

Nombre: _____

Fecha: _____

Actividad NUMB3RS: El favorito para ganar

En "Apuesta arriesgada" un hombre es asesinado en el hipódromo después de ganar la "Apuesta de 6" (acertar al ganador en seis carreras consecutivas). Don le trae a Charlie una libreta que se halló en el cuerpo del muerto. Ésta contiene una página de datos sobre carreras de caballos y ecuaciones. Charlie determina que las ecuaciones eran para escoger el caballo que llegaría en *segundo* lugar, no en primero, quizá para ser menos notorio. Partes de estas ecuaciones utilizan la función "logit" ($L\bar{O} - j\bar{i}t$), una función de probabilidad específica que se vale de logaritmos y razones de probabilidades. Como la función logit se puede complicar mucho, esta actividad sienta sus bases, a saber, en la relación entre probabilidad, probabilidad relativa y razón de probabilidades.

Considera los siguientes datos tomados de un estudio para determinar las bebidas preferidas por estudiantes de secundaria. Se le preguntó a cada uno si prefería refrescos o jugos de fruta. Usa la tabla para responder las siguientes preguntas. Expresa todas las respuestas en esta actividad como decimales a la centésima más cercana.

	Refresco	Jugo	Total
Chicos	46	71	117
Chicas	37	83	120

La *probabilidad* se define como la razón del número de *resultados favorables* (los que estás contando, sean buenos o no) al número de *resultados totales* (todos los resultados posibles). Supón que los datos de arriba se consideran lo bastante válidos para hacer predicciones sobre las preferencias de los estudiantes de secundaria en general.

1. Completa esta tabla con la probabilidad de cada evento. Por ejemplo, la probabilidad de que un chico prefiera refrescos es el número de chicos que prefieren refrescos dividido entre el número total de chicos.

	Refrescos	Jugos
Chicos		
Chicas		

2. ¿Cuál es la suma de las probabilidades para las preferencias de los chicos? ¿Para las preferencias de las chicas? ¿Por qué es así?

Nota que hay una diferencia entre las preferencias de chicos y chicas. ¿Es una diferencia significativa, o puede deberse al azar? Hay muchos modos de responder, y uno relativamente sencillo, que se relaciona con las matemáticas de "Longshot", se

llama la *razón de las probabilidades*. Como su nombre lo indica, es la razón de las probabilidades relativas de que un evento ocurra en un grupo a la probabilidad relativa de que ocurra en otro. La probabilidad relativa es otro modo de expresar la probabilidad. Es la razón del número de resultados *favorables* al número de resultados *desfavorables* para cada grupo. Por ejemplo, la probabilidad relativa a favor de que las chicas prefieran refrescos es la razón del número que elige refrescos al número de aquellas que no los eligen. Igualmente, la probabilidad relativa de que las chicas prefieran jugos se halla calculando la razón recíproca.

- Completa la misma tabla usando la probabilidad relativa (número de chicos que prefieren lo uno/número de chicos que prefieren lo otro).

	Refresco	Jugo
Chicos		
Chicas		

Finalmente, el modo de medir realmente la diferencia entre las preferencias de chicos y chicas toma la forma de una *relación de probabilidades*: la razón de la preferencia de los chicos a la de las chicas.

- ¿Qué significa si la razón de las probabilidades relativas de los chicos a aquellas de las chicas es exactamente 1?
- Completa la tabla usando las razones de las respectivas probabilidades relativas (probabilidad relativa para chicos/probabilidad relativa para chicas para cada opción).

	Refresco	Jugo
Chicos/Chicas		
Chicas/Chicos		

Supongamos que una de las razones de la preferencia de las chicas a la de los chicos fue 2.35 (lo que no es cierto en realidad). Esto significa que las chicas tendrían una probabilidad 2.35 veces la probabilidad de los chicos de tener aquella preferencia.

- Haz un enunciado similar sobre los resultados de la tabla para el #5. Los científicos procuran distinguir entre las diferencias que ocurren al azar y las que indican que las preferencias de chicos y chicas realmente son distintas. Decimos que la diferencia es "estadísticamente significativa" si es muy improbable que viéramos tal diferencia al azar. En esta situación necesitamos una razón de probabilidades que sea por lo menos de 1.8 para concluir que la diferencia entre las preferencias de los chicos y las chicas es estadísticamente significativa. (Ver las "Extensiones" para más detalles sobre el origen de la razón de probabilidades de 1.8.)
- ¿Podemos estar seguros de que hay una verdadera diferencia entre las preferencias de los chicos y de las chicas?

El objeto de esta actividad es dar a los estudiantes un vistazo breve y sencillo de un tema matemático muy extenso. TI y NCTM lo invitan a usted y a sus estudiantes a aprender más sobre este tema con las extensiones que se ofrecen abajo y con su propia investigación independiente.

Extensiones

Introducción

Esta actividad ha sentado las bases para la función de logit. Los estadísticos emplean la función logit, entre muchas otras pruebas y cálculos, para determinar la fuerza de la relación entre dos variables.

Para el estudiante

1. Demuestra que si el porcentaje de chicos que prefieren refrescos es p y el porcentaje de chicas que prefieren refrescos es q , la razón de probabilidades es $\frac{p(1-q)}{q(1-p)}$.
2. La función logit, para toda probabilidad p entre 0 y 1, se define como $\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$. Si el primer grupo tiene la probabilidad p y el segundo tiene la probabilidad q , entonces el logaritmo de la razón de probabilidades, $\frac{p(1-q)}{q(1-p)}$, es igual a la diferencia de $\text{logit}(p) - \text{logit}(q)$. La función logit es el inverso algebraico de una *función logística* o curva logística. Las curvas logísticas se emplean en muchos tipos de modelos estadísticos, entre ellos los de crecimiento demográfico y propagación de enfermedades o rumores. Halla un ejemplo de una función logística y grafícala. Describe cómo sus propiedades son apropiadas para las aplicaciones citadas aquí.
3. Los medios de comunicación suelen hablar de la propagación de alguna enfermedad (como el VIH, SRSA o gripe aviar) como posiblemente exponencial. ¿Por qué resulta más realista una curva logística que una exponencial para un modelo matemático acerca de la propagación de una enfermedad?

Temas relacionados

La razón de probabilidades se considera estadísticamente significativa si un intervalo de confianza del 95% no contiene el valor 1 (razón de 1:1). En otras palabras, hay confianza del 95% de que la razón real sea mayor (o menor) que 1. Hay una calculadora que está basada en la función logit para hacer esto. Para una explicación de cómo funciona, ver el artículo "The odds ratio" por J. Martin Bland y Douglas G. Altman, *British Medical Journal*, volumen 320, mayo 27 de 2000, pág. 1468. Para que la calculadora halle un intervalo de confianza del 95% para la razón de probabilidades, ver <http://www.hutchon.net/ConfidOR.htm>

Para una actividad relacionada sobre la probabilidad de eventos complementarios, ver la actividad de NUMB3RS "Sorpresa de cumpleaños" del episodio "Tránsito". Para bajar esta actividad, visita <http://education.ti.com/exchange> y busca "7463."

Recursos adicionales

Para los interesados en estudiar la función logit (o "transformación logit", como se llama en "Apuesta arriesgada") y sus aplicaciones, ver:

<http://mathworld.wolfram.com/LogitTransformation.html>