

La probabilidad condicional y la reforma educativa: ¿Son compatibles ?

Allan J. Rossman
Dickinson Collage

Thomas H. Short
Villanova University

Journal of Statistics Education v. 3, n. 2 (1995)

Palabras Clave : Teorema de Bayes, Aprendizaje activo, Tecnología.

Resumen:

Demostramos que se puede enseñar probabilidad condicional de acuerdo con muchas de las propuestas del movimiento de reforma en la educación estadística. Al presentar una variedad de aplicaciones de la probabilidad condicional a problemas realistas, planteamos que las actividades interactivas y el uso de tecnología hacen a la probabilidad condicional más entendible, interactiva e interesante para alumnos con distintos niveles de habilidad matemática. Junto con ejemplos específicos, proporcionamos una guía para la implementación de actividades en el aula, así como recomendaciones pedagógicas para promover el interés y la discusión entre los estudiantes.

1. Perspectiva general

1. La década pasada fue testigo del surgimiento y desarrollo de un movimiento de reforma en la educación estadística. Algunas de las características comunes de muchos de los proyectos de reforma específicos en este campo educativo son la promoción de un aprendizaje activo que involucre los estudiantes; el énfasis en el entendimiento conceptual de las ideas fundamentales de la estadística; la introducción de aplicaciones interesantes que involucren datos genuinos para su investigación por parte de los estudiantes; y la motivación para que estos trabajen colaborativamente con sus pares utilizando la tecnología como herramienta para alcanzar cada uno de estos objetivos. El lector que se interese en leer sobre las reformas en la impartición de la educación estadística puede consultar a Cobb (1992), Cobb (1993), Gordon y Gordon (1992) , y Hoaliglin y Moore (1992).

2 Algunos temas tradicionales en probabilidad con frecuencia se sacrifican en los textos y cursos introductorios “reformados” de estadística, ello con el fin de dar espacio a la inclusión de más análisis de datos e inferencia estadística. A menudo, el único contacto de los estudiantes con los conceptos probabilísticos es una breve introducción a las uniones y las intersecciones, junto con las reglas correspondientes para los eventos disjuntos e independientes. La probabilidad condicional y el teorema de Bayes se consideran, en el mejor de los casos, opcionales, por ser innecesarios para la comprensión del contenido estadístico subsecuente.

3 Estamos sin duda de acuerdo con el renovado énfasis en los datos como el componente central de un curso de introducción a la estadística. Así mismo sostenemos que el pensamiento probabilístico es esencial para una ciudadanía educada, lo que por ende garantiza su inclusión en el currículo de estadística. Uno de nosotros (Rossman),

incluye conceptos de probabilidad condicional en un curso de matemáticas para humanidades titulado “Razonamiento Cuantitativo”, mientras que el otro (Short), los incluye en los cursos de introducción a la estadística para estudiantes de humanidades y enfermería.

4 En este artículo, nos proponemos demostrar que las aportaciones de la reforma para la educación estadística pueden aplicarse de manera bastante natural y productiva al estudio de la probabilidad condicional. Presentamos aplicaciones por medio de las cuales los estudiantes pueden desarrollar una comprensión intuitiva de la probabilidad condicional y del Teorema de Bayes, empleando tecnología para explorar sus propiedades, y aplicándolos reflexivamente a una gran variedad de problemas del mundo real.

2. LAS TABLAS DE DOBLE ENTRADA

5 Al discutir las dificultades de los estudiantes para entender la probabilidad condicional, Moore (1992) plantea que la inferencia bayesiana, que se construye a partir de los fundamentos de la probabilidad condicional, no tiene cabida en un curso de introducción a la estadística. Nuestra primera respuesta a esto es que la distinción, sutil pero crucial, entre la $Pr(A|B)$ y la $Pr(B|A)$ se presenta inclusive cuando los alumnos estudian estadística clásica. Muchos estudiantes sucumben a la natural tentación de considerar el p -valor como la probabilidad condicional de que la hipótesis nula sea verdadera dados los datos de la muestra, más que como la probabilidad de haber obtenido tales datos extremos si la hipótesis nula fuera verdadera. El conocimiento de ejemplos de probabilidad condicional aplicada ayudaría a esclarecer la lógica, la interpretación y las limitaciones subyacentes de la inferencia estadística clásica.

