LA PARADOJA DE SIMPSON

Carmen Batanero, Gustavo R. Cañadas y M. Magdalena Gea

Suma, 71, 27-34, 2012

En estadística y probabilidad encontramos diferentes paradojas, de solución asequible a los estudiantes que permiten organizar actividades didácticas en la enseñanza y aprendizaje. En este trabajo describimos la paradoja de Simpson, que produce múltiples errores en la interpretación de la asociación y correlación. Describimos la paradoja y su historia, algunas soluciones y ejemplos. También analizamos los contenidos estadísticos trabajados en su solución, así como los posibles razonamientos erróneos de los estudiantes.

Palabras clave: *Estadística, paradojas, correlación.*

In statistics and probability we find different paradoxes, whose solution is attainable by the students and that serve to organize didactic activities for teaching and learning. In this paper we describe the Simpon's paradox, which produces many errors in interpreting association and correlation. We describe the paradox and its history, some solution and examples. We also analyze the statistical contents implicit in its solution and some possible erroneous reasoning on the part of the students.

Keywords: *Statistics, paradoxes, correlation.*

Introducción

Aunque la enseñanza de la estadística en secundaria tiene ya una gran tradición, en las últimas orientaciones curriculares se propone renovar su enseñanza, y se insiste en la educación del razonamiento estadístico de los estudiantes (ej. NCTM, 2000; M.E.C., 2006). En estas orientaciones curriculares también se recomienda reforzar las intuiciones de los estudiantes y el razonamiento estadístico.

Una herramienta didáctica posible es utilizar algunas de las paradojas clásicas para crear situaciones didácticas que sirvan para enfrentar a los estudiantes con sus intuiciones incorrectas y hacerlas evolucionar en forma positiva. La historia de la probabilidad presenta situaciones muy atractivas que pueden conducir a reflexionar sobre la presencia del azar en la cotidianidad además de servir de motivación hacia el estudio por parte de los alumnos, algunas de las cuáles hemos analizado en trabajos previos (Contreras, Batanero, Arteaga y Cañadas, 2011 a y b). En este trabajo, describimos la paradoja de Simpson, a la que se ha dado mucha importancia en la literatura sobre alfabetización estadística. En lo que sigue, comenzamos con una reflexión general sobre el interés de las paradojas en la clase de estadística, describimos algunas formulaciones de esta paradoja, analizamos los contenidos matemáticos que se trabajan con ella y posibles dificultades.

Interés didáctico de las paradojas

Algunos autores sugieren el interés de utilizar algunas paradojas sencillas de estadística para plantear situaciones motivadoras en el aula que puedan beneficiar al desarrollar la motivación y meta cognición de los estudiantes y hacerles descubrir conexiones entre la historia y la vida cotidiana.

El uso inteligente de paradojas en la clase de matemáticas apoya una pedagogía constructivista, promoviendo un aprendizaje profundo a partir de las creencias previas y dando al profesor el papel de facilitador del aprendizaje (Lesser; 1998). A la vez, el análisis y discusión de las soluciones a las mismas exige al alumno una reflexión sobre sus propios procesos de pensamientos, lo que es tan importante como el aprendizaje de la solución correcta y un paso vital para alcanzar la capacidad matemática abstracta (Falk y Konold, 1992).

León (2009) indica que podemos servirnos de algunas de estas paradojas clásicas para crear situaciones didácticas que sirvan para provocar la reflexión didáctica de los alumnos. Es sencillo encontrar este tipo de situaciones, ya que la historia de la probabilidad y estadística está repleta de ejemplos, pues su construcción no ha sido sencilla (Batanero, Henry y Parzysz, 2005). Un proceso similar se desarrolla en el aprendizaje de los alumnos que deben construir su conocimiento mediante un proceso gradual, a partir de sus errores y esfuerzo.

En lo que sigue describimos la paradoja de Simpson, mostrando una formulación intuitiva de las soluciones correctas. Analizamos los objetos matemáticos que se trabajan en la solución de esta paradoja y las posibles dificultades de los estudiantes. También presentamos algunos ejemplos en que se ha informado sobre esta paradoja.

Historia de la paradoja

Esta paradoja toma su nombre de Edward H. Simpson, quien la describió en detalle en 1951 en relación con ciertas pruebas médicas, aunque también fue mencionada a principios del siglo XX por el estadístico británico G. Udny Yule, por lo que también es conocida como efecto de Yule-Simpson.

