

Nombre: _____ Fecha: _____

Actividad NUMB3RS: ¿Diferente o no?

En "Desechos" el FBI descubre que los niños cerca de un tipo de falla geológica llamada pozo ciego parecen presentar un índice inusualmente alto de cáncer. La agente del FBI Megan Reeves cree que esto podría representar una "agrupación" de casos de cáncer. Charlie le advierte que no se apresure a sacar esa conclusión. El FBI puede haber utilizado un técnica de estadística llamada la *prueba de dos proporciones* para decidir si el índice de cáncer parece ser demasiado alto en este lugar. Una prueba de dos proporciones es una técnica estadística que sirve para decidir si el índice de cáncer parece ser diferente en distintas zonas del país. Esta actividad presenta una introducción a dicha prueba.

Considera los siguientes datos sobre casos de cáncer, basados en una muestra aleatoria hipotética de 50, 000 personas en cada uno de los seis estados de Nueva Inglaterra. En este caso se emplea una muestra grande porque el índice de cáncer es muy bajo.

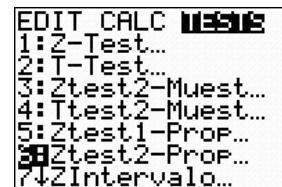
Estado	Casos de cáncer
Connecticut	268
Maine	291
Massachussets	268
New Hampshire	251
Rhode Island	281
Vermont	249

Según estos datos, ¿parece existir una relación entre lugar e índice de cáncer? Un recurso para tratar de responder esta pregunta es el procedimiento estadístico llamado prueba de dos proporciones. Este es uno de muchos procedimientos utilizados para probar hipótesis estudiadas en clases de estadística.

1. Maine y New Hampshire son estados vecinos. ¿Qué porcentaje de la muestra de Maine tenía cáncer? ¿Qué porcentaje de la muestra de New Hampshire tenía cáncer? ¿Te parece que la diferencia entre estos porcentajes es considerable?

Los datos muestran que la proporción de los casos de cáncer en la muestra de Maine es mayor que la proporción de casos de cáncer en la muestra de New Hampshire. Si las muestras se escogieron al azar, entonces una prueba de dos proporciones puede determinar si es probable que la proporción de casos para la población total del estado sea mayor en Maine que en New Hampshire. Una prueba hipotética supone que no hay diferencia y examina los datos para ver si éstos, efectivamente, no respaldan tal suposición. En ese caso, la hipótesis es que no existen diferencias entre índices de cáncer para los dos estados. La hipótesis alternativa supone que el índice de cáncer en Maine es mayor que el de New Hampshire. Una prueba de dos proporciones puede indicar que la hipótesis inicial es cuestionable.

Para hacer esta prueba en tu calculadora, presiona **STAT** y ve al menú de **TESTS**. Luego, selecciona **6:Ztest2-Prop....**



Introduce los datos como sigue:

x1 representa el número de casos de cáncer en la muestra de Maine. **x2** representa el número de New Hampshire. **n1** y **n2**, representan el número de personas en cada muestra (50,000). Como nos interesa saber si la proporción de casos de cáncer es mayor en Maine que en New Hampshire, escoge **>p2**.

```
ZTest2-Prop
x1:251
n1:50000
x2:251
n2:50000
P1:#P2 <P2 >P2
Calcular Dib
```

Luego, selecciona **Calcular** para producir los resultados que ves a la derecha.

```
ZTest2-Prop
P1>P2
z=1.72282231
P=.0424602946
P1=.00582
P2=.00502
↓P=.00542
```

Si las muestras fueran verdaderamente aleatorias, el conjunto de todas las diferencias posibles entre las proporciones de la muestra formaría una distribución normal. La descripción informal de una distribución normal es "una curva de campana". Los resultados de la prueba muestran si la diferencia observada en los datos es típica de la mayoría de las diferencias o si la diferencia es, de hecho, considerable.

