y la otra que solo da lugar a una observación el espacio está vacío. La regresión está dominada por la última observación: Salga donde salga, pasará por ella.

b) ¿Hay alguna de estas gráficas en las que parezca adecuado hablar de asociación entre variables? ¿De entre ellas, cuáles responden a un modelo de regresión lineal simple? ¿Y cuál de ellas parece responder a la ecuación de regresión hallada en el ejercicio 1º? Trata de describir qué clase de procesos o experimentos han dado como resultados los valores de las cuatro gráficas.

Las tres primeras parecen mostrar, ciertamente, alguna clase de asociación entre variables. Sólo la primera y la tercera pueden responder a un modelo de regresión lineal simple; la segunda muestra una fuerte asociación, pero según una curva polinómica. Por otra parte, la tercera muestra un valor anómalo que merece ser investigado, ya que los demás responden muy estrechamente al modelo lineal (aunque con una ecuación de recta de regresión distinta a la que se obtiene teniendo en cuenta esta anomalía).

La descripción se solicita del alumno, una vez más, porque en un análisis exploratorio de datos se hace imprescindible especular (controladamente) sobre la razón de ser de la morfología del diagrama de puntos y de sus posibles ajustes. De tal manera los resultados anómalos u otros accidentes que no responden, o responden solo parcialmente a los modelos previstos (que siempre se tiene en mente cuando se realiza un estudio) pueden ser estimados y controlados en una siguiente toma de datos.

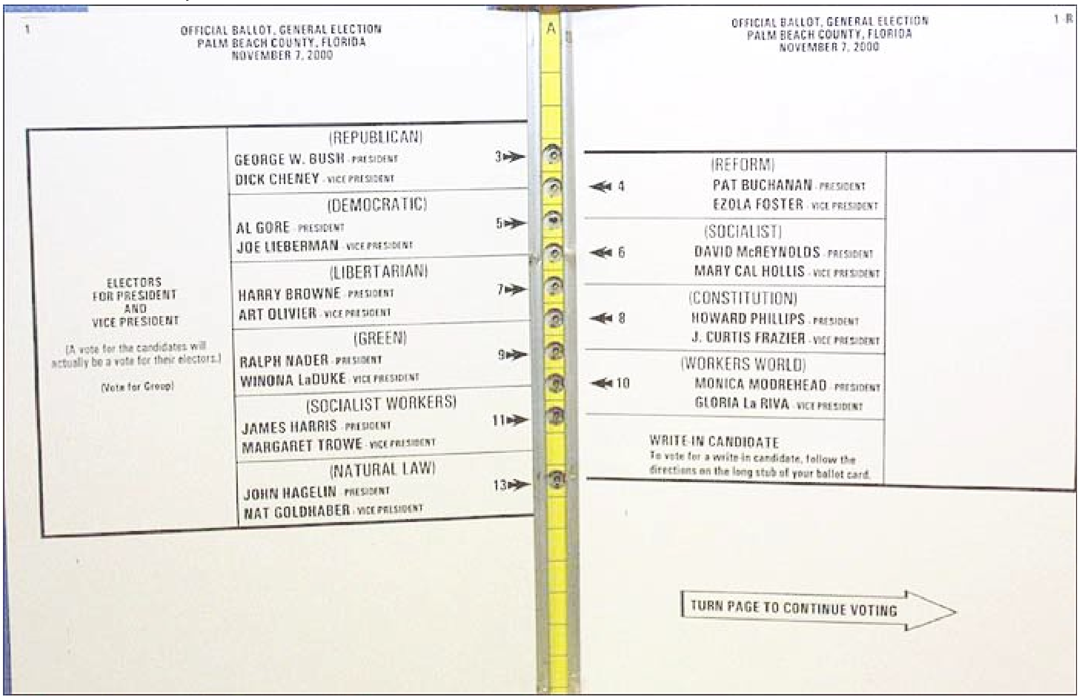
Así, la primera de las gráficas parece ser resultado de una toma de datos coherente, entre dos variables con cierta asociación. La segunda nos muestra claramente que no hay un modelo lineal subyacente, con lo cual considerar la correlación como un buen descriptor de la relación entre variables nos lleva a error. Es preciso ensayar alguna otra clase de modelo de regresión no lineal: las variables, por tanto, estarán relacionadas por alguna clase de ley, (muy determinista), pero no lineal — quizá cuadrática. La tercera, como ya hemos dicho, indica una relación muy fuerte, si hacemos obvia la observación atípica: sería preciso averiguar si procede de algun error de medición o transcripción de datos, o, en caso contrario, hipotetizar alguna razón de ser de la anomalía: quizá el modelo lineal (previsto o no) no sea exactamente erróneo, sino sencillamente incompleto. Por último, la cuarta gráfica responde claramente a una mala elección de variables de estudio: el dominio de la variable independiente es ridículamente estrecho y desequilibrado: es básicamente un problema de una sola variable, la y.

