El Teorema del límite central

El teorema del límite central plantea que para una población con cualquier distribución, la distribución de la medias muestrales se aproxima a una distribución normal conforme aumenta el tamaño de la muestra. En otras palabras, si el tamaño de la muestra es lo suficientemente grande, la distribución de las medias muestrales puede aproximarse por medio de una distribución normal, incluso si la población original no está distribuida normalmente. Además, si la población original tiene media μ y desviación estándar σ, entonces la media de las medias muestrales también será μ, pero la desviación estándar de las medias muestrales será $\frac{\sigma}{\sqrt{n}}$, donde n es el tamaño de la muestra.

Cuando se selecciona una muestra aleatoria simple de n sujetos a partir de una población con media μ y desviación estándar σ, es esencial conocer los siguientes principios.

1. Para una población con cualquier distribución, si n>30, entonces las medias muestrales tienen una distribución que se puede aproximar por medio de una distribución normal, con media μ y desviación estándar $\frac{\sigma}{\sqrt{n}}$.

2. Si n≤30 y la población original tiene una distribución normal, entonces las medias muestrales tienen distribución normal con media μ y desviación estándar $\sigma/\sqrt{}$ n..

El teorema del límite central implica dos distribuciones diferentes: la distribución de la población original y la distribución de las medias muestrales. Ahora se necesita una nueva notación para la media y la desviación estándar de las medias muestrales.

Notación para la distribución muestral de $\overline{x}$

Si se seleccionan todas las muestras aleatorias posibles de tamaño n a partir de una población con media μ y desviación estándar σ, la media de las medias muestrales se denota con $\mu_{\bar{x}}$, de manera que:

$$\mu_{\bar{x}} = \mu$$

Asimismo, la desviación estándar de las medias muestrales se denota con $\sigma_{\bar{x}}$, de manera que:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$\sigma_{\bar{x}}$ Suele denominarse el *error estándar de la media.*

Aplicación del teorema del límite central

Muchos problemas prácticos importantes se resuelven mediante el teorema del límite central. Cuando trabaje con este tipo de problemas, recuerde que si el tamaño de la muestra es mayor que 30, o si la población original se distribuye normalmente, debe tratar la distribución de medias muestrales como si fuera una distribución normal con media μ y desviación estándar $\frac{\sigma}{\sqrt{n}}$.

Valor individual: Cuando trabaje con un valor individual de una población normalmente distribuida, utilice: $z = \frac{x-\mu}{\sigma}$.

Muestra de valores: Cuando trabaje con una media de alguna muestra (o grupo), asegúrese de utilizar el valor de $\frac{\sigma}{\sqrt{n}}$ para la desviación estándar de las medias muestrales. Utilice: $z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$.

Ejemplo:

Suponga que un taxi acuático se hundió en el Inner Harbor de Baltimore y murieron varios pasajeros. Los hombres suelen ser más pesados que las mujeres y los niños, por lo tanto supongamos que al cargar un taxi acuático, la situación extrema es aquella en la que todos los pasajeros son hombres.

El Teorema del límite central

En concordancia con los datos de la National Health and Nutrition, la media es de 172 libras y una desviación estándar de 29 libras. Es decir, suponga que la población de pesos de hombres se distribuye de manera normal con media $\mu = 172$ libras y $\sigma = 29$ libras.

a) Calcule la probabilidad de que, si se selecciona a un solo hombre al azar, su peso sea mayor que 175 libras.

b) Calcule la probabilidad de que 20 hombres elegidos al azar tengan un peso medio mayor que 175 libras (de manera que su peso total exceda la capacidad segura de 3500 libras).

a) Método del valor individual, ya que se está trabajando con un solo valor de una población distribuida normalmente. Si se utiliza una tabla de la distribución normal estándar, hay que convertir el peso de 175 a su puntuación z correspondiente:

$$z = \frac{x - \mu}{\sigma} = \frac{175 - 172}{29} = 0.10$$

Nos remitimos a la tabla de la distribución normal estándar y utilizamos z = 0.10 para encontrar el área acumulada a la izquierda de 175 libras es 0.5398. Por tanto, la región sombreada es 1 – 0.5398 = 0.4602. La probabilidad de que un hombre elegido al azar pese más de 175 libras es de 0.4602. (Si se utiliza una calculadora o un programa de cómputo en vez de la tabla de la distribución normal estándar, el resultado más exacto es 0.4588 en lugar de 0.4602).

