

*firma invitada*


  
 Newton Leibniz Riemann Euler Bernoulli Fermat



**MONOGRÁFICO ESTADÍSTICA**

## Estadística: Enseñar y crear actitudes positivas a través de casos prácticos

Pere Grima Cintas

### Resumen

Una forma de despertar interés por la estadística es presentando casos prácticos que atraigan la atención de los estudiantes y a partir de ellos sacar las ideas y los conceptos que se quieran transmitir. En este artículo se presentan diez ejemplos, casos o situaciones prácticas, que se pueden utilizar en la enseñanza secundaria y que además de transmitir conocimientos pretenden despertar una actitud positiva, de interés y curiosidad, por todo lo que la estadística puede aportar.

### Abstract

One way to create interest in statistics is presenting case studies to attract student's attention and draw from them the ideas and concepts that we want to convey. This article presents ten examples, cases and practical situations, that can be used in high school and that as well as transmitting knowledge can awaken a positive attitude, interest and curiosity, for all that statistics can provide.

### Resumo

Uma forma de criar interesse nas estatísticas é apresentando casos práticos para atrair a atenção dos alunos e a partir deles tirar deles as idéias e os conceitos que se queira transmitir. Este artigo apresenta dez exemplos, casos e situações práticas, que podem ser utilizados no ensino secundário e também pretendem despertar uma atitude positiva, interesse e curiosidade, para todos os que as estatísticas podem fornecer.

### 1. Introducción: Estadística a través de casos prácticos

La estadística es una de las asignaturas que aparece con más frecuencia en los planes de estudio de titulaciones universitarias. Incluso carreras que no suelen incluir las matemáticas en su currículum, como medicina, sociología, ciencias políticas o psicología, sí incluyen la estadística y tienen en esta materia una de sus herramientas de trabajo más importantes. Y no podía ser de otra forma, la estadística trata sobre cómo recoger datos (cuántos, de qué forma) y cómo analizarlos para obtener la información que permita tomar las decisiones más adecuadas.

Pero a pesar de su destacada presencia en las titulaciones universitarias y en muchas actividades profesionales, la estadística suele tener poco protagonismo en la enseñanza secundaria (pongamos, para entendernos, que entre los 14 y los 18 años). Rara vez existe una asignatura específica de estadística, y más bien se presenta incluida en el libro de matemáticas, mucha veces al final, de forma que si no da tiempo a verlo todo esta es la parte que se queda sin impartir.

Otro problema, según mi parecer, es que muchas veces se enseña persiguiendo solo objetivos de conocimientos: se pretende que los estudiantes aprendan algunas cosas de estadística descriptiva (cómo se construyen determinados gráficos, por ejemplo), cálculo de probabilidades (donde lo que se considera más importante –y difícil– es aprenderse las fórmulas de la combinatoria) y quizá también algo sobre el coeficiente de correlación o sobre el cálculo de los coeficientes de una ecuación de regresión. Otra cosa es que esos conocimientos permanezcan y estén disponibles para sacarles provecho cuando se presente la ocasión o que, por el contrario, se derrumben una vez pasado el examen y no solo no quede nada, sino que dejen un poco de rechazo a la estadística, una sensación de que “esto no es lo mío”.

Creo que cuando se estudia estadística en la enseñanza secundaria los objetivos no deben estar solo orientados a los conocimientos sino también a crear actitudes positivas hacia esta materia. Me parece muy importante que cuaje la idea de que la estadística es una herramienta para conocer mejor la realidad (natural, física, social,...) que nos rodea y que esto se realiza a través del análisis de datos que reflejen de forma objetiva aquello que se desea conocer<sup>1</sup>. Una forma de crear esa actitud es desarrollando las clases en torno a casos prácticos que atraigan la atención de los estudiantes. Estos casos deben poner de manifiesto las posibilidades de la estadística para resolver problemas de interés y también se deben poder sacar lecciones (conocimientos) sobre la metodología usada y sobre sus posibilidades de aplicación en otras situaciones similares.

Por supuesto que este enfoque no es nuevo. Existen libros de introducción a la estadística (a nivel universitario) que plantean situaciones prácticas para poner en contexto la teoría que a continuación explican (como el de David Moore o el de Gary Smith). También existen libros excelentes solo con casos prácticos, como el de J.M. Tanur *et al.* o el de R. Peck *et al.*, que se comentan más adelante y que también se pueden utilizar en la enseñanza secundaria.

A continuación se presentan 10 situaciones prácticas, distribuidas en cinco apartados según la temática o el tipo de mensaje que se desee transmitir y que creo que cumplen ese doble objetivo de aportar conocimientos y generar una actitud positiva hacia la estadística.

## 2. Casos para destacar la importancia de tomar decisiones en base a datos convenientemente recogidos y analizados

Usar la estadística no necesariamente es sinónimo de utilizar palabras raras o de hacer cálculos complicados. Significa que deseamos ver la realidad de forma

---

<sup>1</sup> Y, por supuesto, hay que saber identificar cuáles son esos datos y cómo recogerlos, y también hay que saber analizarlos aplicando las técnicas más adecuadas, que no necesariamente deben ser complicadas.

objetiva, a través de datos que reflejen de la mejor manera posible qué es lo que está ocurriendo. Una vez se tienen los datos hay que saber sacarles la información y saberla plasmar de forma clara y convincente.

## 2.1. John Snow y el fin del cólera en Londres

En 1854 se desató una terrible epidemia de cólera en el centro de Londres que ocasionó más de 500 muertos en apenas 10 días. En aquella época imperaba la teoría de la miasma como vía de transmisión de la enfermedad (se suponía que “el veneno” estaba en los vapores que se desprendían de los cuerpos en descomposición) pero al Dr. John Snow esta teoría no le cuadraba con sus observaciones y sospechó que el problema podía estar en el agua.