El Caso de las "Papeletas Mariposa" (Elecciones Gore-Bush 2000 en Palm Beach, FL)

En el caso anterior ya se ha indicado que las observaciones atípicas dentro de un grupo de datos no deberían (como tantas veces sucede) obviarse o descartarse como errores o meros accidentes. Un análisis de datos consecuente debe investigar precisamente estas anomalías y tratar de averiguar su origen. En el caso que vamos a exponer a continuación, de hecho, la importancia del análisis recae precisamente en el resultado atípico, y no en la norma de la distribución.

En las elecciones presidenciales estadounidenses del año 2000, en las que figuraban como candidatos principales AL Gore y George W. Bush, fueron unas de las más reñidas confrontaciones electorales de la historia del país. En este contexto, se produjo una profunda controversia —que llegó a los tribunales— sobre los sistemas de votación y el recuento de votos en ciertas circunscripciones clave. Uno de los condados bajo interdicto fue Palm Beach, Florida, donde se produjeron resultados electorales anómalos. Las sospechas se centraron en su bastante peculiar sistema de votación, concretamente en la papeleta de votación, o "papeleta mariposa" (*butterfly ballot*)[2]:

Papeleta electoral durante las elecciones del 2000, condado de Palm Beach



Fuente: G. Adams, *Voting Iregularities in Palm Beach, Florida*

Como se puede observar, la papeleta puede inducir a error: mientras que para el candidato republicano (George W. Bush) la indicación no tiene pérdida por estar ubicada en la posición superior, y su marca en primer lugar, para votar por el candidato demócrata (Al Gore) es preciso discriminar entre el lugar que ocupa en el listado de candidatos (el 2º) y la posición de la marca, indicada por la flecha (la 3ª). Entre el partido demócrata y el republicano se reparten el 99% de los votos, con lo cual la atención de la inmensa mayoría de los votantes se centra en el lado izquierdo de la papeleta, sin preocuparse de las opciones de la derecha. La comprobación de que, en efecto, este peculiar diseño indujo a error a un gran número de votantes se desarrollará en la siguiente actividad.

ACTIVIDADES

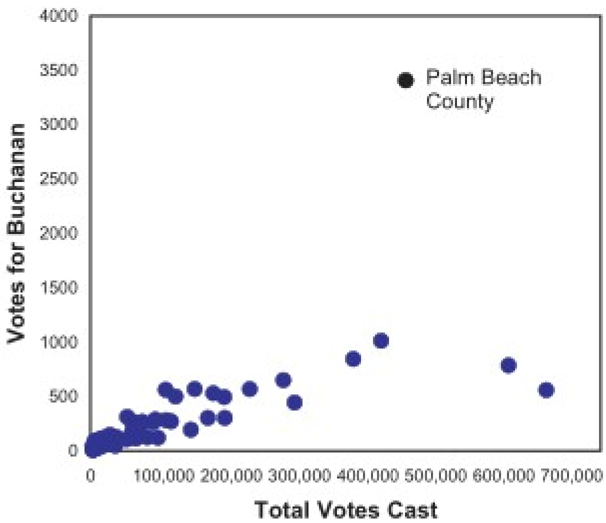
1º) Las elecciones presidenciales entre George W. Bush y Al Gore en el año 2000 fueron, además de muy reñidas, objeto de controversia pública, en parte debido al voto en uno de los condados que dieron la victoria a Bush, el de Palm Beach, Florida. Aquí (vid.