Las pantalla muestra un valor de p de aproximadamente 0.042. Este valor significa que si no hubiera diferencia real entre los índices de cáncer de los dos estados, la diferencia observada en estas muestras sucedería al azar aproximadamente el 4.2% de las veces. Muchos estadísticos utilizan el 5% (0.05) como el valor máximo para decidir si una observación es anómala; por lo que, en este ejemplo, concluirían que es improbable obtener al azar resultados como estos. Los estadísticos usan el muestreo para hacer predicciones sobre la población entera. En este ejemplo, es probable que el índice real de cáncer en Maine sea más alto que en New Hampshire, según los datos de la muestra. Para un valor de p mayor que 5%, los estadísticos concluirían que no hay suficiente información para afirmar que uno de los índices es realmente mayor que el otro. Sin embargo, no concluirían que no hay diferencias entre los índices.

2. Connecticut y Rhode Island también son estados vecinos. Compara sus datos de cáncer usando la prueba de dos proporciones. Selecciona **<p2** o **>p2**, según corresponda. ¿Cuál es el valor de p para esta prueba? Explica el significado de esta cifra en el contexto de este ejemplo. ¿Tienes suficiente información para concluir que un índice es mayor que el otro?
3. Los datos de muestra en esta actividad indican que 268 de cada 50,000 personas fueron diagnosticadas con cáncer tanto en Connecticut como en Massachussets. ¿Significa esto que los índices *reales* de cáncer en la población total de los dos estados fueron los mismos? Compara sus datos de cáncer mediante una prueba de dos proporciones. Selecciona **<p2** o **>p2**. ¿Cuál es el valor de p para esta prueba? Explica el significado de tal cifra en el contexto de este ejemplo.
4. ¿Cuál es la relación entre el valor de p obtenido para la prueba **< p2** o **> p2** y el valor para la prueba **≠ p2**?

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

Para el estudiante

1. Compara el índice de cáncer de tu estado con el de estados vecinos usando los datos de los Centros para Control de Enfermedades. El sitio Web para estadísticas sobre cáncer de Estados Unidos trae información basada en género, grupo étnico y tipo de cáncer. Visita: <http://apps.nccd.cdc.gov/uscs>
2. Completa la actividad de identificar la fuente de un brote de la enfermedad de Legionnaire en Louisiana, en el sitio Web de los Centros para Control de Enfermedades: <http://www.cdc.gov/excite/classroom/legionnairesQ.htm>
3. Realiza una encuesta en tu escuela para saber si entre los muchachos y las muchachas existe una diferencia significativa en cuanto a puntos de vista sobre algún tema. Analiza los datos obtenidos empleando las técnicas de esta actividad.

Recursos adicionales

El proyecto de Jóvenes Estudiosos de Epidemiología de la Fundación Robert Wood Johnson realiza una competencia anual para otorgar becas. El sitio Web contiene un buen número de unidades educativas sobre epidemiología. Una actividad, "Investigación de Brotes de Enfermedades", trata de una agrupación de casos de leucemia en Massachussets que sirvió de inspiración para la película "A Civil Action" protagonizada por John Travolta. Visita: <http://www.collegeboard.com/yes>

Para más información sobre índices de cáncer y factores de riesgo, visita: http://www.cancer.org/docroot/STT/STT_2.asp

A veces se sospecha que los casos de cáncer se deben a algo en está en el ambiente, porque aparece una agrupación de casos en una zona geográfica específica. Para más sobre agrupaciones de casos de cáncer, visita: <http://www.cdc.gov/nceh/clusters/about.htm>

Burrill, Gail, Christine A. Frankin, Landy Godbold, y Linda J. Young; *Navigating through Data Analysis in Grades 9–12*, [NCTM, 2003.] Las actividades en este libro familiarizan al estudiante con el muestreo aleatorio simple, las técnicas de muestreo y la simulación como recurso para analizar datos tanto categóricos como numéricos.