Lección 17: Margen de Error al Estimar la Proporción de una Población

Trabajo en clase

En esta lección, encontrarás e interpretarás la desviación estándar de una distribución simulada para una proporción de muestras y utilizarás esta información para calcular el margen de error para la estimación de la proporción de la población

Ejercicios 1-6: Desviaciones Estándar para Proporciones

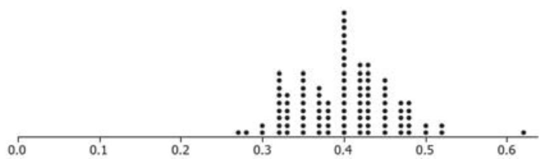
En la lección anterior, utilizaste distribuciones muestrales simuladas para aprender acerca de la variabilidad del muestreo en la proporción de muestra y el margen de error cuando se utiliza una muestra aleatoria para estimar una proporción poblacional. Sin embargo, encontrar el margen de error utilizando la simulación puede ser engorroso y llevar mucho tiempo para cada situación. Afortunadamente, dado el comportamiento consistente de la distribución muestral de la proporción de muestra para las muestras aleatorias, los estadísticos han desarrollado una fórmula que te permitirá encontrar el margen de error de forma rápida y sin simulación.

1. 30% de los estudiantes que participan en deportes en la Secundaria Union son mujeres (una proporción del 0.30).
   a. Si tomaste varias muestras aleatorias de 50 estudiantes que practican deportes e hiciste un diagrama de puntos de la proporción de mujeres en cada muestra, ¿dónde crees que se centrará esta distribución? Explica tu razonamiento.

   b. En general, para cualquier tamaño de la muestra, ¿dónde crees que estará el centro de la distribución simulada de la proporción de muestra de mujeres que practican deporte en la Secundaria Union?

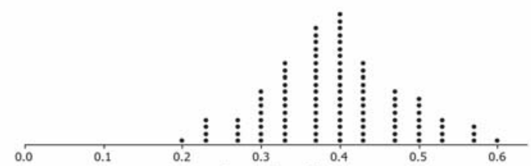
2. A continuación se presentan dos distribuciones muestrales simuladas para la proporción de muestra de mujeres en muestras aleatorias de todos los estudiantes en la Secundaria de la Union.

Distribución Muestral Simulada 2:



Proporción de Muestra

Distribución Muestral Simulada 2:



Proporción de Muestra

a. Basándote en las dos distribuciones muestrales mencionadas anteriormente, ¿cuál crees que es la proporción de población de mujeres?

b. Una de las distribuciones muestrales anteriores se basa en muestras aleatorias del tamaño de 30, y la otra se basa en muestras aleatorias del tamaño de 60. ¿Qué distribución muestral corresponde a la muestra del tamaño de 30? Explica tu elección.

3. Recuerda de tu trabajo anterior en las estadísticas donde las distribuciones se describen utilizando la forma, el centro y la dispersión. ¿Cómo se midió la dispersión?

4. En las lecciones anteriores, conociste la fórmula para la desviación estándar de la distribución muestral de la media de la muestra. También hay una fórmula para la desviación estándar de la distribución muestral de la proporción de la muestra. Para muestras aleatorias del tamaño de n, la desviación estándar se puede calcular utilizando la siguiente fórmula:

$\text{desviación estándar} = \sqrt{\frac{p(1-p)}{n}}$, donde p está el valor de la proporción de la población y n es el tamaño de la muestra.

   a. Si la proporción de mujeres en la Secundaria Union es de 0.4, ¿cuál es la desviación estándar de la distribución de las proporciones de muestra de mujeres para muestras aleatorias del tamaño de 50? Redondea tus respuestas a tres cifras decimales.

   b. La proporción de hombres en la Secundaria Union es de 0.6. ¿Cuál es la desviación estándar de la distribución de proporciones de la muestra de hombres para muestras del tamaño de 50? Redondea tus respuestas a tres cifras decimales.

   c. Piensa en las gráficas de las dos distribuciones en las partes (a) y (b). Explica la relación entre tus respuestas utilizando el centro y la dispersión de las distribuciones.