Para estudiar la situación marcó sobre un plano de la zona el lugar donde vivían los fallecidos y observó que las marcas se distribuían en torno a una fuente (la de *Broad Street*). Los datos y la forma de presentarlos era tan clara y convincente que, a pesar de lo estafalaria que parecía su teoría, las autoridades ordenaron inutilizar la fuente y a partir de aquel momento los contagios cayeron en picada. Aunque no se entendía muy bien cómo el agua podía contagiar la enfermedad, las ciudades de los países desarrollados prestaron mucha atención a la higiene en la distribución de las aguas y el cólera dejó de ser una amenaza en los países desarrollados<sup>2</sup>.



Figura 1: Localización de los muertos por cólera realizada por John Snow sobre un plano de la época (vista parcial). Fuente: Wikipedia: “1854 Broad Street cholera outbreak”

<sup>2</sup> No puedo dejar de comentar que mientras escribo estas líneas están apareciendo en los periódicos noticias e imágenes dramáticas sobre la epidemia de cólera en Haití. Es evidente que la lucha contra las enfermedades no sólo es un problema de conocimiento científico.

**¿Qué se aprende?** La importancia de basar las decisiones en base a datos recogidos de forma inteligente y rigurosa. Un buen diseño de la recolección de datos facilita su análisis posterior.

**Otras actividades:** Discutir las posibilidades de uso de este tipo de representación gráfica en otros ámbitos: Identificación de la locación más frecuente de los defectos en un producto (por ejemplo, defectos de soldadura en un circuito impreso). Identificación de puntos negros en las carreteras...

**Más información en:**

Wikipedia: “John Snow” (más amplio en inglés) y “1854 Broad Street cholera outbreak”.

Página web dedicada al Dr. Snow elaborada por el Departamento de Epidemiología de la UCLA: <http://www.ph.ucla.edu/epi/snow.html>

Tufte, E.R. (1997) Visual Explanations. Graphics Press, pp: 27-37.

## 2.2. Florence Nightingale: Datos para tomar decisiones que salvan vidas

Se dice que la Guerra de Crimea (1853-1856) fue la guerra de la era moderna: se usó por primera vez el ferrocarril, era la primera vez que se disponía de telégrafo y es la primera guerra de la que se tienen fotografías. Pero también está considerada como la peor dirigida de todas aquellas en las que ha participado Inglaterra, los soldados morían a causa de enfermedades infecciosas, falta de comida, de agua y de condiciones higiénicas. El ambiente en los hospitales de campaña era caótico y la mortalidad muy elevada. Pero el uso del telégrafo no solo estaba a disposición de los militares. También lo usaban los reporteros de guerra para enviar crónicas a sus diarios explicando la incompetencia de los mandos militares y la falta de cuidados de los soldados heridos, generando un estado de opinión que obligó al ministro de la guerra a enviar un cuerpo de enfermeras para mejorar la atención a los heridos.

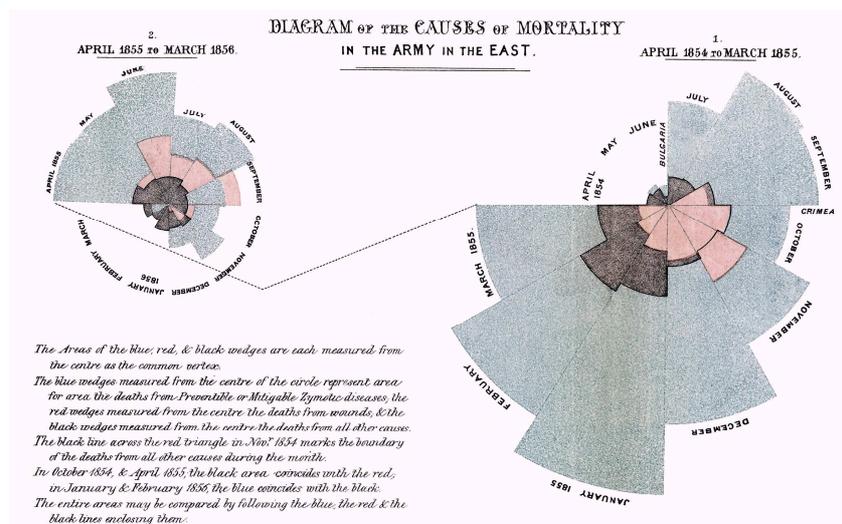


Figura 2: Diagrama sobre las causas de mortalidad y su evolución realizado por F. Nightingale (Fuente: Wikipedia: “Florence Nightingale”, en inglés). Una versión animada puede verse en la [página web](#) (al final) de la revista Science News

Al frente estaba una mujer entregada e inteligente llamada Florence Nightingale. Cuando llegaron el caos era total, pero Florence Nightingale supo observar, entender y documentar qué era lo que estaba ocurriendo y la asociación entre la masificación de los enfermos y la tasa de mortalidad. Sus conclusiones eran tan claras que convenció a la burocracia militar de la necesidad de realizar un profundo cambio en las formas de actuación: las prioridades fueron el orden y la limpieza, y la tasa de mortalidad bajó de una forma espectacular. Florence Nightingale es reconocida como experta en el uso de la estadística, utilizó nuevas formas de representación gráfica de datos y fue la primera mujer admitida en la Royal Statistical Society.

**¿Qué se aprende?:** La importancia de los datos para conocer lo que está pasando y para convencer a los que tienen que aportar los recursos (que muchas veces son reticentes a los cambios) de cuáles son las medidas que se deben tomar.

**Otras actividades:** Profundizar en la vida y las aportaciones de Florence Nightingale. Discutir cómo la estadística pone de manifiesto, incluso en nuestros días, que sin grandes inversiones se pueden salvar muchas vidas.

**Más información en:**

Wikipedia: "Florence Nightingale".

Florence Nightingale: "Measuring Hospital Care Outcomes", Joint Commission, 1999. Muy interesante la introducción escrita por D. Neuhauser: "Florence Nightingale: A Passionate Statistician", también en [Google Books](#).

