LECTURA:
UN EJERCICIO DE LANZAMIENTO SUCESIVO DE UNA MONEDA, PARA INTRODUCIR EL VALOR DE *P*

Nicholas P. Maxwell
Universidad de Washington, Bothell

Journal of Statistics Education v.2, n.1 (1994)

Resumen

El *p*-valor puede introducirse con un ejercicio basado en el lanzamiento de una moneda. El instructor lanza al aire una moneda 10 veces sucesivas y pide a un estudiante predecir en voz alta el resultado de cada lanzamiento. Los estudiantes llevan un registro de lo que piensan que está sucediendo en cada uno de los lanzamientos de la moneda. El instructor reporta que cada uno de los resultados dados por el estudiante seleccionado fue correcto. Ante ello, los estudiantes, de manera intuitiva, rechazan la hipótesis nula debido a que el *p*-valor es muy pequeño. Así, con este ejemplo concreto se logra que los estudiantes se percaten de que siguieron intuitivamente la lógica de la inferencia estadística aún antes de estudiar esta materia.

1.- El significado del *p*-valor es esencial para entender la estadística inferencial. Sin embargo, muchos estudiantes tienen problemas para comprender qué es un *p*-valor. Es muy común confundir el *p*-valor con la probabilidad de que la hipótesis nula sea correcta, o con la probabilidad de que la hipótesis experimental alternativa sea incorrecta. (Phillips 1971, p. 80; Freedman et al. 1991, p. 435).

2. El método usual para presentar el *p*-valor es conducir a los estudiantes a través de los pasos de la prueba *t*, explicando todo el vocabulario técnico que identifica cada elemento en el procedimiento (e.g.Witte 1993, p. 303; Freedman et al. 1991, p. 435; Wonnacott & Wonnacott 1990, p. 293; Winkler & Hays 1975, p. 443; Moore & McCabe 1989, p. 466; Hildebrand 1986, p. 333). Una presentación de este tipo es comúnmente muy precisa en su formulación del *p*-valor (es decir, la probabilidad de obtener al menos un valor extremo de *t* como el que se encontró, si la hipótesis nula es verdadera). Para los estudiantes, una introducción al *p*-valor como ésta puede ser poco clara, debido a que no muestra que la regla de decisión es razonable. Como si fueran bayesianos, los estudiantes buscan la probabilidad de que la hipótesis experimental sea correcta. Tal probabilidad parecería apoyar una regla de decisión muy intuitiva: acéptese la hipótesis experimental si es altamente probable que sea cierta. Más que describir una regla inferencial tan razonable y lógica, los textos introductorios parecen describir algo misterioso y peculiar.

3. En realidad, la regla de decisión que subyace a la inferencia estadística es bastante natural, y no necesita en lo absoluto parecer algo misterioso a los estudiantes. Estos podrían apreciar cuán razonable y lógica es la inferencia estadística si se percataran de que las personas con frecuencia piensan intuitivamente de la misma manera. (Hong & O'Neil 1992).

4. Se han sugerido algunas alternativas a la formas tradicionales de enseñanza inicial de las pruebas de significancia. Morgan y Morgan (1984) describen el uso de ensayos de Monte Carlo para ilustrar el *p*-valor, pero sus ejercicios no parecen mostrar con claridad la naturaleza intuitiva de las pruebas de significancia. Bonsangue (1992) se centra en el carácter intuitivo de estas pruebas; presenta elaboradas actividades en el aula que ilustran cómo las personas tratan de minimizar de manera natural los errores Tipo I y Tipo II. A diferencia del método Bonsangue, la técnica que se presenta aquí es eficaz y se apoya en muy pocos materiales.

5. Popham y Sirotnik (1992) proporcionan un ejemplo convincente de una intuición que sigue la lógica de la inferencia estadística. Así, ellos escriben:

> Suponga que Joe y June apuestan tazas de café lanzando una moneda: quien pierde paga el café. Joe se encarga de lanzar la moneda, y June decide pedir cruz en cada ocasión, pensando en ganar aproximadamente la mitad de las tazas de café. Si la moneda cae cara diez veces seguidas, June podría empezar a pensar que hay algo sospechoso en esa moneda... ¡o en la persona que la lanza! (Popham and Sirotnik 1992, p. 48)

Cualquiera puede entender esta historia, aun sin ningún conocimiento previo de estadística o probabilidad. Este ejemplo puede presentarse en la clase, y los estudiantes pueden colocarse en la posición de June. Pueden darse cuenta por sí mismos de que, como June, ellos rechazarían una idea cuando ocurra algo que haga muy poco probable que la idea sea verdadera.