5. Piensa en las simulaciones que tu clase realizó en la lección anterior y las simulaciones anterior Ejercicio 2 .

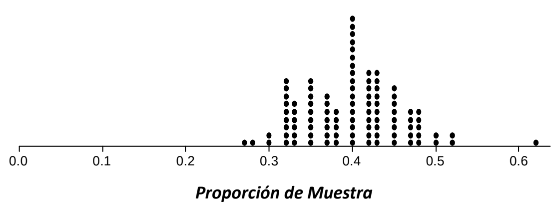
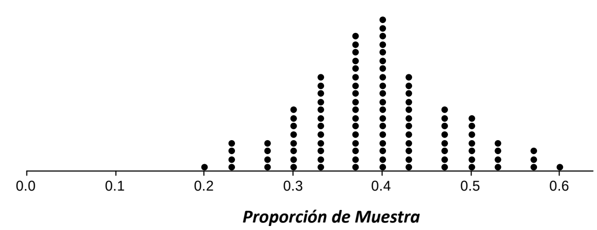
   a. ¿La variabilidad de la muestra en la proporción muestral fue mayor que las muestras del tamaño de30 o que las muestras del tamaño de50? En otras palabras, ¿la proporción de muestra tiende a variar más de una muestra aleatoria a otra cuando el tamaño de la muestra es 30 o 50?

   b. Explica la manera en que la fórmula para la desviación estándar de la proporción de muestra confirma que la observación de la variabilidad en las proporciones de muestra disminuye a medida que se aumenta el tamaño muestral.

6. Considera las dos distribuciones muestrales simuladas de la proporción de mujeres en el Ejercicio 2, donde fue la proporción de la población fue del 0.4. Recuerda que has encontrado $n = 60$ para la Distribución 1, y $n = 30$para la distribución 2.

   a. Encuentra la desviación estándar para cada distribución. Redondea tus respuestas a tres cifras decimales.

   b. Haz un dibujo y marca los intervalos de una desviación estándar de la media para cada una de las dos distribuciones. Interpretar los intervalos en los términos de la proporción de mujeres en una muestra.

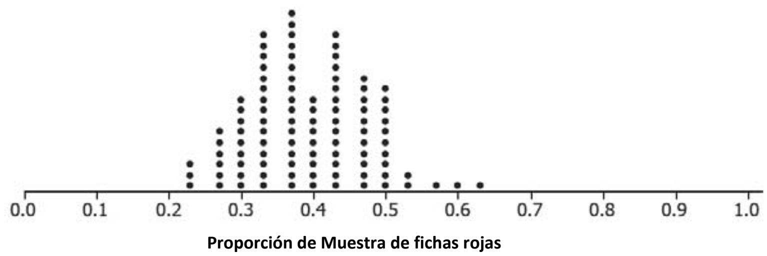
Distribución Muestral Simulada 1: $n = 60$	Distribución Muestral Simulada 2: $n = 30$
	

En general, los tres resultados sobre la distribución muestral de la proporción de muestra son conocidos:

- La media de la distribución muestral de las proporciones de muestra será aproximadamente igual a el valor de la proporción de la población, p.
- La distribución muestral de la proporción de muestra es menos variable para muestras más grandes que para muestras pequeñas. La variabilidad en la distribución muestral se describe por la desviación estándar de la distribución, y la desviación estándar de la distribución muestral para muestras aleatorias del tamaño de $n$es $\sqrt{\frac{p(1-p)}{n}}$, donde p es el valor de la proporción de la población. Esta desviación estándar se calcula generalmente utilizando la proporción de muestra, que se denota por $\hat{p}$ (se lee como p-sombrero), para distinguirla de la proporción de la población. La fórmula para la desviación estándar estimada de la distribución muestral de las proporciones de muestra es $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.
- Mientras el tamaño de la muestra es lo suficientemente grande que la muestra incluye, al menos 10 éxitos y fracasos, la distribución muestral es aproximadamente normal en forma. Es decir, una distribución normal sería un modelo razonable para la distribución muestral.

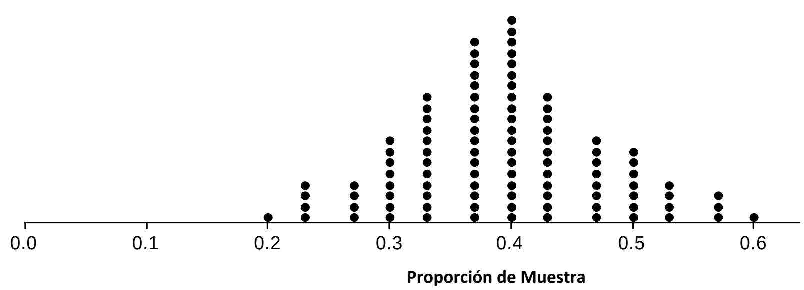
Ejercicios 7–12: Uso de la Desviación Estándar con Margen de Error

7. En el trabajo anterior, investigaste una distribución muestral simulada de la proporción de mujeres en una muestra del tamaño de 30, extraídas de una población con una proporción conocida de 0.4 mujeres. La distribución simulada de la proporción de fichas rojas en una muestra del tamaño de 30, extraídas de una población con una proporción conocida de 0.4 se muestra a continuación.