La paradoja es una de las muchas asociadas al estudio de la correlación y muestra que en determinados casos se produce un cambio en la asociación o relación entre un par de variables, ya sean cualitativas o cuantitativas, cuando se controla el efecto de una tercera variable; o, como indican Malinas y Bigelow (2009), dependiendo de los datos con los que trabajemos, la asociación entre dos variables se puede invertir cuando la población de estudio se divide en subpoblaciones. Esto ocurre, según Blyth (1972) cuando se analiza una variable dependiente respecto a otras variables independientes en algún estudio o experimento. Incluso algunas veces una de las variables (la que hace cambiar el tipo o intensidad de correlación entre las otras) es una variable extraña o no controlada, por lo que el investigador puede no ser consciente de este efecto y llegar a una conclusión incorrecta en su estudio.

Simpson (1951) describió la paradoja a partir de tres variables dicotómicas, aunque también se produce en el caso de variables continuas (Gaviria, 1999). Un ejemplo de una situación concreta donde se puede darse viene dado por el enunciado del Problema 1.

Problema 1. *Supongamos que se quiere realizar un estudio comparado de la efectividad de una cierta cirugía en dos hospitales A y B, para lo cuál se obtienen los datos presentados en la Tabla 1. Se pide analizar los datos y determinar cuál hospital da mayor tasa de supervivencia*

Tabla 1. Comparación de supervivencia a una cierta cirugía

	Hospital A	Hospital B
Mueren	63	16
Sobreviven	2037	784
Total de pacientes operados	2100	800

Si analizamos los datos de la Tabla 1, se puede observar que en el hospital A, muere el 3% (63/2100) de los pacientes que se somete a la cirugía y en el hospital B el 2% (16/800), por lo que inicialmente estos resultados nos podrían llevar a pensar que el hospital más seguro para someterse a dicha operación sería el hospital B. La paradoja aparece cuando controlamos los resultados teniendo en cuenta otras variables que influyen en la supervivencia, por ejemplo, el “estado de salud de los pacientes antes de la operación”.

Problema 2. *Estudiemos la efectividad de la misma cirugía en los hospitales A y B, controlando el estado de salud del paciente antes de ser hospitalizado. En las Tablas 2 y 3 se presentan los datos de supervivencia para pacientes de buena salud y salud delicada. Se pide analizar los datos y determinar cuál hospital da mayor tasa de supervivencia en cada tipo de enfermo.*

Tabla 2. Supervivencia de pacientes con buena salud inicial

Pacientes con buena salud	Hospital A	Hospital B
Mueren	6	8
Sobreviven	594	592
Total de pacientes operados	600	600

Tabla 3. Supervivencia de pacientes con salud inicial delicada

Pacientes con salud delicada	Hospital A	Hospital B
Mueren	57	8
Sobreviven	1443	192
Total de pacientes operados	1500	200

Si analizamos ahora la tasa de mortalidad de personas que han sido operadas, para pacientes con buena salud (Tabla 2), se tiene que, del total de los que se operaron en el hospital A, fallecieron un 1% (6/600) y de los que se operaron el B un 1,3% (8/600). Por lo tanto, parece mejor el hospital A para los pacientes de buena salud.

Estudiemos ahora el caso de pacientes con salud delicada: Del total de los que se operaron en el hospital A un 3,8% (57/1500) murieron y en el hospital B el 4% (8/200), por lo que también podríamos concluir que el hospital A es preferible para los pacientes de salud delicada. Es aquí donde se presenta la paradoja, ya que sin tener en cuenta la variable estado de salud de los pacientes, hemos de recomendar el hospital B, pero si separamos los pacientes en dos grupos, los que tenían buena salud y los de salud delicada, en cada uno de estos dos grupos hay que recomendar el hospital A.