6 En segundo lugar, entender esta distinción en las probabilidades condicionales es fundamental para analizar los datos categóricos presentados en una tabla de dos entradas (también conocidas como tablas de contingencia). Por ejemplo, la siguiente tabla clasifica a los miembros del senado de los EE.UU. de 1994, según su partido político y su género:

	Hombres	Mujeres	Total fila
Republicanos	42	2	44
Demócratas	51	5	56
Total columna	93	7	100

7 Es apropiado e importante pedir que los estudiantes evalúen la legitimidad de afirmaciones como “la mayoría de los senadores demócratas son mujeres” y “la mayoría de las mujeres senadoras son demócratas”. La capacidad para interpretar estas dos afirmaciones es una habilidad esencial para analizar las tablas de doble entrada de datos; éste no es un ejercicio extraño en probabilidad condicional.

3. DESCUBRIENDO EL TEOREMA DE BAYES

8 Con la similitud entre las probabilidades condicionales y el análisis de las tablas de doble entrada en mente, planteamos que los estudiantes pueden descubrir por sí mismos el Teorema de Bayes --uno de los resultados más importantes y más aplicables de la

probabilidad condicional--, por medio de la elaboración de tablas de doble entrada. Una aplicación genérica, basada en un ejemplo encontrado en DeGroot (1986), trata de la identificación del origen de una pieza defectuosa.

9 Supóngase que se utilizan 3 máquinas en una fábrica para producir una gran cantidad de piezas idénticas. Las máquinas de producción tienen diversas capacidades. La máquina A tiene una gran capacidad y produce el 60 % de las piezas, mientras que las máquinas B y C producen el 30% y el 10%, respectivamente. El historial de datos indica que el 10% de las piezas producidas por la Máquina A son defectuosas, comparado con el 30% para la máquina B y el 40% para la máquina C. Si una pieza se examina y se encuentra que está defectuosa, ¿cuál máquina tiene la mayor probabilidad de haberla producido? ¿Cuál es la de menor probabilidad? ¿Cuál es la probabilidad condicional, actualizada a la luz de la aparición de la pieza defectuosa, para cada máquina de haberla producido?

10 Para desarrollar su sentido intuitivo para la probabilidad condicional, primero solicitamos a los estudiantes que adivinen las respuestas a estas preguntas. Luego, en vez de presentarles el teorema de Bayes y de hacer que introduzcan las probabilidades apropiadas, les pedimos que construyan una tabla de doble entrada para una población hipotética de piezas en las cuales los porcentajes ajusten exactamente. (Hacemos hincapié en que datos reales de una muestra de piezas exhibirían variabilidad y no seguirían los porcentajes perfectamente). Las siguientes preguntas guían a los estudiantes para llenar la tabla:

	Defectuosas	No-defectuosas	Total de fila
Máquina A			
Máquina B			
Máquina C			
Total de columna			100

(a) ¿De cada 100 piezas producidas, cuántas fueron hechas por la máquina A? ¿Cuántas por la máquina B? ¿Cuántas por la C? Ponga estos datos en los totales de fila de la tabla.

(b) ¿De esas piezas producidas por la máquina A, cuántas esperarías encontrar defectuosas? Repita lo mismo para las máquinas B y C, y anote sus resultados en la columna “Defectuosas”.

(c) ¿Cuántas del total de 100 piezas en su tabla son defectuosas? Incorpore el resultado como el total de columna para la columna “Defectuosas”.

(d) ¿Del número de piezas que se esperaba que fueran defectuosas, qué proporción fueron producidas por la máquina A? ¿Cuántas por la máquina B? ¿Cuántas por la C?

11 La tabla resultante es la siguiente:

	Defectuosas	No-defectuosas	Renglón total
Máquina A	6	54	60
Máquina B	9	21	30
Máquina C	4	6	10
Columna total	19	81	100

12 Los estudiantes pueden leer directamente en esta tabla que entre las piezas defectuosas, 6/19 son producidas por la máquina A, 9/19 por la máquina B, y 4/19 por

la máquina C. Esto puede también entenderse como las probabilidades actualizadas de que cada máquina haya producido una pieza defectuosa dada la información (los datos) de que la pieza es defectuosa. En este proceso, los estudiantes aplican esencialmente el teorema de Bayes sin percatarse de ello.

13 Contrariamente a la intuición de muchos estudiantes, es la máquina B la que tiene más probabilidad de haber producido la pieza defectuosa. A pesar de ser la menos confiable, la máquina C es la que tiene menos probabilidad de haberla producido, gracias a que, en primer lugar, produce muy pocas piezas. Sin embargo, la probabilidad de que la máquina C haya producido la pieza aumenta un poco más del doble (del 10% al 4/19), a la luz de la evidencia de que la pieza resultó defectuosa.