**3. Ejemplos de la utilidad del análisis exploratorio de datos**

Una vez se tienen los datos, realizando gráficos sencillos de entender y de interpretar puede obtenerse una información que ayuda a tomar mejores decisiones.

**3.1. El caso de la panadería**

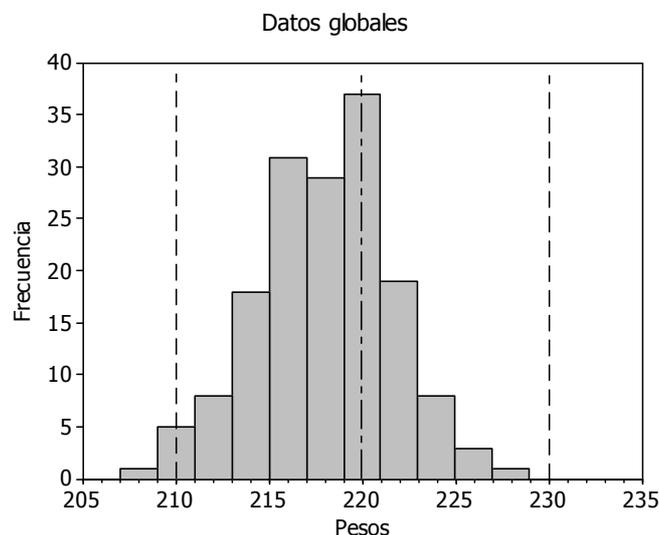
El dueño de una panadería está preocupado porque sospecha que las barras de pan que produce tienen un peso demasiado variable, de forma que algunas pueden estar incluso por debajo de los límites que marca la ley.

Día	Operario	Máquina 1				Máquina 2			
1	A	220.3	215.5	219.1	219.2	220.3	208.0	214.4	219.2
2	B	215.8	222.0	218.9	213.6	216.9	213.4	217.7	217.7
3	B	220.4	218.7	218.6	219.6	222.9	219.7	209.4	221.6
4	B	221.5	227.0	219.5	222.5	223.1	215.3	220.4	215.6
5	A	215.7	225.3	223.0	218.0	216.0	210.9	221.4	210.9
6	A	222.7	215.1	219.6	217.3	212.1	213.0	218.0	216.5
7	A	216.0	218.8	217.9	213.0	216.9	216.0	213.5	219.2
8	B	219.4	218.3	216.7	224.1	216.2	218.4	216.6	214.9
9	B	219.8	222.6	219.1	217.7	216.2	212.2	216.9	214.9
10	A	220.2	219.5	222.4	219.9	222.9	214.3	219.1	216.7
11	B	218.0	223.9	219.6	221.9	214.9	212.6	219.4	213.3
12	B	219.3	219.6	218.8	219.9	219.0	216.7	216.4	213.5
13	B	220.0	214.1	224.3	217.4	218.0	219.5	219.5	222.3
14	A	223.9	220.6	219.5	219.6	211.8	218.2	218.3	217.4
15	A	218.1	218.8	218.4	217.9	214.6	215.7	218.0	216.4
16	B	216.9	221.6	220.6	222.6	215.6	220.4	217.3	216.2
17	B	217.9	225.7	222.2	216.1	212.5	214.6	209.7	211.3
18	A	224.2	216.2	219.9	220.4	215.8	219.9	216.5	211.9
19	A	214.1	219.7	222.4	224.5	213.7	209.7	216.9	213.1
20	A	221.1	225.0	222.7	222.2	212.5	217.5	217.4	215.7

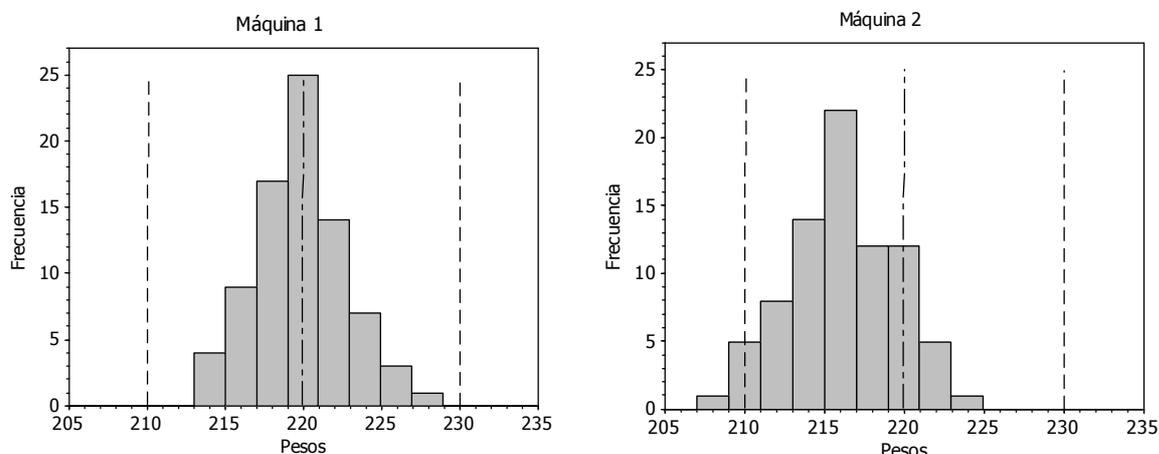
Se utilizan dos máquinas para hacer el pan y también trabajan dos operarios, unos días hace el pan uno y otros días el otro. La siguiente tabla contiene los pesos (en gramos) de una muestra de barras de pan que se han ido recogiendo durante 20 días.

El peso debe ser de  $220 \pm 10$  g y supondremos que estos datos son representativos de la producción general. Las preguntas que nos planteamos son: ¿Existe algún problema? ¿Qué está pasando? ¿Qué hay que hacer para resolver el problema, si es que existe?