En esta paradoja observamos un propiedad contra intuitiva de las probabilidades condicionadas (Blyth, 1972). Para ciertos sucesos *M, A* (con *A’* complementario de *A*) es posible que se dé que

$$P(M / A) > P(M /A')$$

Y que al mismo tiempo

$$P(M \ / \ A \cap S) \leq P(M \ / \ A' \cap S)$$
$$P(M \ / \ A \cap S') \leq P(M \ / \ A' \cap S')$$

Para otro suceso S y su respectivo complementario *S'*. En nuestro ejemplo, sea

M = "Morir en la operación", *A*= "operarse en el hospital *A* y *S*= "tener buena salud". Observamos que:

$$P(M / A)=0{,}03 > P(M / A')=0{,}02$$

$$P(M / A \cap S) = 0{,}01 \leq P(M / A' \cap S) = 0{,}013$$

$$P(M / A \cap S') = 0{,}038 \leq P(M / A' \cap S') = 0{,}04$$

Aunque esta propiedad sea contra intuitiva, la paradoja se resuelve observando la frecuencia total de pacientes que se opera en uno y otro hospital en las tablas 2 y 3. La mayoría de los pacientes que se operan en el hospital A tienen una salud delicada y es de esperar que este tipo de pacientes tenga mayor dificultad en sobrevivir tras una operación, por lo que la tasa de mortalidad en dicho hospital es mayor que en el hospital B, que tiene menos pacientes y la mayoría gozan de buena salud. Es decir el suceso *S* tener buena salud no es independiente del suceso *A* y de ahí la paradoja.

Razonamientos que pueden explicar la paradoja

Solución analítica

Probablemente al llegar a este punto, los alumnos se sentirán sorprendidos y querrán una justificación del cambio de signo de la asociación al controlar por una tercera variable. Un posible razonamiento que les puede convencer es el siguiente (Malinas y Bigelow, 2009): Supongamos que para ciertos números enteros se cumplen las condiciones:

$$\frac{a}{b} < \frac{A}{B} ; \ \frac{c}{d} < \frac{C}{D}$$

Es posible, a pesar de ello, que si sumamos los distintos numeradores y denominadores de las fracciones que aparecen en ambas desigualdades el orden varíe.

$$\frac{(a+c)}{(b+d)} > \frac{(A+C)}{(B+D)}$$

Este hecho, conocido como "reversión de las desigualdades de Simpson" se produce para un gran número de valores *a, b, c, d, A, B, C, D*. Un ejemplo citado por Malinas y Bigelow es el siguiente: $\frac{1}{5} < \frac{2}{8}$ y $\frac{6}{8} < \frac{4}{5}$, mientras que $\frac{7}{13} < \frac{6}{13}$

Solución gráfica

La paradoja puede explicarse gráficamente mediante vectores. Havil (2008) representa mediante letras en una tabla el número de pacientes de uno y otro sexo que se han curado cuando se les han aplicado dos medicamentos alternativos, X e Y:

Tabla 4. Tabla de pacientes curados

	Hombres curados	Mujeres curadas	Curados totales
Medicamento M_1	(a) de (b)	(p) de (q)	(a + p) de (b + q)
Medicamento M_2	(c) de (d)	(r) de (s)	(c + r) de (b + d)

Havil explica la paradoja gráficamente (Gráfico 1) en un diagrama cartesiano, mostrando en el eje X el número de curados y en Y el número de pacientes al que se le ha administrado cada medicamento. Podemos representar las proporciones de curados con cada medicamente por las pendientes de los segmentos representados en trazo grueso.

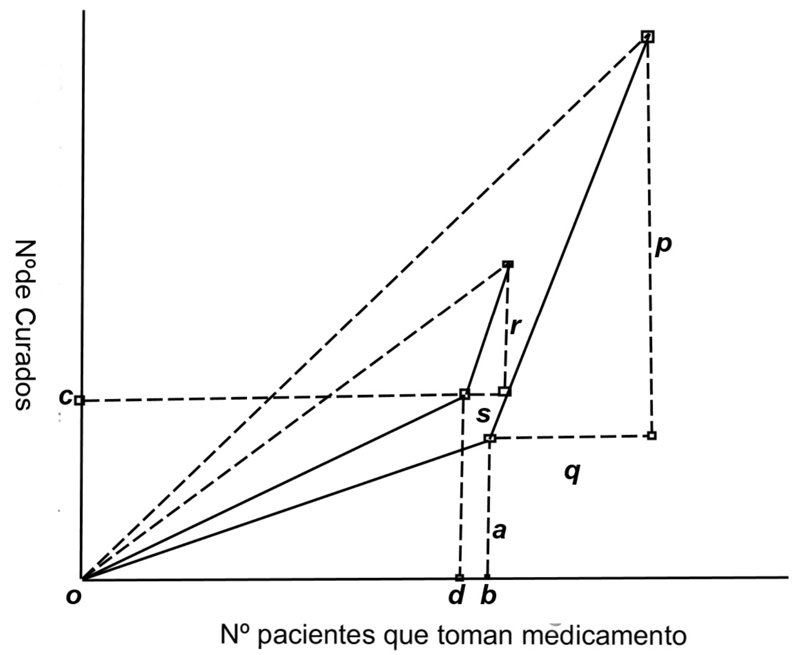

Gráfico 1. Representación gráfica de la tabla (Havil, 2008)

