

NOMBRE

FECHA

PERIODO

Materiales de apoyo familiar

Estadística inferencial

En esta unidad, el estudiante usará una pequeña *muestra* de datos para hacer una estimación sobre la información de un grupo más grande llamado *población* y utilizar la simulación para determinar un rango de valores para hacer la estimación. Una población es el conjunto completo de temas de interés para una pregunta y una muestra es un grupo más pequeño dentro de esa población.



Por ejemplo, es posible que queramos averiguar la cantidad *media* (o promedio) que las familias en Estados Unidos gastan en alimentos cada mes. La población incluye a todas las familias de Estados Unidos, pero recopilar información de todos sería muy difícil y costaría mucho dinero, por lo que podríamos comenzar a recopilar datos con una muestra de 50 familias.

Una cuestión importante a considerar al comenzar a recopilar información de una muestra es cómo se seleccionará la muestra. Los datos que recopile pueden ser muy diferentes si pregunta a familias que están comprando en un supermercado local en comparación con

NOMBRE

FECHA

PERIODO

preguntas a personas fuera de un restaurante elegante. De manera similar, la cantidad gastada en alimentos en San Francisco probablemente sea muy diferente de la cantidad gastada en las zonas rurales de Iowa. Incluso puede haber algunos hábitos de gasto ocultos de formas en las que aún no hemos pensado. Entonces, ¿cómo puede asegurarse de que su muestra sea representativa de las familias en Estados Unidos sin utilizar demasiadas familias de grupos que no son típicos en sus gastos?

La solución es utilizar la *aleatoriedad*. Podemos seleccionar 50 familias usando un proceso aleatorio, como hacer que una computadora seleccione las familias de una base de datos al azar sin considerar otros factores. Esto debería reducir el *sesgo* que podría ser introducido por humanos que intentan obtener información sobre las personas y probablemente incluirá proporciones más precisas de los diferentes tipos de familias en los Estados Unidos. Si bien la aleatoriedad puede no eliminar por completo el sesgo de la selección de la muestra lo reducirá significativamente en comparación con la selección sin aleatoriedad.

Los investigadores han realizado estudios como estos y han descubierto que la cantidad media que se gasta en alimentos cada mes. Un informe dice que la cantidad media gastada en alimentos cada mes es de \$600 con un *margen de error* (MdE) de \$150. El margen de error se utiliza para decir que no esperamos que cada familia de la muestra gaste exactamente \$600.

Es importante buscar el margen de error en los resultados estadísticos. Es irresponsable discutir estadísticas sin proporcionar un margen de error para describir cuánto se espera que varíe el valor. Muchas gráficas incluidas en los informes de noticias lo muestran en letra pequeña en el gráfico. Busca algo como $\pm 3\%$, cuando hay una gráfica sobre el índice de aprobación de un funcionario o encuestas durante las próximas elecciones. Esto significa que los porcentajes que se muestran en la gráfica podrían ser hasta un 3 % más bajos o más altos que el número mostrado.

Aquí hay una tarea para hacer con el estudiante:

Una ciudad tiene próxima una votación sobre si se deben aumentar los impuestos sobre la renta de las sociedades en un 2 % para aumentar la financiación de las escuelas públicas. El noticiero local muestra una imagen que indica que el 52 % de la población votante está a favor del aumento de impuestos y en la esquina muestra el “margen de error $\pm 3.5\%$ ”. El periodista cree que los impuestos de las sociedades aumentarán porque cualquier voto superior al 50 % a favor de los impuestos aprobará la ley.

1. El periodista que encontró la cifra del 52 % llegó a ella yendo a 4 de los 20 barrios diferentes de la ciudad y preguntando a los residentes su opinión. ¿Hay algún problema con la forma en que se hizo eso? ¿Se te ocurre una mejor manera de recopilar datos?
2. ¿Qué significa el margen de error en esta imagen?
3. ¿Deberías creer que aumentarán los impuestos? Explica tu razonamiento.

NOMBRE

FECHA

PERIODO

Soluciones:

1. Ir sólo a 4 barrios de la localidad podría dejar fuera las opiniones de muchos electores de otros barrios a los que el periodista no fue. Una mejor manera de recopilar información podría ser seleccionar aleatoriamente varios hogares de la ciudad para realizar una encuesta sobre su opinión. Es más probable que la selección aleatoria evite cualquier sesgo que tenga el periodista sobre qué vecindarios visitar.
2. El margen de error significa que el porcentaje real a favor del aumento de impuestos podría ser un 3.5 % mayor o un 3.5 % menor que el 52 % reportado según la muestra. Esto significa que el porcentaje real caería entre el 48.5 % y el 55.5 %.
3. Ejemplos de respuestas:
 - Creo que todavía hay muchas posibilidades de que aumenten los impuestos. Aunque el porcentaje real podría ser tan bajo como el 48.5 % según el margen de error, también podría llegar al 55.5 %. La mayoría de los porcentajes posibles están por encima del 50 %, por lo que creo que el aumento se producirá.
 - Creo que todavía no está claro si se producirá el aumento. Según el margen de error, el porcentaje real podría ser tan bajo como el 48.5 %, lo que haría que el aumento no se produjera. Tampoco estoy seguro de los métodos utilizados por el periodista para recolectar una muestra en este informe, por lo que es posible que el informe no sea muy preciso.



© CC BY 2019 by Illustrative Mathematics®