Si se intenta sacar conclusiones simplemente mirando los datos es fácil equivocarse. Aunque en este caso solo hay 160 valores, intentar sacar conclusiones “a ojo” es arriesgado. Tampoco es necesario empezar haciendo grandes cálculos o aplicando técnicas sofisticadas, basta con representar los datos gráficamente (mejor si se realizan con algún programa, o con una hoja de cálculo tipo Excel):



Estratificando por operario y máquina se observa que el problema está en la máquina 2, que está descentrada. No hay problema en la máquina 1, y los dos operarios se comportan prácticamente igual.



**¿Qué se aprende?:** Los datos “en bruto” ocultan información que se puede hacer visible utilizando gráficos sencillos. Esa información permite realizar el diagnóstico correcto y tomar la decisión más adecuada.

**Otras actividades:** Plantarse preguntas, buscar datos para responderlas y representarlos gráficamente para ver la respuesta de forma clara (ver, por ejemplo, el siguiente caso). Los centros de estadística oficial (en España el Instituto Nacional de Estadística: <http://www.ine.es/inebmenu/indice.htm>) son una fuente abundante de datos.

**Más información en:**

Capítulos de estadística descriptiva de buenos textos, como el de David Moore “Introduction to the Practice of Statistics”.

Sobre malos usos en las representaciones gráficas (una forma divertida de tratar el tema) un clásico es el libro de Darrell Huff: “How to Lie with Statistics”.

**3.2. ¿Qué coches resultan baratos y cuáles caros respecto a su potencia?**

A muchos jóvenes les interesa el mundo de la tecnología (coches, motocicletas, ordenadores,...) y es fácil obtener datos sobre estos temas a través de internet.

Por ejemplo, en la página web del Real Automóvil Club de España se pueden obtener listados de coches que cumplan unas determinadas condiciones ([http://www.race.es/servicios/ofertas\\_de coches/buscador\\_de coches\\_nuevos/](http://www.race.es/servicios/ofertas_de coches/buscador_de coches_nuevos/)).

Buscando coches nuevos tipo “turismo-berlina” con motor diesel y 4 puertas se obtienen los datos de 449 coches (obtenido el 10 de noviembre de 2009) cuyo análisis con una hoja de cálculo permite responder a la pregunta planteada en este apartado.

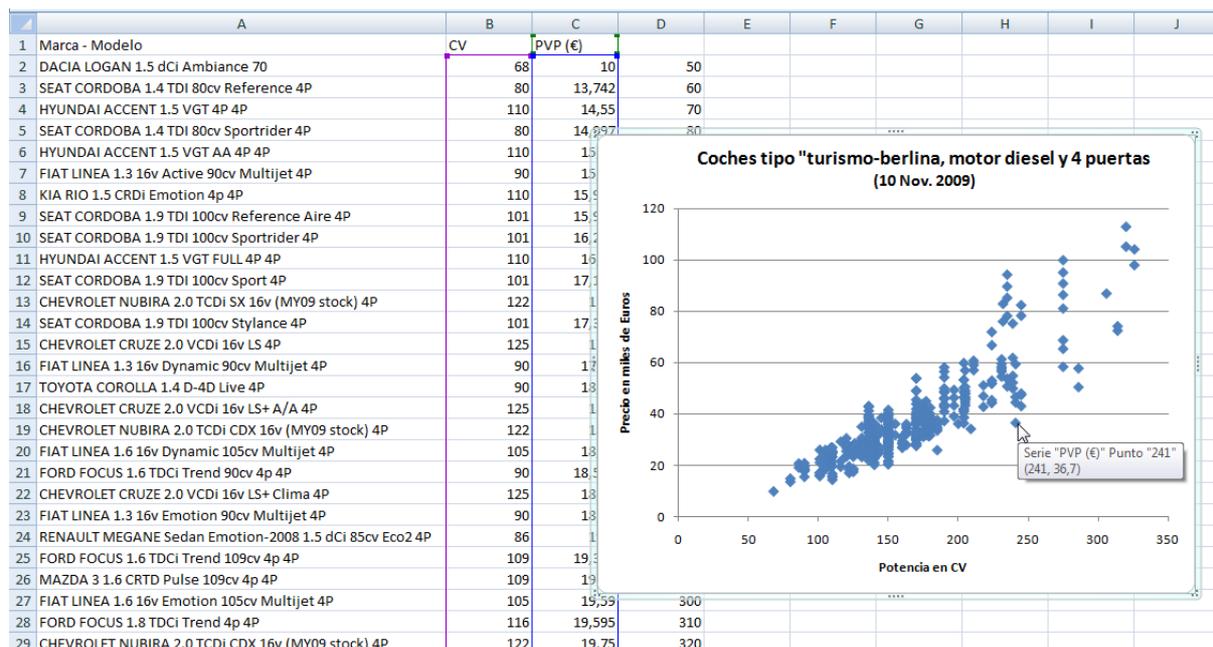


Figura 3: Diagrama bivalente construido con la hoja de cálculo Excel que representa el precio frente a la potencia de un conjunto de coches.

**¿Qué se aprende?:** Técnicas de análisis gráfico de datos. Uso de la hoja de cálculo para construir gráficos.

**Otras actividades:** Pueden realizarse investigaciones similares para analizar la relación entre consumo y potencia, o entre consumo y peso del coche. Por supuesto también se pueden realizar con motocicletas e incluso con aviones.

#### 4. Problemas relacionados con la estimación de las características de la población

Estos son problemas típicos: estimar medias, proporciones,... (¿qué proporción de hogares tiene conexión a internet?, ¿qué proporción de votos obtendrá tal partido en las próximas elecciones?). En este apartado se presentan dos casos en los que el objetivo es estimar el tamaño de la población.