Observamos que dichas pendientes son menores entre los curados con el medicamento M_1 que entre los curados con medicamento M_2 tanto en el caso de hombres (es decir, $\frac{a}{b} < \frac{c}{d}$) como en el de mujeres (es decir, $\frac{p}{q} < \frac{r}{s}$). Sin embargo, cuando consideramos el porcentaje global de pacientes curados (con independencia de su sexo; que se representa por las pendientes de los segmentos con trazos) puede resultar paradójicamente $\frac{a+p}{b+q} > \frac{c+r}{d+s}$ mayor para el medicamento M_1 que para el M_2.

La clave está en que, al sumar vectores, la pendiente de la suma depende no sólo de la pendiente de los sumandos, sino también de su longitud (obviamente, si todas las pruebas se hicieran sobre el mismo número de pacientes no se podría dar la paradoja).

Trabajo en el aula

En el trabajo en el aula con esta paradoja se usarán implícita o explícitamente diferentes objetos matemáticos, que dependen de la solución. Por ejemplo, si se explica usilizando la probabilidad condicional se utilizarían los siguientes (en la clasificación de Godino, Font y Wilhelmi, 2008):

- *Lenguaje matemático:* Se utilizan expresiones verbales, y numéricas de las frecuencias en las distintas celdas de la tabla y los sucesos implicados, así como las tablas de contingencia. También se usará lenguaje simbólico para calcular las probabilidades asociadas.
- *Conceptos*: En esta paradoja los alumnos trabajan la idea de experimento aleatorio, suceso, espacio muestral, complementario, frecuencia absoluta y relativa, doble, marginal y condicional, probabilidad condicional, dependencia e independencia.
- *Propiedades:* Algunas propiedades que aparecen en la resolución de estos problemas son: diferencia entre probabilidad condicionada y simple, relación entre probabilidad condicionada, conjunta y simple, complementario, regla de la unión, del producto.
- *Procedimientos:* Algunos procedimientos que podemos encontrar en la resolución de estas paradojas son: cálculo de probabilidades simples, compuestas y condicionadas,
- *Argumentos*: La actividad permite usar el razonamiento deductivo y empírico.

Obviamente estos objetos cambiarían notablemente si se utilizan otras soluciones; en la hemos denominado analítica se trabajan predominantemente fracciones y desigualdades y en la gráfica conceptos geométricos, vectores, pendientes de rectas.

También podemos observar los siguientes procesos matemáticos:

1. *Procesos de materialización - idealización* (pasar de algo que se percibe a algo que no se percibe). Por ejemplo, los hospitales, enfermos y situación de supervivencia a los que hacen referencia la paradoja son objetos imaginarios, que podemos materializar, si por ejemplo, llevamos a cabo una simulación de la situación.
2. *Procesos de particularización – generalización.* Es cuando pasamos de un caso particular, generalizando a una propiedad de un conjunto o viceversa, cuando una propiedad que sabemos es general, la aplicamos a un caso particular. Por ejemplo, sabemos que la suma total de todas las probabilidades de los sucesos en un experimento es la unidad. En cada ejemplo, particularizando llegamos a las probabilidades de los sucesos dados. Así sabemos que la suma de la probabilidad de supervivencia y muerte ha de ser una en cada caso sin tener que calcularla.
3. *Procesos de representación – significación.* Los procesos de representación y significación aparecen continuamente en el trabajo matemático, pues como no podemos operar directamente con objetos ideales, representamos las operaciones sobre los mismos por medio de símbolos o por medio de otros objetos. Por ejemplo, el objeto “probabilidad” lo representamos por la letra P; la probabilidad de un suceso que denominamos A lo representamos mediante $P(A)$.
4. *Procesos de descomposición – reificación.* El alumno que trata de resolver el problema tiene que pasar constantemente de considerar objetos elementales (unitarios) a considerar objetos compuestos de varios objetos elementales (sistémico). Así, cada suceso de un experimento aleatorio es elemental, pero el espacio muestral del experimento es sistémico.