2 La imagen siguiente y las correspondientes en las actividades, G. Adams, *Voting Iregularities in Palm Beach, Florida* [artículo en línea], 2001 [consultado Junio 2015], en *CHANCE* (ASA), vol. 14 nº1 pp. 22-23. (pub. en línea en www.amstat.org) Disponible en: http://www.amstat.org/misc/VotingIrregularitiesArticle.pdf

supra) *se muestran las instrucciones para rellenarla. ¿Qué casilla corresponde al partido republicano y cuál al demócrata? Comenta brevemente si has encontrado alguna dificultad en averiguar la respuesta a la anterior pregunta debido a alguna característica del diseño de la papeleta.*

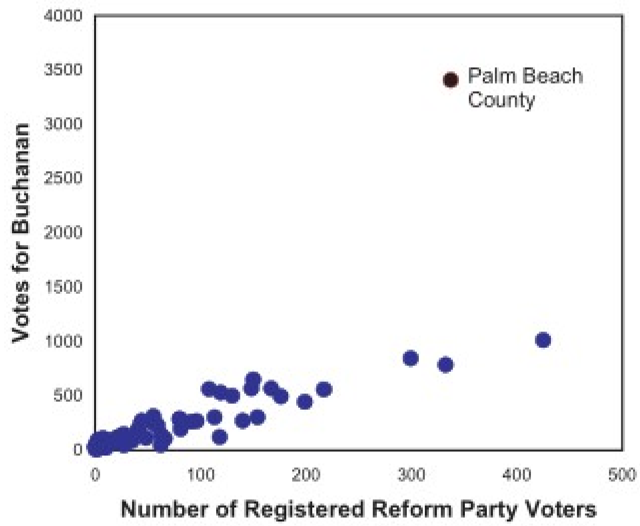
Al partido republicano le corresponde la casilla 3, y al demócrata la 5. No obstante, es preciso prestar atención a la papeleta para encontrar la correspondencia. En un primer momento puede suponerse que el orden de presentación de las candidaturas por escrito corresponde al orden descendente de las casillas, de manera que, aunque el voto republicano está claro, el voto demócrata podría realizarse por error a favor del candidato del Partido Reformista de Pat Buchanan, cuya casilla está en 2º lugar (en orden descendente) en lugar de la correspondiente, que está en 3er lugar.

2°)Observa la siguiente gráfica de distribución de votos por condado correspondiente al partido reformista de Pat Buchanan. Cada estado viene representado por un punto yel eje de abscisas corresponde a su número de votantes ¿Encuentras alguna anomalía? Imagina que no has hecho la primera actividad. ¿A qué crees que correspondería esa resuktado en Palm Beach? Si llegaras a la conclusión de que es un error o un efecto de la mera variabilidad intrínseca a cualquier recogida de datos, ¿Descartarías esa observación o investigarías sobre un origen alternativo a tus primeras suposiciones?



Decididamente, si se descarta el error en transcripción o recogida de datos, el resultado en Palm Beach, incluso a primera vista, no puede ser fruto de la mera variabilidad. Debe haber alguna razón subyacente a esa observación. Con lo cual, no sería prudente descartarla en un primer momento, sino realizar algunas pesquisas para averiguar su razón de ser. Alguna línea de investigación podría llevarnos a considerar el diseño de las papeletas electorales del condado en cuestión, y llevarnos a formular la hipótesis provisional de que, efectivamente, es la ubicación equíivoca de casillas en la misma la que ha dado lugar a este resultado anómalo.

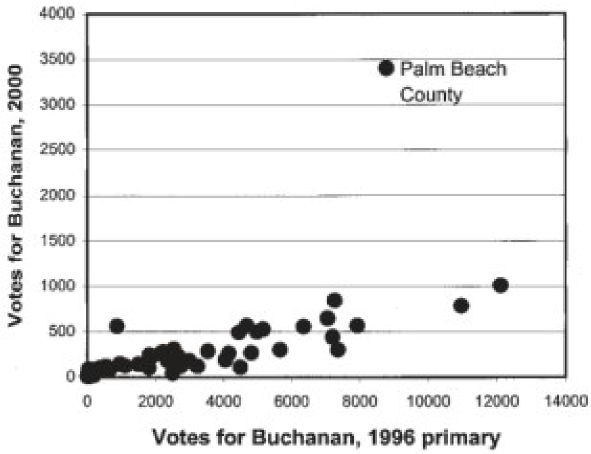
3º) Durante tus averiguaciones sobre el asunto confirmas que en Palm Beach hay un número desusado de votantes afiliados al partido reformista ¿Sería suficiente esta consideración para explicar la observación atípica? ¿Se te ocurre alguna manera de contrastar la nueva hipótesis resultante de esta información?