##### 4.1. ¿Cuántos peces hay en un lago?

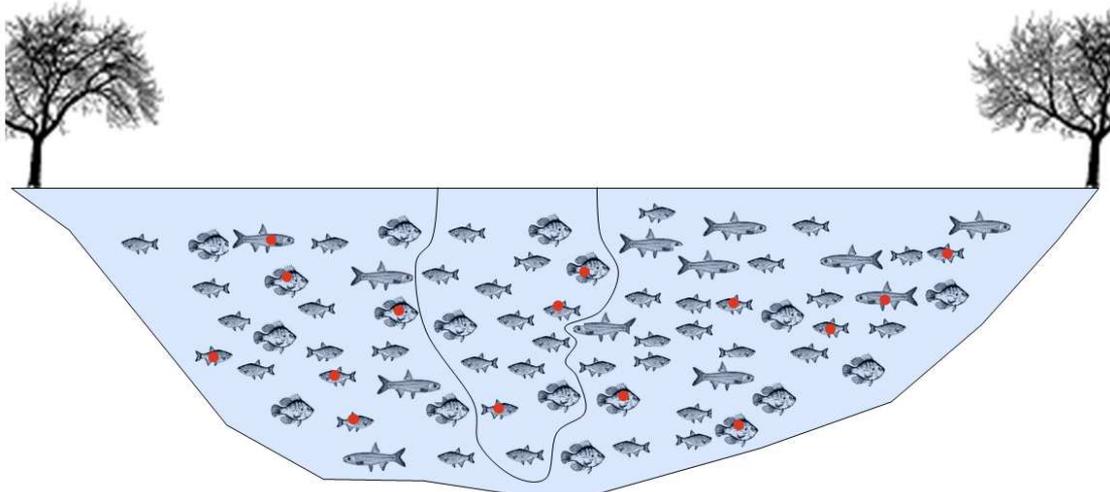
Parece imposible que se pueda responder a esta pregunta, especialmente si el lago es grande y las aguas turbias, pero utilizando métodos razonables (no hace falta vaciar el lago, ni pescar todos los peces, como dicen algunos estudiantes cuando se les pregunta) se puede tener una aproximación suficientemente buena y, lo que también es muy interesante, podemos conocer cuál es la “calidad” de la estimación realizada.

Se trata del método de la pesca y repesca: se capturan  $M$  peces, se marcan y se devuelven al agua<sup>3</sup>, se deja pasar un tiempo en que se supone que los peces marcados se han dispersado por el lago y se capturan otros  $C$ , de los cuales  $R$  aparecen marcados. Si  $N$  es el número total de peces, es razonable considerar que la proporción  $M/N$  (total de peces marcados respecto al total de peces del lago) será parecida a  $R/C$  (peces que aparecen marcados en la segunda captura –repesca– respecto al total de peces capturados) y ya podemos despejar  $N$  para tener una primera aproximación del número de peces que hay en el lago.

$$\frac{M}{N} \cong \frac{R}{C} \rightarrow \hat{N} = \frac{M \cdot C}{R}$$

Por supuesto existen métodos de estimación más precisos, incluso los hay que tienen en cuenta la posible tasa de mortalidad debida a las marcas o la pérdida de ellas. Pero la estadística no se conforma con dar un valor concreto (lo que llamamos “estimación puntual”, más o menos aproximado), sino que trata de dar un intervalo dentro del cual muy probablemente estará el verdadero valor buscado.Cuál es la calidad de la estimación y de qué depende puede indagarse a través de la simulación, mediante *applets* que están disponibles a través de internet o usando una hoja de cálculo.

<sup>3</sup> Esto no es fácil. Hay que hacerlo de forma que la marca no merme las posibilidades de supervivencia y de movilidad del pez, y también debe ser una marca que no desaparezca fácilmente.



**Figura 4: Repesca:** dejamos que se dispersen los peces marcados y volvemos a pescar

Por supuesto, esta técnica no sólo sirve para estimar cuantos peces hay en un lago, en general puede servir para estimar la abundancia de una determinada especie en un territorio.

**¿Qué se aprende?:** Existen técnicas “ingeniosas” que permiten estimar valores que parecen imposibles de conocer. No se pretende acertar el número exacto, pero sí tener un orden de magnitud, que será tanto más preciso cuantos más recursos (número de peces pescados y número de repescados, e incluso número de repescas que se hacen) se destinan a estimarlo. Es posible evaluar la calidad de la estimación realizada.

**Otras actividades:** ¿Cómo se podría tener una buena estimación de cuantos granos de arroz hay en un kilo si no tuviéramos ningún aparato de medida (en concreto, no se puede pesar ninguna cantidad).

#### **Más información en:**

Una buena descripción, en tono de divulgación, se encuentra en el artículo: “*How Many Fish are in the Pond?*” de Roger W. Johnson, que se puede bajar de: <http://ts.rsscse.org.uk/gtb/contents.html>.

En general, todos los artículos de esta página (selección de los mejores artículos de la revista *Teaching Statistics*) contienen buenas ideas para la enseñanza de la estadística a nivel introductorio.

Muchos ejemplos y ejercicios de matemáticas y estadística aplicados a la ecología se pueden encontrar el libro de los profesores de Universidad Autónoma de Barcelona Josep Piñol y Jordi Martínez-Vilalta: “*Ecología con Números*”, Ed. Lynx. Contiene un CD con *applets* muy interesantes, que también se pueden bajar de: <http://www.ecologiaconnumeros.uab.es/> (Applets de los modelos > Listado de applets > Tema 3. Tamaño y estructura de las poblaciones > Marcaje y recaptura).

Hoja de Excel: <http://www-eio.upc.es/~grima/PecesSimulacionExcel.xls>

Presentación en Power Point: [http://www-eio.upc.es/~grima/Peces\\_y\\_Taxis.ppt](http://www-eio.upc.es/~grima/Peces_y_Taxis.ppt)  
Incluye otros temas pero puede ser útil.