14 La probabilidad condicional y el teorema de Bayes en ocasiones se introducen utilizando los diagramas de árbol de probabilidad. Mientras que los árboles se pueden construir para representar la estructura de los problemas de probabilidad condicional, creemos que el uso de tablas de doble entrada es más conducente a la organización y al cálculo interactivo de las probabilidades apropiadas. Las tablas de doble entrada también conectan las ideas de la probabilidad condicional con el análisis de los datos de las variables categóricas.

4. APLICACIONES: LA PRUEBA DEL SIDA

15 Habiendo descubierto este análisis de una tabla de doble entrada, los estudiantes pueden aplicar el procedimiento a situaciones más interesantes y más relevantes. Considérese, por ejemplo, la interpretación de los resultados de una prueba de diagnóstico médico. Una prueba muy común para detectar SIDA es la prueba ELISA. Un estudio de Gastwirth (1987) estima que cuando una persona realmente es portador del virus del SIDA, esta prueba da un resultado positivo el 97.7% de las veces. Cuando una persona no es portadora del virus del SIDA, el resultado de la prueba es negativo el 92.6% de las veces. Estos porcentajes se conocen como la sensibilidad y especificidad de la prueba, respectivamente. Más aún, el estudio estima que una tasa base de cerca del 0.5% de la población americana es portadora del virus del SIDA. Esta tasa base proporciona una probabilidad inicial de que un individuo aleatoriamente seleccionado porte el virus. Los datos bajo la forma de resultados de la prueba permiten a uno actualizar la probabilidad inicial para los individuos que se analizan..

16 Una pregunta natural es ¿cuál es la probabilidad de que un estadounidense, aleatoriamente seleccionado, que resulte positivo sea realmente un portador del virus del SIDA? Incluso los estudiantes con solamente habilidades aritméticas básicas pueden abordar esta cuestión construyendo una tabla de doble entrada para una población hipotética de 1,000,000 de personas en la que los porcentajes se ajusten con exactitud. Los estudiantes trabajan con las siguientes preguntas:

	Prueba positiva	Prueba Negativa	Total renglón
Portador de SIDA			
No SIDA			
Total de columna			1.000.000

(a) Utilice la tasa base de la enfermedad en la población para determinar cuántas de este 1.000.000 de personas portarían el virus del SIDA. ¿Cuántas de ellas quedan como no portadoras?

- (b) Utilice la sensibilidad de la prueba para determinar cuántos de los portadores del virus del SIDA darían positivo en la prueba. ¿Cuántos de ellos serían negativos?
- (c) Utilice la especificidad de la prueba para determinar cuántos de los no-portadores resultarían negativos en la prueba. ¿Cuántos darían positivo?
- (d) ¿Cuál es el número total resultante de personas positivas?
- (e) ¿Del total de personas que resultaron positivas, cuál es la proporción que realmente sería portadora del virus del SIDA?

17 La tabla resultante es:

	Prueba positiva	Prueba Negativa	Total renglón
Portador de SIDA	4,885	115	5,000
No SIDA	73,630	921,370	995,000
Total de columna	78,515	921,485	1.000.000

18. A partir de la tabla, los estudiantes pueden ver con facilidad el resultado contraintuitivo de que la mayoría de los resultados positivos en la prueba va a personas que no son portadoras de la enfermedad. Solamente cerca de un 6.22% de los resultados positivos en la prueba va a personas que portan realmente el virus del SIDA. Los estudiantes pueden comentar con sus demás compañeros para elaborar una explicación por escrito de tan sorprendente resultado.

19 La tecnología computacional permite que los estudiantes automaticen este análisis. Pedimos a los estudiantes que incorporen fórmulas en una hoja de cálculo la cual producirá una tabla para cualquier tasa base, sensibilidad y especificaciones que el usuario introduzca. Los estudiantes pueden entonces investigar fácilmente los efectos de los cambios en los datos, la tasa base, la sensibilidad y la especificidad. Por ejemplo, les pedimos a los estudiantes que utilicen el .0622 como la nueva tasa base para encontrar la probabilidad de ser portador del SIDA para una persona que ha dado positivo a la prueba dos veces (asumiendo que las pruebas sean independientes). También solicitamos a los estudiantes que utilicen la hoja de cálculo para producir las gráficas de las probabilidades iniciales y actualizadas.