Como hipótesis alternativa, parece bastante razonable; una anomalía explicaría a la otra. Pero es muy diferente si se representan y exploran los datos: la gráfica muestra que, si bien puede haber una correlación positiva entre número de afiliados y número de votos al partido en cuestión en los demás condados, esta relación se dispara una vez más en el caso de Palm Beach; si esa observación fuera consitente con las demás, el partido reformista habría tenido alrededor de 1000 votos, y no 3500, como aparece en el diagrama de puntos. Por lo tanto, esta hipótesis sigue teniendo un apoyo más débil que nuestra hipótesis inicial.

4º)Otra averiguación que realizas es que Pat Buchanan se afilió él mismo al partido reformista en un breve plazo anterior a las elecciones, aproximadamente de un año: ¿Podría suceder que Pat Buchanan tuviera alguna clase de apoyo personal localizado en Palm Beach, y que fuera este apoyo el que hubiera alzado el numero de votantes del partido reformista, que hasta las elecciones del 2000 no contaba con él entre sus filas? ¿Se te ocurre alguna manera de contrastar esta hipótesis?

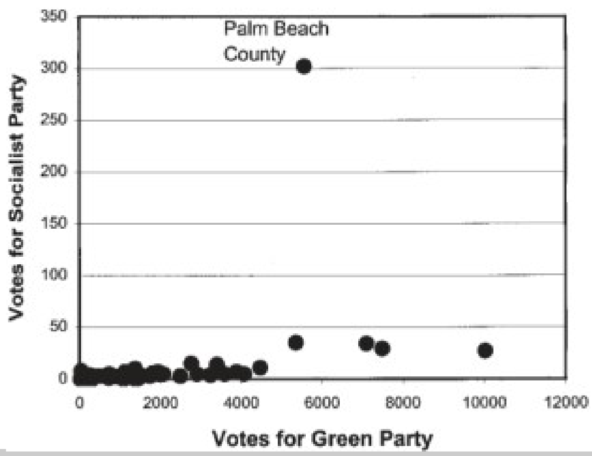
Efectivamente, podría ser plausible: quizá hiciera una campaña intensiva en ese condado, o fuera la cabeza de algún grupo comercial o empresarial localizado en Palm Beach, lo cual también le habría dado apoyo personal. Una manera de contrastar esto es comprobar alguna otra elección en la que Pat Buchanan hubiera sido candidato por otro partido que no fuera el reformista, y comprobar si este apoyo personal no tenía nada que ver con su afiliación. Pero como se puede comprobar en el diagrama de puntos siguiente, que refleja una comparación entre los votos recibidos en las elecciones primarias del partido republicano en 1996 (fecha en la que Buchanan pertenecía a este partido) y los recibidos en 2000 como miembro del reformista, se sigue observando otra vez una anomalía: los votos por Buchanan en 1996 (abscisas) reflejan otra vezque son ligerammente correlatados según el número de votantes del condado, pero asimismo no existe tal correlación entre el número de votos en 2000 con los de 1996; de otra forma, la abscisa correspondiente de Palm Beqach estaría mucho más ala derecha, y el punto representado en la gráfica podría encajar perfectamente en la misma recta de regresión lineal que los demás condados. Como este no es el caso, la observación atípica queda otra vez sin explicar según esta última hipótesis.



Con lo cual, nuestra hipótesis incial sigue siendo la más plausible.

5º) Pongamos que se te han acabado los recursos de documentación o de imaginación y no se te ocurre ninguna otra hipótesis alternativa a la inicial ¿Se te ocurre, no obstante, alguna línea de investigación que diera un apoyo más sólido a la primera hipótesis?

Como no queremos ser desleales con el alumno, pasaremos rápidamente a exponer (siguiendo el artículo que hemos referenciado en la nota [2]) que, si nuestra hipótesis incial incluía un posible error en la elección del orden de las casillas entre Gore y Buchanan sería razonable suponer un error similar entre otras casillas igualmente susceptibles de ser confundidas. por ejemplo, si en las elecciones del 2000 el partido Socialista (siguiente en el orden de especificación de partidos), y al que le corresponde la <u>4º casilla</u>, hubiera recibido una cantidad inusual de votos respecto al partido que se halla a su izquierda en <u>4º lugar según el texto</u>, tendríamos un fenómeno equivalente al que hemos hipotetizado inicialmente. Entonces deberíamos representar mediante un diagrama de puntos