6. Yo utilizo en el aula una variante del ejemplo de Popham y Sirotnik (1992). Para introducir la inferencia estadística, lanzo una moneda al aire diez veces seguidas y selecciono a un estudiante para que adivine el resultado de cada lanzamiento. Después de cada lanzamiento, reporto cómo le fue al estudiante seleccionado y entonces pido al resto de la clase que haga un registro de sus ideas acerca de lo que está ocurriendo. Mientras los estudiantes hacen este registro, yo escribo en el pizarrón si el colaborador acertó y dejo espacio para añadir lo que pensaron los estudiantes en cada momento. El truco para enseñar estadística inferencial es que yo miento: reporto que el estudiante seleccionado dijo correctamente cada resultado del lanzamiento de la moneda.

7. Al principio este ejercicio les parece poco interesante, pero conforme el colaborador va acertando en tres, cuatro, y luego cinco lanzamientos de la moneda, la clase se muestra extremadamente inquieta. Algunos de los estudiantes se alborotan tanto que tengo que pedirles que se tranquilicen para que permitan a los otros escribir lo que piensan a medida que el ejercicio avanza. Durante este ejercicio, trato de actuar con espontaneidad, pero no es necesario un gran esfuerzo de actuación. El ejercicio depende de que los estudiantes se den cuenta de que algo raro está ocurriendo, así que si mi expresión delata el truco, no hay problema. Después del último lanzamiento, les pregunto a los estudiantes qué pensaron y cuándo lo pensaron, y escribo sus opiniones acerca de cada lanzamiento en el pizarrón. En mis clases, todos los estudiantes reportan que al llegar al décimo lanzamiento, ellos han llegado a la conclusión de que algo raro está ocurriendo.

8. Para guiar a los estudiantes hacia el descubrimiento de su regla de decisión en esta situación, les pregunto varias cosas. "¿Qué pueden concluir acerca de lo que ocurrió?". Los estudiantes concluyen que yo mentí o que tenía un truco de magia sobre el que me puse de acuerdo de antemano con el estudiante seleccionado. "¿Por qué llegan a esas conclusiones?" Ellos responden que es debido a la poca probabilidad de que alguien adivine correctamente diez lanzamientos de moneda consecutivos. Cuando los presiono, ellos responden que, asumiendo que la moneda sea legal, que yo no esté mintiendo, y que el estudiante seleccionado no sea un telépata, es extremadamente improbable que éste adivine los diez lanzamientos correctamente. Yo pregunto: "Tomando en cuenta que hay varios

aspectos de la concepción inicial de ustedes acerca lo que estaba pasando, ¿cuál es la mejor manera de resumir sus conclusiones?". La mayoría concuerda (presionándolos un poco) que todo lo que pueden concluir a partir del experimento mismo, es que al menos algún aspecto de su concepción inicial es probable que haya sido falso; para llegar a una conclusión final, deben de considerar lo que saben de cada uno de sus supuestos iniciales para decidir cual descartar.

9. Entonces les pregunto: ¿por qué no decidieron que ocurría algo extraño a partir del segundo o el tercer lanzamiento de la moneda, y por qué no aguardaron hasta el undécimo lanzamiento para llegar a una conclusión. La mayoría de los estudiantes señalan que adivinar tres lanzamientos correctamente no es suficientemente improbable, y que no es necesario esperar hasta los once resultados correctos, porque acertar en diez es ya bastante improbable. Yo señalo que parece que ellos tienen una especie de límite para lo improbable: cualquier cosa menos probable que eso y ellos cuestionan lo que pensaban inicialmente del evento que está sucediendo. La mayoría de los estudiantes están de acuerdo de que esto es razonable.