#### 4.2. ¿Cuántos taxis hay en una ciudad?

En España lo habitual es que los taxis lleven a la vista un número de licencia (al menos en las grandes ciudades) y que sea un número correlativo desde el 1 hasta un valor igual al número de licencias que existan. El número de licencias es fijo, cuando alguien quiere ser taxista debe comprar la licencia a otro que la desee vender, cambiando el titular de la licencia pero no su número. En estos casos de población numerada no es necesario utilizar el método de la pesca y repesca. Por ejemplo, si  $\mu$  es la media de una población numerada de forma correlativa siempre ocurre que su número de elementos es igual a  $2\mu - 1$  y utilizando la media de la muestra  $\bar{x}$  como estimador de la media de la población podemos proponer que el estimador sea  $2\bar{x} - 1$ . Por ejemplo, si los datos son 16, 28, 45, 48, 68, 72 y 81, como  $\bar{x} = 51,14$ , nuestra estimación podría ser  $N = 2 \cdot 51,14 - 1 \cong 101$ .

Parece un buen método pero tiene un punto débil muy evidente: supongamos que los valores de la muestra son 3, 4, 6 y 15, la media es 7 y por tanto nuestra estimación sería 13, un valor evidentemente falso. Una opción que no presenta este problema es considerar que, por simetría, el número de valores que irán a continuación del último será similar al número de los que hay antes del primero. Esta estrategia también es razonable y nunca dará resultados evidentemente falsos, sin embargo, tiene el inconveniente de que no aprovecha toda la información disponible. La única incógnita es saber cuántos elementos hay a continuación de la última observación y una buena idea es añadirle el promedio que se calcula con los elementos que hay antes del primero, entre el primero y el segundo, entre el segundo y el tercero, etc. Si  $n$  es el número de observaciones que se tienen y  $X_i$  representa el valor de la observación  $i$ -ésima, este promedio será:

$$\frac{(X_1 - 1) + (X_2 - X_1 - 1) + (X_3 - X_2 - 1) + \dots + (X_n - X_{n-1} - 1)}{n} = \frac{X_n}{n} - 1$$

Y, por tanto, la estimación del número total de elementos queda:  $\hat{N} = X_n + \frac{X_n}{n} - 1$

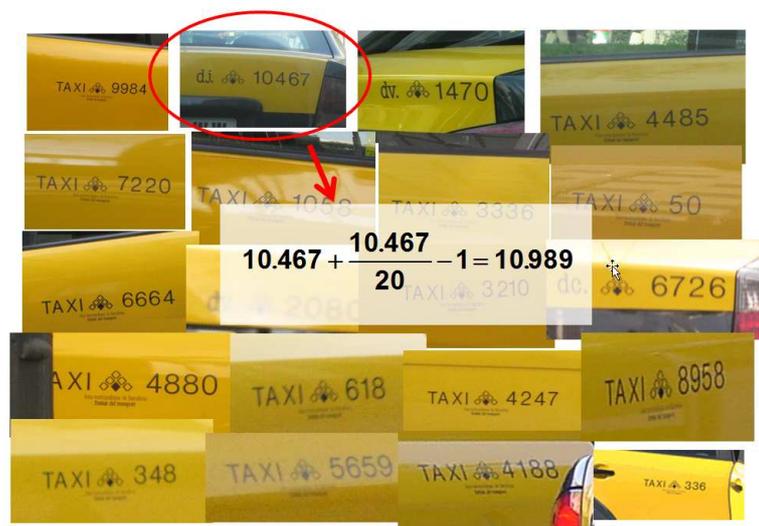


Figura 5: Estimación del número de taxis que hay en Barcelona

Este es un estimador con un comportamiento excelente, con todas las propiedades deseables. Y no solo sirve para contar taxis, también puede servir para estimar el número de participantes en una carrera popular, en la que a cada participante se le da un dorsal con un número correlativo.

**¿Qué se aprende?:** El método de razonamiento seguido para deducir la expresión del estimador. Que en una situación como ésta es posible realizar una buena estimación con muy poco esfuerzo.

**Otras actividades:** Se puede practicar en clase construyendo pequeñas tarjetas de cartulina numeradas, del orden de 1000 o 1500, introducirlas en un recipiente y sacando unas pocas deducir cuantas hay. Es relativamente fácil si se utiliza una hoja de Excel y de forma automática se coloca un número en cada celda (hay que cambiar el tamaño, 2x4 cm está bien), a continuación se imprimen y se recortan con una guillotina.

Si se quiere estimar el número de taxis que hay en Barcelona utilizando este método no hace falta que visiten la ciudad (aunque cualquier excusa es buena). Se cita una revista para los taxistas del área metropolitana y hacia el final se incluye un apartado que publica los números de licencia de los taxis que han encontrado objetos perdidos en su vehículo. Esta lista se puede considerar una muestra aleatoria de números de licencia y... ya es muy fácil.

#### Más información en:

La idea está sacado del artículo "*Estimating the Size of a Population*" de Roger W. Johnson, que se puede bajar del mismo lugar que el comentado para el método de la pesca y repesca: <http://ts.rsscse.org.uk/gtb/contents.html>.

La presentación en Power Point citada en el apartado anterior también contiene una parte que trata este tema.

### 5. Problemas de cálculo de probabilidades que despiertan interés

La estadística se apoya en el cálculo de probabilidades para sacar conclusiones pero no creo que sea acertado introducir al principio de los cursos introductorios de estadística unas cuantas lecciones de cálculo de probabilidades ya que desvía el centro de atención que debe tener este tipo de curso y creo que tiene más inconvenientes que ventajas.

De todas formas, también es verdad que este tema puede ser apasionante, que forma parte del contenido obligatorio de muchos cursos de estadística y que se pueden resolver problemas, incluso complicados, de una forma que puede ser divertida y motivadora.