la relación de votos entre ambos partidos y ver si el resultado en Palm Beach es también atípico, con la condición adicional de serlo favorablemente al partido socialista, con lo cual la equivalencia entre los errores quedaría confirmada. Pues bien, esto es exactamente lo que sucede. Mientras que la relación entre votantes del partido socialista y el verde es más o menos proporcional en el resto de los condados, en el caso de Palm Beach los socialistas reciben una cantidad atípica de votos con respecto a los verdes. Con lo cual queda, no confirmada, pero sí mucho más asegurada sobre bases más firmes nuestra primera hipótesis, la cual, de todas formas, no está completamente demostrada. Los datos sugieren que estamos en lo cierto con nuestro primer aserto, pero, como sabemos de los anteriores casos, puede haber variables desconocidas que expliquen todos estos fenómenos.

En todo caso, lo que indican los datos es que el asunto merece una investigación profunda, por medios estadísticos u otros más directos. Y otra cosa: que, sin los procedimeintos de naturaleza estadística, la aplicación sistemática estado a estado y condado a condado de cada tipo de papeleta sería poco menos que inviable. Cuando la naturaleza y la cantidad de los datos es en principio inmanejable, entonces es cuando la estadísitca se convierte en una herramienta imprescindible en la investigación, lo cual incluye no solo la política, sino muchas otras ramas del concoimiento: medicina, farmacología, biología, sociología, psicología, economía, geología, antropología, etc.

Si bien hay que decir que en este caso la atipicidad de los datos es tan alta, que corremos el riesgo de comportarnos como cierto profesor de matemáticas de secundaria que, al contemplar un diagrama de dispersión de talla-peso de unos alumnos de instituto y encontrar que uno de ellos pesa 70kg. y mide 6 metros de alto, se sorprenda de no haber visto nunca a tal individuo rondando por los pasillos…

BIBLIOGRAFÍA

[1] *Ley 14/1970, de 4 de agosto, General de Educación y Financiamiento de la Reforma Educativa* [LGE] Jefatura del Estado. Gobierno de España. Disponible en: http://www.boe.es/diario_boe/txt.php?id=BOE-A-1970-852 *13*

[2] *Ley Orgánica 1/1990, de 3 de octubre de 1990, de Ordenación General del Sistema Educativo* [LOGSE] Jefatura del Estado. Gobierno de España. Disponible en: http://www.boe.es/buscar/doc.php?id=BOE-A-1990-24172 *13*

[3] *Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa* [LOMCE], Jefatura del Estado. Gobierno de España. Disponible en: https://www.boe.es/diario_boe/txt.php?id=BOE-A-2013-12886 *12, 30*

[4] *Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato,* Ministerio de Educación, Cultura y Deporte. Gobierno de España. Disponible en: http://www.boe.es/boe/dias/2015/01/03/pdfs/BOE-A-2015-37.pdf *37*

[5] P.K. Andersen y M. Væth, *Survival Analysis,* en *Encyclopedia of Statiscal Sciences*, S. Kotz y N.L. Johnson, (eds. lit.)2ª ed.,Vol. 12, Wiley & Sons, Hoboken, NJ, 2006, pp. 8474-8483.*54*

[6] C. Batanero, *La comprensión de la probabilidad en los niños, ¿Qué podemos aprender de la investigación?* [artículo en línea], en *Atas do III encontro de probabilidades e estatística na ecola. Braga : Centro de Investigação em Educação da Universidade do Minho,* J. A. Fernandes , P. F. Correia , M. H. Martinho, & F. Viseu , (eds.) (2013) [consultado Junio 2015]. Disponible en:*www.ugr.es/~batanero/pages/ARTICULOS/1Batanero.pdf* *3*

[7] K. Bury, *Statistical Models in Applied Science,* Wiley & Sons, New York [etc.], 1975. *50*

[8] G.C. Canavos, *Probabilidad y Estadística: Aplicaciones y Métodos,* McGraw-Hill, Madrid, 1993 Traduccion de Edmundo Gerardo Urbina Medal. *50, 58*

[9] K.L. Chung, *Teoría Elemental de la Probabilidad y de los Procesos Estocásticos,* (3ª ed.), Reverté, Barcelona, 1983. Traducida por L. B. Bou García, rev. de E. Linés Escardó. *23, 103*

[10] G.W. Cobb y D.S. Moore, *Mathematics, Statistics, and Teaching,* Nov. 1997, [consultado Mayo 2015] Am. Math. Month. Vol. 194, 9, pp. 801-823. Disponible en [previo registro]:http://links.jstor.org/sici?sici=0002-9890%28199711%29104%3A9%3C801%3AMSAT%3E2.0.CO%3B2-7 *3, 28, 30, 151*