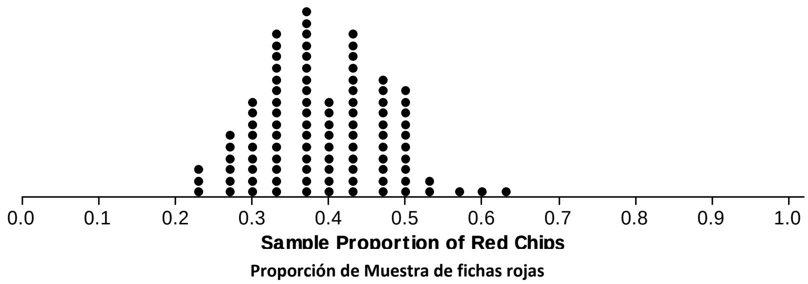


   a. Utiliza la fórmula para la desviación estándar de la proporción de muestra para calcular la desviación estándar de la distribución muestral. Redondea tus respuestas a tres cifras decimales.

   b. La distribución del Ejercicio 2 para una muestra del tamaño de 30se encuentra a continuación. ¿Cómo se comparan las dos distribuciones?



c. ¿Cuántos (... ʼiación estándar de 0.4? ¿Cuá



En general, para una proporción de población conocida, acerca del 95% de los resultados de una distribución muestral simulada de una proporción de muestra estará dentro de dos desviaciones estándar de la proporción poblacional. Una advertencia es que si la proporción es cercana a 1 o a 0, esta regla general no se puede sostener a menos que el tamaño de la muestra sea muy grande. A partir de esto puedes construirla para estimar una proporción de éxitos para una proporción de población desconocida y calcular un margen de error sin tener que llevar a cabo una simulación.

Si la muestra es lo suficientemente grande para tener por lo menos 10 de cada uno de los dos posibles resultados de la muestra, pero lo suficientemente pequeña para no ser más del 10% de la población, la siguiente fórmula (basada en una proporción de muestra observada $\hat{p}$) se puede utilizar para calcular el margen de error. La desviación estándar involucra al parámetro p que comienza ha estimarse. Debido a que p es a menudo desconocido, los estadísticos reemplazan p con su estimación $\hat{p}$ en la fórmula de desviación estándar. A esta desviación estándar estimada se le llama error estándar de la proporción de muestra.

8.

a. Supongamos que sacas una muestra aleatoria de 36 fichas de la bolsa misteriosa y encuentras 20 fichas rojas. Encuentra $\hat{p}$, la proporción de muestra de fichas rojas, y el error estándar.

b. Interpreta el error estándar.

Cuando la estimación de una proporción de la población, el **margen de error** puede ser definido como la **diferencia máxima esperada** entre el valor de la proporción de población y una estimación de la muestra de esa proporción (la más alejada del valor real de la población que crees que su estimación probablemente sea) .

Si $\hat{p}$ es la proporción muestral de una muestra aleatoria del tamaño de n de alguna población, y si el tamaño de la muestra es lo suficientemente grande:

$$\textbf{estimado margen de error} = 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

9. Henri y Terence tomaron muestras del tamaño de 50 de una bolsa misteriosa. Henri sacó 42 fichas rojas, y Terence sacó 40fichas rojas. Encuentra los márgenes de error para cada estudiante.

10. Divida los problemas a continuación de tu grupo, y encuentra la proporción muestral de éxitos y el margen de error estimado en cada situación:
    a. Muestra del tamaño de 20, 5fichas rojas

    b. Muestra del tamaño de 40, 10fichas rojas

    c. Muestra del tamaño de 80, 20 fichas rojas

    d. Muestra del tamaño de 100, 25fichas rojas

11. Mira tus respuestas del Ejercicio 2.
    a. ¿Qué conjeturas puedes hacer acerca de la relación entre el tamaño de la muestra y el margen de error? Explica por qué tu conjetura tiene sentido.

    b. Piensa en la fórmula para el margen de error. ¿Cómo apoya esto o refutar tu conjetura?