20 Como ejercicio final con esta aplicación, un instructor puede desafiar a los estudiantes a que reflexionen sobre las implicaciones de este análisis para el examen de SIDA requerido por empleadores y los antecedentes para donadores de sangre. Es de particular importancia la selección de la tasa base de SIDA en las poblaciones en cuestión. La tasa base del 0.5% en este ejemplo se aplica a la población de los Estados Unidos, pero las probabilidades base ideales para los individuos varían dependiendo de sus factores de riesgo de VIH.

5. APLICACIONES : EVIDENCIA LEGAL

21 Otro contexto importante que pide un razonamiento bayesiano involucra la evidencia legal de naturaleza cuantitativa. Tanto a los jueces como a los miembros de los jurados se les pide a menudo actualizar su evaluación subjetiva de la culpabilidad del acusado con base en la introducción de evidencia probabilística. Los estudiantes con habilidades matemáticas algo avanzadas pueden derivar que el teorema de Bayes indica que

$$\Pr(E|G)\Pr(G)$$

$$\Pr(G|E) = \frac{\Pr(E|G)\Pr(G)}{\Pr(E|G)\Pr(G) + \Pr(E|\text{no } G)\Pr(\text{no } G)}$$

donde G representa la culpabilidad del acusado y E la evidencia en cuestión.

22 Considere el caso de Joseph Jamieson, que fue juzgado en un proceso penal en 1987 en la corte de Juzgado del fuero común de Pittsburgh bajo los cargos de violación de siete mujeres del Distrito de Shadyside de la ciudad, durante el período del 18 de Abril de 1985 al 30 de Enero de 1986. Fienberg (1990) reporta que al analizar las secreciones corporales que se tomaron como evidencia en las escenas de los crímenes, una experta forense concluyó que el asaltante poseía las características de la sangre y las marcas genéticas del tipo B, secretor, PGM 2 + 1-. Posteriormente, ella atestiguó que solamente el .32% de la población masculina del condado de Allegheny tenía estas características sanguíneas y que Jamieson mismo era un tipo B, secretor, PGM 2+1-. Resulta natural preguntar cómo un miembro del jurado debería actualizar la probabilidad de la culpabilidad de Jamieson a la luz de esta evidencia forense cuantitativa.

23 En este caso, la $\Pr(E|G) = 1$ y la $\Pr(E|\text{no } G) = .0032$, ya que si Jamieson no cometió los crímenes, entonces algún otro hombre en el condado de Allegheny presumiblemente lo hizo. Aplicados éstos al Teorema de Bayes según lo presentado anteriormente y, simplificando, se llega a la expresión:

$$\Pr(G|E) = \frac{\Pr(G)}{.9968 \Pr(G) + .0032}$$

donde $\Pr(G)$ representa la evaluación subjetiva de los miembros del jurado de la culpabilidad de Jamieson antes de escuchar la evidencia forense. Los estudiantes pueden utilizar un paquete de hoja de cálculo o una calculadora gráfica para graficar esta probabilidad actualizada de la culpabilidad como una función de la probabilidad anterior. También pedimos a los estudiantes que utilicen la tecnología para calcular la probabilidad actualizada de la culpabilidad para ciertos valores de la probabilidad previa; éstos se convierten en:

Prob. Previa	.5	.2	.1	.01	.001	.00000278	
-----+-----+-----+-----+-----+-----+-----+							
Prob. Actual	.9968	.9874	.9720	.7594	.2383	.0009	

24 Esta tabla revela que si la probabilidad que uno asignaría a la culpabilidad de Jamieson antes de oír la evidencia forense es del 50%, entonces uno debería estar convencido en un 99.68% de su culpabilidad después de oír esta evidencia. Aun si uno considera que la probabilidad de su culpabilidad antes de conocer la evidencia es solamente de 1 en 10, después esta evidencia todavía sube la probabilidad de culpabilidad hasta un 97.2%.

25 La última columna de la tabla requiere de una explicación especial. La defensa en este caso argumentó que la probabilidad de culpabilidad previa debería ser de 1 en 360.000, el número estimado de hombres en el grupo de edad apropiado en el condado

de Allegheny. La probabilidad de la culpabilidad actualizada entonces resulta de, apenas, 1 en 1150, el número de varones con las mismas características de la sangre en el grupo de edad apropiado en el condado de Allegheny. Esta columna de la tabla destaca la importancia de la selección de la probabilidad inicial o de base en este análisis.