10. Etiqueto entonces elementos de sus ideas con términos estadísticos. Señalo que iniciaron con ideas específicas de lo que iba a ocurrir: que se trataba de una moneda legal, que el colaborador no era clarividente y que yo estaba reportando correctamente los aciertos del colaborador. Les explico que la idea inicial de lo que está ocurriendo se denomina "Hipótesis nula", que la probabilidad de que ocurran esos resultados, si la hipótesis nula es verdadera, es llamado "*p*-valor", y que el umbral utilizado para decidir cuando rechazar la hipótesis nula se conoce como "alfa".

11. Utilizo este ejercicio justo antes de introducir las pruebas de significancia y lo retomo más tarde, durante el curso, cada vez que los estudiantes parecen perder de vista lo que es el *p*-valor. Debido a que el ejercicio puede recordarse con facilidad, solamente se necesitan breves recordatorios para devolver al estudiante a una clara comprensión del *p*-valor.

12. Hacer este ejercicio antes de introducir las pruebas de significancia es útil con algunos alumnos; otros por su parte disfrutan de que se les permita descubrir la lógica de las pruebas de significancia por sí mismos. Otros estudiantes podrían beneficiarse más al efectuar el ejercicio directamente, después de haber aprendido los procedimientos y la terminología de las pruebas de significancia. Para llegar a todos los estudiantes, puede ser ideal tener dos ejercicios como este, utilizar uno justo antes de introducir este tema y el segundo justo al terminar. De esta manera, la técnica presentada aquí podría muy bien complementar la técnica presentada en Eckert (1994).

13. La mayoría de mis estudiantes parecen sentir que las pruebas de significancia son como pensar a la inversa, pero durante este ejercicio ellos siguen la lógica de estas pruebas sin ningún esfuerzo especial. Saber que esta regla de decisión es intuitiva y que ellos la han utilizado aún antes de que la estudiaran en clase resulta muy importante para muchos estudiantes. De esta manera nos aseguramos de que aprendan que la lógica de la estadística inferencial es algo bastante natural.

Reconocimientos

Gracias a los tres árbitros anónimos por sus muy útiles comentarios al borrador de este documento. Gracias por sus comentarios también a Rachel Maxwell, Janet Robertson, y James Robertson.

Referencias:

Bonsangue, M.V. (1992), "*Is it true that 'blonds have more fun'*?, Mathematics Teacher, 85, 579-581.
Eckert, S. (1994), "*Teaching Hipótesis Testing UIT Playing Cards: A Demostration*", *Journal of Statistical* Education *,* v. 2, n. 1.
Freedman, D., Pisan, R., Purves, R., Adhikari, A. (1991), *Statistics* (2da. Ed.), Nueva York: W. W. Norton.
Hildebrand, D. K. (1986), *Statistical Thinking for behavioral Scientists*, Boston: Duxbury Press.
Hong, E. Y O'Neil, H. F. , Jr. (1992), "Instructional strategies to help learners build relevant mental models in inferential statistiscs", *Journal of Educational Psychology*, 84, 150-159.
Moore, D. S. Y McGabe, G. P. (1989), "*Introductionb to the Parctice of Statistics*", New York: W. H. Freman.
Morgan L. A. Y Morgan F. W.,(1984), "Personal computers, p-values and hypothesis testing", *Mathematics Teacher*, 77, 473-478.
Phillips, J. L. (1971), "*How to Think about Statistics*" (Edición revisada), New York: W. H. Freeman.
Popham, W. J., y Sirotnik, K. A. (1992), "*Understanding Statistics in Education*", Itasca, Illinois: F. E. Peacock Publishers.
Winkler, R. L., y Hays, W. L. (1975), *Statistics Probability, Inference, and decisión* (Segunda edición), New York: Holt, Rinehart, y Winston.
Witte, R. S. (1993), *Statistics* (4° ed.), Fort Worth: Harcourt Brace Jovanovich College Publishers.
Wonnacott, T. H., y Wonnacott, R. J. (1990), *Introductory Statistics*(5° ed.), New York: John Wiley and Sons.

Nicholas p. Maxuell
University of Washington, Bothell
Department of liberal Studies, XB-05
22011 26th Avenue S. E.
Bothell, Wa 98021

NMaxwell@U.Washington.edu