Posibles dificultades en la actividad

Un razonamiento erróneo se comete cuando se interpretan los datos sin considerarlos globalmente, es decir, sin tener en cuenta todas las variables que aparecen en la situación. Posibles explicaciones de este razonamiento son las siguientes:

1. Al calcular la probabilidad no tenemos en cuenta todos los elementos del espacio muestral. Según tomemos las variables que afectan a la asociación entre “supervivencia” y “hospital”, los sucesos toman una probabilidad u otra (Gras y Totohasina, 1995). Supone un fallo en pasar de la idea espacio muestral (intensivo) al espacio muestral concreto (extensivo).
2. En relación al estudio de la asociación entre las dos variables, se olvida una

condición necesaria en dicho estudio que es el aislamiento de otras variables. Esto quiere decir, que los pacientes que asisten a cada hospital, en el ejemplo, debieran suponerse elegidos al azar (lo cual evidentemente no se cumple, pues la tasa de pacientes con buena salud es muy diferente en cada uno de los hospitales). Si existe una variable no controlada que afecta en forma desigual a los dos grupos, el valor real del coeficiente de asociación entre dos variables puede aumentar, disminuir o incluso cambiar de signo.

Casos reales relacionados con la paradoja de Simpson

La paradoja de Simpson tiene una larga tradición en el mundo académico y ha aparecido en muchos casos célebres. Filósofos como Cohen y Nagel (1934) describieron los resultados paradójicos de diferentes tasas de mortalidad por tuberculosis de unos estudios realizados en ciudades americanas teniendo en cuenta la raza, preguntándose si las poblaciones del estudio realmente eran comparables, es decir, si eran son homogéneas, pues al agregar los datos, cambiaron los resultados. Es también bien conocido por los responsables de políticas laborales y no infrecuente es que los cambios en la composición del empleo hagan que el crecimiento global de los salarios y costes laborales difiera del de los distintos grupos laborales, considerados individualmente.

Otro ejemplo, descrito por Bickel, Hammel y O'Connell (1975) se produjo en la Universidad de Berkeley, acusada de discriminar a las mujeres, pues se comprobó, que solo aceptaba el 34% de las solicitudes de ingreso efectuadas por mujeres, frente al 44% en el caso de hombres. En un intento de comprobar la causa, y tratando de ver si la aceptación o no dependían de los departamentos se desglosaron las cifras de solicitudes y aceptaciones de hombres y mujeres en diferentes departamentos, constatándose con sorpresa que en casi todos los departamentos el porcentaje de mujeres admitidas era superior al de los hombres. La explicación se obtuvo por la diferente proporción de mujeres que solicitaron admisión en diferentes departamentos.

Otro caso reciente (Saari, 2001) se produjo cuando en 1999 el Estado de California estableció un sistema de primas para los profesores de los colegios públicos en los que se lograra mejorar el rendimiento de los estudiantes. La evaluación de los centros mostró (con satisfacción de las autoridades educativas) que el rendimiento de los alumnos había mejorado en la mayoría de los colegios, tanto para el grupo racial dominante, como para la minoría (hispanos). Sin embargo, al mezclar los datos de los dos grupos, el rendimiento global bajó en casi todos los centros, debido a una variación de la proporción de hispanos en los centros a lo largo del periodo. Gaviria (1999) indica que es muy importante considerar la posibilidad de la aparición de la paradoja de Simpson al interpretar resultados de evaluación del rendimiento. No hacerlo así llevaría necesariamente a la mala interpretación de los datos.

Implicaciones para la enseñanza

Diferentes autores recomiendan precaución en la práctica estadística, para evitar los interrogantes que surgen de los datos paradójicos, y precaución antes de hablar de relaciones de causalidad, pues el concepto de causalidad es difícil desde el punto de vista filosófico y no es necesario, para los métodos científicos de investigación (Pearl 2000). Cartwright (1979) por su parte afirma que la interpretación de las relaciones causales requeridas por la investigación científica y la toma de decisión racional han de ser cuidadosas, ya que la dependencia de las regularidades y las frecuencias en que los juicios de probabilidad se basan, no son suficientes para representar dichas relaciones.