[11] D.R. Cox y E.J. Snell, *Applied Statistics: Principles and Examples,* Chapman & Hall, London, 1992. *49*

[12] H. Cramer, *Métodos Matemáticos de Estadística,* 2ª ed., Aguilar, Madrid, 1953. Traducción de Enrique Casado. *7, 29, 49*

[13] C. Dagum, *Income Distribution Models,* en *Encyclopedia of Statiscal Sciences,* S. Kotz y N.L. Johnson, (eds. lit.)2ª ed., Vol.5, Wiley & Sons, Hoboken, NJ, 2006, pp. 3378-3386. *50, 54*

[14] F.N. David, *Games, Gods, and Gambling: The origins and history of probability and statistical ideas from the earliest times to the Newtonian era,* Hafner Pub. Co., New York, 1962. *1,2*

[15] P. Diaconis, *Theories of Data Analysis: From Magical Thinking to Classical Statistics,* en *Exploring Data Tables, Trends and Shapes,* D.C. Hoaglin, F. Mosteller y J.W. Tukey (eds. lit.) Wiley &Sons, Hoboken, NJ, 1985 *27*

[16] W. Eckhardt, *Paradoxes in Probability Theory* [extracto de libro en línea], Springer, Dordrecth [etc.], 2013. Disponible en [consultado Junio 2015]: https://books.google.es/books?id=hbm8prMLBmYC&pg=PA2&lpg=PA2&dq#v=onepage&q&f=false *3, 125*

[17] W. Feller, *Introducción a la teoría de probabilidades y sus aplicaciones, vol. 1,* 3ª ed.(1968), Limusa-Wiley, México, 1973. *5, 125*

[18] R.A. Fisher, *The Design of Experiments,* 9ª ed., Hafner Press, New York, 1971, *15*

[19] D. Freedman, R. Pisani y R. Purves, *Statistics,* 4ª ed., Norton&Co, New York, 2007. *149, 150*

[20] G. Gigerenzer, Z. Swijtink, *et. al., The Empire of Chance: How Probability Changed Science and Everyday Life,* Cambridge University Press, Cambridge, 1989. [Extracto] Disponible en:*https://books.google.es/books?id=Bw2yKfpvts8C&lpg=PP1&pg=PA12#v=onepage&q&f=false* [consultado Mayo 2015]. *6, 8, 9, 10, 12, 29, 83, 83, 103*

[21] D. Gillies, *Philosophical Theories of Probability* [extracto de libro en línea], Routledge, London, 2000. Disponible en [consultado Mayo 2015]: https://books.google.es/books?id=iPR2fTfgylYC&lpg=PP1&hl=es&pg=PP1#v=onepage&q&f=false *6*

[22] P. Gorroochurn, *Classic Problems of Probability* [recurso en línea] (pub. en línea 7 Mayo 2012 [consultado Mayo 2015] en Wiley Online Library) John Wiley & Sons, Inc., Hoboken NJ, 2012. DOI: 10.1002/9781118314340.,*3, 104, 106, 107, 125, 145*

[23] I. Hacking, *The Taming of Chance,* Cambridge University Press, Cambridge, 1990. *3, 6, 9, 103*

[24] R. Hastie y R. M. Dawes, *Rational Choice in an Uncertain World: The Psychology of Judgement and Decision Making* [extracto de libro en línea], 2ªedición, SAGE, Thousand Oaks (California), 2010. Disponible en [consultado Junio 2015]: https://books.google.es/books?id=w_fYZZPcgpkC&printsec=frontcover&hl=es&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false *128*

[25] M. Hazewinkel (ed. lit.), *Encyclopaedia of Mathematics Vol. 2: Coproduct — Hausdorff — Young Inequalities* [libro electrónico], Springer Science & Business Media, 1995. DOI: 10.1007/978-1-4899-3795-7. *105*

[26] D.C. Hoaglin, F. Mosteller y J.W. Tukey, *Understanding Robust and Exploratory Data Analysis,* Wiley & Sons, New York, 2000. *49*

[27] D. Huff y I. Geis, *How to take a chance,* Norton & Co., New York, 1959. *91, 115*

[28] D. Huff y I. Geis, *How to lie with Statistics,* Norton & Co., New York, 1954. *26, 84, 91*