#### 5.1. Coincidencia en las fechas de cumpleaños

En una clase de 30 estudiantes ¿cuál es la probabilidad de que dos o más celebren su cumpleaños el mismo día? Empezamos calculando la probabilidad de que dos personas no hayan nacido el mismo día. La primera no tiene restricciones, ha podido nacer cualquier día del año (365 casos favorables sobre 365 posibles), pero la segunda ha podido nacer cualquier día menos el que ha nacido el primero (364 días favorables sobre 365 posibles):

$$\frac{365}{365} \cdot \frac{364}{365} = 0,9973$$

De forma análoga, la probabilidad de que tres hayan nacido en días distintos será:

$$\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} = 0,9918$$

Y la probabilidad de que 30 hayan nacido en días distintos:

$$\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \dots \cdot \frac{336}{365} = 0,2937$$

Solo hay dos posibilidades: todos han nacido en días distintos o al menos dos han nacido el mismo día. Luego la probabilidad de que en un grupo de 30 al menos dos hayan nacido el mismo día será:

$$1 - 0,2937 = 0,7063$$

La probabilidad es mayor de lo que intuitivamente se supone. Esta es una carta que tiene en la manga el profesor para impresionar a los alumnos. Con ayuda de una hoja de cálculo se puede construir la curva que relaciona el número de personas y la probabilidad de coincidencia.

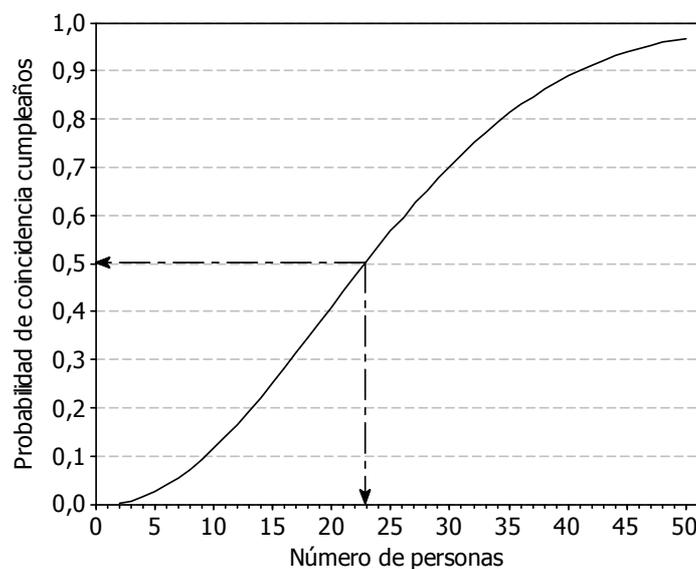


Figura 6: Probabilidad de dos o más personas de un grupo celebren su cumpleaños el mismo día, en función del número de integrantes del grupo

**¿Qué se aprende?:** Las reglas del cálculo de probabilidades permiten deducir probabilidades que evaluadas de forma intuitiva conducen a resultados erróneos. (Moraleja: las reglas para el cálculo de probabilidades son útiles)

**Otras actividades:** En un grupo de 23 personas la probabilidad de que dos o más celebren su cumpleaños el mismo día ya es ligeramente mayor del 50% (exactamente del 50,7%). ¿Dónde hay 23 personas?: en un campo de futbol: 11+11+1. En la prensa deportiva (que se puede consultar por internet) es fácil encontrar la alineación de los partidos, y también es fácil encontrar las fechas de

nacimiento de los jugadores (en las páginas web de los clubes o incluso en la wikipedia). Si se elige una jornada de la primera división en España, hay 20 equipos y se juegan 10 partidos, casi seguro que en algún partido dos personas de las que estaban en el campo celebraban su cumpleaños el mismo día.

Otra actividad puede ser estimar estas probabilidades por simulación con una hoja de cálculo <http://www-eio.upc.es/~grima/SimulacionCumpleaños.xls> (Excel 2007).

## 5.2. Lotería 6/49: ¿Cuál es la probabilidad de que una misma combinación ganadora salga dos veces?

Una persona juega toda su vida adulta (pongamos 50 años) a la lotería primitiva, si se realizan dos sorteos por semana, ¿cuál es la probabilidad de que durante ese periodo salga más de una vez la misma combinación ganadora?

Aunque hay otras variantes, lo habitual es elegir 6 números del 1 al 49, y hay 13.983.816 formas de hacerlo (son combinaciones de 49 elementos tomados de 6 en 6) de las cuales solo una es la ganadora.

Suponiendo que esta persona juegue 100 veces al año, jugará 5000 veces a lo largo de su vida y resulta que el problema planteado es análogo al del cumpleaños, pero es como si tuviéramos un año con 13.983.816 días y 5000 personas cada una de las cuales ha nacido uno de esos días ¿Cuál es la probabilidad de que dos o más hayan nacido el mismo día? Aplicado las fórmulas que hemos visto (una hoja de cálculo es imprescindible) resulta que esta probabilidad es del 59%. Luego no es raro que, si lleva bien las cuentas, descubra que ha tocado dos veces la misma combinación.

**¿Qué se aprende?:** Como en el caso anterior, que a veces la intuición falla al evaluar probabilidades. También se tiene un ejemplo de problema que, aunque aparentemente es muy distinto al anterior, en esencia es el mismo y se resuelven de la misma forma.

**Otras actividades:** En España se empezó a realizar este tipo de sorteo en 1985. En <http://www.onlae.es/primitiva/historicoTodosLosSorteos.aspx> se pueden consultar todas las combinaciones ganadoras hay hoy ¿Ha salido alguna vez repetida? Buscarlo es laborioso aunque con una hoja de cálculo se puede hacer de forma eficiente. La respuesta es sí, el 22 de agosto de 2002 y el 10 de diciembre de 2009 salió la misma combinación ganadora.

## 6. Miscelánea

Existen muchos casos prácticos que se pueden utilizar para enseñar y motivar sobre las posibilidades de la estadística. En este apartado se comenta uno de ellos, el primero que aparece en un libro muy conocido, y también las posibilidades de usar la prensa como fuente de material para la discusión y el aprendizaje.