El trabajo con la paradoja de Simpson puede servir para aumentar la precaución de los estudiantes ante la interpretación precipitada de la correlación como causalidad. Para finalizar, recordamos que González (2004) señala que el uso de la historia con fines didácticos depende del conocimiento histórico del profesor y su iniciativa para adaptar este saber a los intereses y necesidades del grupo, por lo que ha de exponer los avances de la disciplina junto con su estado actual teórico y de aplicabilidad. El estudio de la historia de la probabilidad y de las paradojas asociadas a la misma será entonces un componente importante en la preparación de formadores.

Agradecimientos: Proyecto EDU2010-14947 (MCIN), beca FPU-AP2009-2807 (MCIN) y grupo FQM126 (Junta de Andalucía).

Referencias:

Batanero, C., Contreras, J. M., Cañadas, C., y Gea, M. M. (2012). Valor de las paradojas en la enseñanza de las matemáticas. Un ejemplo de probabilidad. *Novedades educativas* 261, 78-84.

Batanero, C., Contreras, J. M. y Díaz, C. (2011). Experiencias y sugerencias para la formación probabilística de los profesores. *Paradigma, 32* (2), 53-68.

Batanero, C., Henry, M. y Parzysz, B. (2005). The nature of chance and probability. En

G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 15-37). New York: Springer.

Bickel, P.J. Hammel E.A. y O' Connell J.W. (1975) Sex bias in graduate admissions: data from Berkeley. *Science 187*, 398–404.

Blyth, C. R. (1972). On Simpson's paradox and the sure-thing principle. *Journal of the American Statistical Association*, 67, 364-366.

Cartwright, N. (1979) Causal laws and effective strategies. *Noûs*, *13*, 419-437.

Cohen, M., y Nagel, E. (1934) An introduction to logic and scientific method. New York: Harcourt.

Contreras, J. M. Batanero, C., Arteaga, P. y Cañadas, G. (2011a). O dilema dos prisioneiros: valor de o paradoxos na clase da matemáticas. *Gamma, 11, 91-96.*

Contreras, J. M., Batanero, C., Arteaga, P. y Cañadas, G. (2011b). La paradoja de la caja de Bertrand: algunas formulaciones y cuestiones didácticas. *Epsilon, 28*(2), 7-20.

Falk, R. y Konold, C. (1992) The psychology of learning probability. En F. Gordon y S. Gordon (Eds.), *Statistics for the twenty-first century* (pp. 151-164). Washington, DC: Mathematical Association of America.

Fernández, J. A., Batanero, C., Contreras, J. M. y Díaz, C. (2009). A simulação em Probabilidades e Estatística: potencialidades e limitações. *Quadrante,* Número monográfico. "As novas tecnologias no ensino e aprendizagem da Matemática", *XVIII* (1-2), 161-183.

Gaviria, J. E. (1999) La paradoja de Simpson y la interpretación de los resultados de las evaluaciones del rendimiento académico en el sistema educativo, *Revista de educación, 318,* 211-223.

Godino, J. D., Font, V. y Wilhelmi, M. R. (2008). Análisis didáctico de procesos de estudio matemático basado en el enfoque ontosemiótico, *Publicaciones, 38,* 25-48.

González, P. (2004) La historia de la matemática como recurso didáctico e instrumento para enriquecer culturalmente su enseñanza". *Suma, 45,* 17-28.

Gras, R. y Totohasina, A. (1995) Chronologie et causalité, conceptions sources d'obstacles épistémologiques à la notion de probabilité conditionnelle. *Recherches en Didactique des Mathématiques, 15*(1), 49-95.

Havil, J. (2008). *Impossible? Surprising solutions to counterintuitive conundrums.* Princeton, Princeton University Press.

León, N. (2009) La historia como elemento motivador hacia el estudio de la probabilidad: el problema de la apuesta interrumpida. *Sapiens 1,* 69-88.

Lesser, L. (1998). Countering indifference – Using counterintuitive example. *Teaching Statistics, 20*(1), 10-12

Malinas, G. y Bigelow, J. (2009). Simpson's paradox. En N. Edward (Ed.), *The Stanford Encyclopedia of Philosophy*. Online: http://plato.stanford.edu/archives/fall2009 /entries/paradox-simpson

MEC (2006). *Real Decreto 1631/2006, de 29 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación Secundaria Obligatoria.*

NCTM (2000). *Principles and standards for school mathematics.* Reston, VA: Autor.

Pearl, J. (2000) *Causality: models, reasoning, and inference* , New York, Cambridge: Cambridge University Press

Saari, D. (2001). *Decisions and elections. explaining the unexpected*. Cambridge University Press.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, 13, 2*38–241.