26 La tecnología también permite a los estudiantes explorar otra faceta probabilística del argumento de la defensa. La experta forense llegó a la caracterización del tipo B, secretor, PGM 2+1- reuniendo la evidencia de la sangre de los siete crímenes. La tabla de abajo revela la información genética que se podría discernir de cada escena del crimen. Los estudiantes pueden utilizar tecnología para investigar la probabilidad actualizada de la culpabilidad de Jamieson para cada uno de los crímenes por separado y descubrir que la evidencia caso-por-caso es mucho menos incriminante para el acusado.

Víctima	Marcador genético atribuible al asaltante	Proporción poblacional con el marcador genético
A	B,secretor	.08
B	B or O,2+ or 2+1+ or 2+1-	.17
C	B,secretor	.08
D	2+1- or 1+1- or 1-	.26
E	B,secretor,2+ or 2+1-	.0056
F	AB or B,secretor,2+1-	.0048
G	B,secretor	.08
compuesto	B,secretor,2+1-	.0032

27 Esta aplicación permite también que los estudiantes examinen un conjunto de cuestiones éticas. ¿El principio de “inocente hasta que se compruebe culpable” significa que la probabilidad de la culpabilidad previa a conocer la evidencia debe ser cero? Si es así, entonces ninguna evidencia en el mundo puede mover esa probabilidad de cero. ¿Cómo se cuantifica los criterios legales de “más allá de una duda razonable” y la “preponderancia de la evidencia”? ¿Puede o debe el sistema de justicia de los EE.UU. esperar que los miembros del jurado apliquen métodos bayesianos desde su posición como tales? ¿Si no es así, entonces cómo lograrán darle sentido a un testimonio como el del experto forense, de que el .0032 de todos los varones tienen los mismos marcadores genéticos del acusado?

6. CONCLUSIÓN

28 Hemos presentado ejemplos por medio de los cuales los estudiantes pueden desarrollar una comprensión intuitiva de la probabilidad condicional y del teorema de Bayes, aplicándolos reflexivamente a una variedad de usos que involucran datos genuinos. Por otra parte, hemos demostrado que la tecnología puede ayudar a los estudiantes a apreciar las relaciones secuenciales que son el fundamento de la probabilidad condicional. La tecnología puede también facilitar la exploración de la

sensibilidad de los resultados al tamaño de la muestra del estudio y a la selección de probabilidades iniciales. Creemos que la enseñanza y el aprendizaje de la probabilidad condicional se pueden mejorar con elementos tales como el aprendizaje activo, la comprensión conceptual, los datos genuinos, y el uso de tecnología, características todos ellos de la reforma de la educación estadística. Uno no necesita dejar atrás la probabilidad condicional cuando se incluyen ejemplos y actividades importantes e interesantes en los cursos de estadística.

Referencias:

Cobb, G. (1992), "Teaching Statistics," in *Heeding the Call for Change: Suggestions for Curricular Action*, ed. L. Steen, MAA Notes No. 22, Washington: Mathematical Association of America, pp. 3-43.

Cobb, G. (1993), "Reconsidering Statistics Education: A National Science Foundation Conference," *Journal of Statistics Education*, v.1, n.1.

DeGroot, M. (1986), *Probability and Statistics* (2da. Ed.), Reading, MA: Addison-Wesley Publishing Co., Inc.

Fienberg, S. (1990), "Legal Likelihoods and A Priori Assessments: What Goes Where?," in *Bayesian and Likelihood Methods in Statistics and Econometrics (Essays in Honor of George A. Barnard)* (1990), eds. S. Geisser, J. S. Hodges, S. J. Press, and A. Zellner, North-Holland, pp. 141-162.

Gastwirth, J. (1987), "The Statistical Precision of Medical Screening Procedures: Application to Polygraph and AIDS Antibodies Test Data," *Statistical Science*, 2, 213-238.

Gordon, S., and Gordon, F. (eds.) (1992), *Statistics for the Twenty-First Century*, MAA Notes No. 26, Washington: Mathematical Association of America.

Hoaglin, D., and Moore, D., (eds.) (1992), *Perspectives on Contemporary Statistics*, MAA Notes No. 21, Washington: Mathematical Association of America.

Moore, D. (1992), "What is Statistics?," in *Perspectives on Contemporary Statistics*, eds. D. Hoaglin and D. Moore, MAA Notes No. 21, Washington: Mathematical Association of America, pp. 1-17.

Allan J. Rossman
Department of Mathematics and Computer Science
Dickinson College
P.O. Box 1773
Carlisle, PA 17013-2896
rossman@dickinson.edu

Thomas H. Short
Department of Mathematical Sciences

Villanova University
Villanova, PA 19085-1699
short@monet.vill.edu