### 6.1. La vacuna contra la poliomielitis

La posibilidad de inmunizarse frente a una enfermedad infecciosa seguramente ha sido uno de los descubrimientos que más impacto ha tenido en la mejora de la salud y en la esperanza de vida. Pero cada enfermedad requiere su vacuna específica, y dar con ella no siempre es fácil. Existen varios procedimientos para prepararlas y aunque en experimentos con animales, o con humanos a pequeña

escala, se pueden tener bastantes pistas sobre su nivel de eficacia, antes de dar una vacuna por buena y recomendar su uso masivo es necesario estar muy seguros de que sus beneficios compensarán los costos y los riesgos que inevitablemente se asumen.

En 1954 se realizó una prueba a gran escala para evaluar la eficacia de una vacuna contra la poliomielitis (la vacuna Salk, desarrollada por el epidemiólogo Jonas Salk) y el proceso que se siguió está muy bien explicado en el libro de J. Tanur et al.: “La estadística: una guía de lo desconocido”<sup>4</sup>. Se justifica la necesidad de realizar una prueba a gran escala, se comentan las diferentes alternativas que se consideraron en el diseño de la recogida de los datos y las dificultades encontradas al llevarlo a cabo, se justifica la necesidad de tener un grupo de control y aparecen conceptos como “control por placebo” o técnica de “doble ciego”. Está escrito en un tono de divulgación, perfectamente al alcance de un estudiante de enseñanza secundaria, y pone de manifiesto el papel clave que juega la estadística en este tipo de estudios. Un aspecto importante es que al final, de éste y de todos los casos que se presentan, se incluye un cuestionario sobre el contenido del caso.



Figura 7: Libros que contienen casos prácticos sobre el uso de la estadística (los dos de arriba son el mismo, en inglés y en español). Ver bibliografía

<sup>4</sup>Capítulo 1: “El mayor experimento de la historia en el campo de la sanidad pública: la gran prueba de la vacuna Salk contra la poliomielitis (1954)” escrito por el profesor Paul Meier de la Universidad de Chicago.

Lamentablemente, la versión española del libro está agotada, aunque seguramente es fácil encontrarlo en las bibliotecas universitarias. En inglés existe un nuevo volumen, editado con mismo espíritu que el primero, y que contiene nuevos casos. En la Facultad de Matemáticas y Estadística de la UPC también editamos un libro de este estilo que puede ser descargado de <http://upcommons.upc.edu/e-prints/handle/2117/7915>.

## 6.2. La estadística en la prensa (y en la publicidad)

A veces en la prensa se encuentran errores o malos usos de la estadística y esos recortes se pueden usar para, de una forma que puede ser divertida, poner de manifiesto los errores a evitar. Creo que lo mejor es que el profesor tenga su propio dossier de prensa local con noticias o titulares que tengan interés a efectos didácticos. Hay que evitar transmitir la idea de que “la prensa nos engaña” (poco se lee...) seguramente en todas partes hay prensa más o menos seria, y cualquiera se puede equivocar. En el dossier, además de muestras de errores debería haber muestras de estudios o gráficos bien hechos, que resalten de forma clara la información que contienen los datos.

# Alerta por la desprotección infantil ante videojuegos violentos

Amnistía Internacional dice que el sector no está bien regulado y pide al Gobierno que intervenga | El 65% de los menores de 10 a 17 años admiten que acceden a programas para mayores de edad

Aun con premisas tan poco edificantes, los adultos pueden hacer lo que quieran. «El problema está en que el 50% de los niños y el 15% de las niñas de entre 10 y 17 años reconocen usar habitualmente videojuegos destinados a mayores de 18 años», dijo Baltà. Estos porcentajes



Figura 8: Ojo con las operaciones con porcentajes. Tomado de: Sara Fontdecaba y María Montón: “Estadística en la prensa. Estudio crítico” (ver referencias)

**¿Qué se aprende?:** La estadística es omnipresente en los medios de comunicación, a veces se hace un buen uso y otras no tanto. A veces las apariencias engañan y hay que estar atento.

**Otras actividades:** Los mismos estudiantes pueden participar aportando recortes de prensa o de publicidad que se pueden comentar.

### Más información en:

La página web: [www.malaprensa.com](http://www.malaprensa.com) contiene ejemplos de gráficos poco afortunados y ejemplos de mala interpretación de estadísticas

Sara Fontdecaba y María Montón: “Estadística en la prensa. Estudio crítico” en “Estadística en Acción” publicado por la Facultad de Matemáticas y Estadística de la UPC. Puede descargarse el [libro completo](#). Este es el capítulo 15.

**Bibliografía**

- Grima, P. (editor) (2008). *Estadística en acción. Publicado por la Facultad de Matemáticas y Estadística de la UPC. Barcelona, España.*
- Huff, Darell (1993). *How to Lie with Statistics.* Editorial W. W. Norton & Company, USA
- Johnson, R.W. (1994). *Estimating the Size of a Population. Teaching Statistics. Volume 16, Issue 2, pages 50–52.*
- Johnson, R.W. (1996). *How Many Fish are in the Pond?. Teaching Statistics. Volume 18, Issue 1, pages 2-5.*
- Moore, David S. (2007). *Introduction to the Practice of Statistics.* Editorial W. H. Freeman, USA.
- Nightingale, F. *et al.* (1999). *Florence Nightingale: Measuring Hospital Care Outcomes.* Editado por Joint Commission on Accreditation of Healthcare.
- Peck, R. *et al.* (2005). *Statistics: A guide to the unknown.* Editorial Duxbury Press. USA.
- Smith, Gary (1997). *Introduction To Statistical Reasoning.* Editorial Mcgraw-Hill College, USA.
- Tanur, J.M. *et al.* (1992). *La Estadística: Una guía de lo desconocido.* Editorial Alianza. Madrid, España. (Versión original: *Statistics: A guide to the unknown.* Editorial Holden Day, USA.
- Tufte, E.R. (1997). *Visual Explanations.* Editorial Graphics Press, USA.