12. Supongamos que una muestra aleatoria del tamaño de 100 se utilizará para estimar una proporción de la población.
    a. ¿El margen estimado de error será mayor si $\hat{p} = 0.4$ o $\hat{p} = 0.5$? Apoya tu respuesta con los cálculos apropiados.

    b. ¿El margen estimado de error será mayor si $\hat{p} = 0.5$ o $\hat{p} = 0.8$? Apoya tu respuesta con los cálculos apropiados.

    c. ¿Para qué valor de $\boldsymbol{\hat{p}}$ piensas que el margen estimado de error será mayor? (Sugerencia: Realiza la gráfica de $\boldsymbol{\hat{p}(1 - \hat{p})}$ en la que $\boldsymbol{\hat{p}}$ va de 0 a 1.)

Resumen de la Lección

- Debido a que las muestras aleatorias se comportan de una manera consistente, un tamaño de muestra suficientemente grande permite encontrar una fórmula para la desviación estándar de la distribución muestral de una proporción de la muestra. Esto se puede utilizar para calcular el margen de error: $M = 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$, donde $\hat{p}$ es la proporción de éxitos en una muestra aleatoria del tamaño de n .
- El tamaño de la muestra es lo suficientemente grande para utilizar este resultado para un margen estimado de error si hay por lo menos 10 de cada uno de los dos resultados.
- El tamaño de la muestra no debe exceder de 10% de la población.
- Cuanto menor sea el tamaño de la muestra, menor será el margen de error.

Conjunto de Problemas

1. Diferentes estudiantes tomaron muestras aleatorias del tamaño de 50 de la bolsa misteriosa. A continuación se presenta el número de fichas rojas cada uno arrojó. En cada caso, encuentra el margen de error para las proporciones de las fichas rojas en la bolsa misteriosa.
   a. 10 fichas rojas
   b. 28 fichas rojas
   c. 40 fichas rojas

2. El periódico de la escuela en una secundaria grande informó que de 120 de 200 estudiantes seleccionados prefieren los lugares de estacionamiento asignados. Calcula el margen de error. Interpreta el intervalo que resulta en el contexto.

3. Un periódico en una gran ciudad pregunto a 500 mujeres lo siguiente: "¿Utiliza comida orgánica (como leche, carnes, verduras, etc.)?" 280 mujeres respondieron "sí". Calcula el margen de error. Interpreta el intervalo que resulta en el contexto.

4. Los resultados de las pruebas de un nuevo medicamento en 1000 personas con una cierta enfermedad encontraron que 510 de ellos mejoraron cuando utilizan el medicamento. Supongamos que 1000 de estas personas pueden ser considerados como una muestra aleatoria de la población de todas las personas con esta enfermedad. Basándote en estos resultados, ¿sería razonable pensar que más de la mitad de las personas con esta enfermedad mejoraría si usarán el nuevo medicamento? ¿Por qué si o por qué no?

5. Un periódico de Nueva York tomó una muestra aleatoria de 500 votantes registrados de la Ciudad de Nueva York y encontró que 300favorecían un determinado candidato a gobernador del estado. Un segundo periódico encuestó a 1000 votantes registrados en el estado de Nueva York y se encontró que 550 personas favorecían a este candidato. Explica cómo interpretarías los resultados.

6. En una muestra aleatoria de $1{,}500$ estudiantes en una gran escuela suburbana, $1{,}125$ informaron tener una mascota, resultando en el intervalo 0.75 ± 0.022. Mientras que en una gran escuela urbana, 840 de $1{,}200$ estudiantes informaron tener una mascota, resultando en el intervalo 0.7 ± 0.026. Debido a que estos dos intervalos no se superponen, no parece haber una diferencia en la proporción de estudiantes suburbanos que tienen una mascota y la proporción de estudiantes urbanos que tienen una mascota. Supongamos que el tamaño de la muestra de la escuela suburbana era sólo de 500 pero aún así el 75% reportó tener una mascota. Supongamos que el tamaño de la muestra de la escuela urbana era sólo de 600 y aún así el 70% reportó tener una mascota. ¿Hay todavía una diferencia en la proporción de estudiantes que tienen una mascota en las escuelas suburbanas y en las escuelas urbanas? ¿Por qué ocurre esto?

7. Encuentra un artículo en los medios de comunicación que utiliza un margen de error. Describe la situación (un experimento, un estudio observacional), e interpretar el margen de error para el contexto.