[29] D. Kahneman, P. Slovic y A. Tversky (eds.), *Judgement under uncertainty: Heuristics and biases,* Cambridge University Press, Cambridge, 1982. *3, 19, 104, 123,137*

[30] M.G. Kendall, *Studies in the History of Probability and Statistics II. The Beginnings of a Probability Calculus,* en *Biometrika,* [recurso en línea. DOI:10.2307/2333573] Junio 1956, [consultado Mayo 2015]Vol. 43, Nº 1/2, pp. 1-14. Disponible en : http://www.jstor.org/stable/2333573 [previo registro].*2, 104*

[31] M. Kline, *El pensamiento matemático de la Antigüedad a nuestros días*, vol.I, Alianza, Madrid, 1992. Traducción de M. Martínez. *104*

[32] M. Kline, *El pensamiento matemático de la Antigüedad a nuestros días*, vol.II, Alianza, Madrid, 1992. Traducción de M. Martínez.,*2, 105*

[33] A.N. Kolmogorov, *La Teoría de la Probabilidad,* en *La Matemática: su contenido, métodos y significado,* A.D. Aleksandrov, A.N. Kolmogorov, *et al.* 7ª ed., vol 2, Alianza, Madrid, 1985. Trad. de Eduardo Abad Rius, cap. XI, *2, 5*

[34] W. Mendenhall, R.J Beaver y B.M Beaver, *Introducción a la Probabilidad y Estadística*, 13ª ed.,Cengage Learning, México, 2010. Trad. de J.H. Romo Muñoz, rev. por A.E. García Hernández. *58*

[35] D.M. Merkle, *Call-in Polls,* en *Encyclopedia of Survey Research Methods,* P.J. Lavrakas ed., SAGE, Los Angeles, 2008, pp. 77-78. Disponible en [consultado Marzo 2015]: *https://books.google.es/books?id=Rhp1AwAAQBAJ&lpg=PP1&dq=inauthor%3A%22Paul%20J.%20Lavrakas%22&hl=es&pg=PP1#v=onepage&q&f=false* *37, 40*

[36] D.S. Moore, *New Pedagogy and New Content: The Case of Statistics* [artículo en línea], Feb. 1997 [consultado Mayo 2015], en *International Statistical Review,* 65, 2, pp. 123-165. Disponible en: www.*iase-web.org/documents/int***stat***review/97.***Moore***.pdf* *3, 28*

[37] D.S. Moore y W.I. Notz, *Statistics. Concepts and Controversies,* W.H. Freeman and Co., New York, 2009. *37, 39, 42, 145, 150, 151*

[38] Organización para la Cooperación y el Desarrollo (OCDE), *Marcos y Pruebas de Evaluación de PISA 2012: Matemáticas, Lectura y Ciencias,* Instituto Nacional de Evaluación Educativa, Dirección General de Evaluación y Cooperación territorial. Secretaría de Estado de Educación, Formación Profesional y Universidades. Ministerio de Educación, Cultura y Deporte. Madrid, 2003. Disponible en:
http://www.mecd.gob.es/dctm/inee/internacional/pisa2012/ marcopisa2012.pdf?documentId=0901e72b8177328d *25, 30*

[39] D. Peña, *Estadística.Modelos y Métodos:1.Fundamentos* (6ª ed.), Alianza, Madrid, 1993. *49*

[40] C. Pérez, *Técnicas de Muestreo Estadísitico,* Ibergaceta, Madrid, 2010. *41*

[41] J.J. Sánchez Carrión, *Manual de Análisis de Datos,* Alianza, Madrid, 1995. *41*

[42] S. Siegel, *Estadística No Paramétrica Aplicada a las Ciencias de la Conducta* (2ª ed.), Trillas, México, 1985. *49*

[43] I. Todhunter, *A History of the Mathematical Theory of Probability from the time of Pascal to that of Laplace* [rep. de libro imp. en línea] (pub. en línea 4 Feb. 2008) [consultado Mayo 2015], McMillan & Co., Cambridge, 1865. Disponible en:
https://archive.org/details/ofmathemahistory00todhrich *105*

[44] E.R. Tufte, *The Visual Display of Quantitative Information,* 2ª ed., Graphic Press, Cheshire CT, 2007. *84, 96*

[45] J.W. Tukey, *Exploratory Data Analysis,* Addison-Wesley, Reading, Mass., 1977. *3, 27, 49, 84, 159*