b) Utilizando el teorema del límite central (porque estamos trabajando con la media de una muestra de 20 hombres y no con un solo hombre). Aun cuando el tamaño de la muestra no es mayor que 30, utilizamos una distribución normal, de manera que las muestras de cualquier tamaño producirán medias distribuidas normalmente; como estamos trabajando con una distribución de medias muestrales, debemos utilizar los parámetros $\mu_{\bar{x}}$ y $\sigma_{\bar{x}}$, que se evalúan de la siguiente manera:

$$\mu_{\bar{x}} = \mu = 172$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{29}{\sqrt{20}} = 6.4845971$$

Queremos calcular el área sombreada que se presenta en la figura.

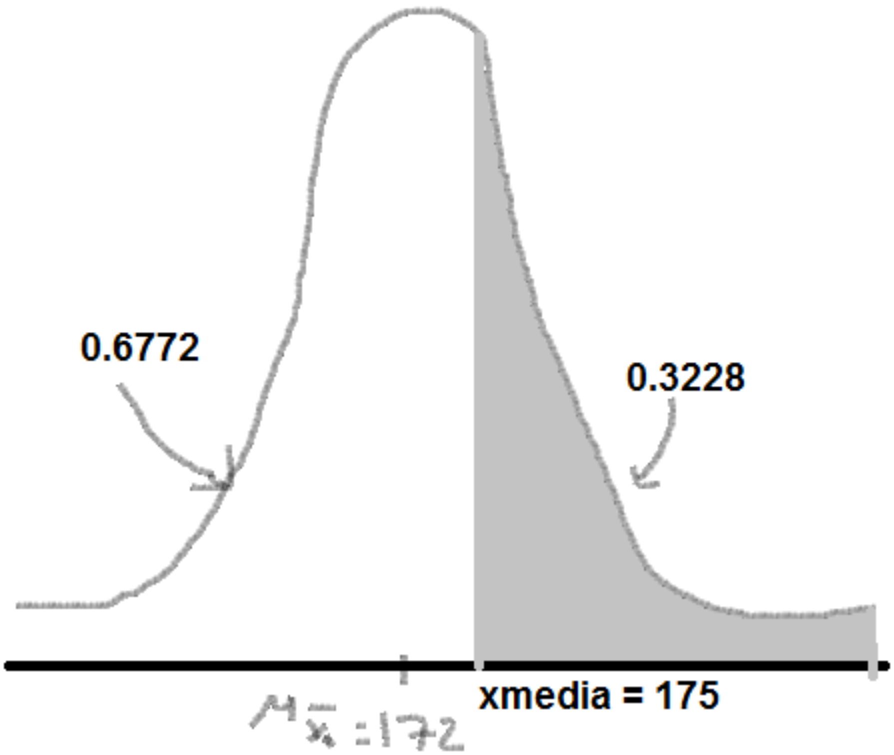

Imagen elaborada por Elsa Rivera (2015), Peso de hombres.

Observe que la distribución de la figura es más estrecha porque la desviación estándar es menor. Si utilizamos la tabla de la distribución normal estándar, encontramos la puntuación z relevante, la cual se calcula de la siguiente manera:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{175 - 172}{\frac{29}{\sqrt{20}}} = \frac{3}{6.4845971} = 0.46$$

Si nos remitimos a la tabla de la distribución normal estándar, encontramos que z = 0.46 corresponde a un área izquierda acumulada de 0.6772, de manera que la región sombreada es 1-0.6772=0.3228. La probabilidad de que 20 hombres tengan un peso medio mayor que 175 libras es de 0.3228 (si se utiliza una calculadora o un programa de cómputo, el resultado es 0.3218 en vez de 0.3228).

Interpretación: existe una probabilidad de 0.4602 de que un hombre pese más de 175 libras y una probabilidad de 0.3228 de que 20 hombres tengan un peso medio mayor que 175 libras. Como la capacidad segura del taxi acuático es de 3500 libras, es muy probable (con una probabilidad de 0.3228) que se sobrecargue si se transporta a 20 hombres elegidos al azar. Puesto que ya han muerto 21 personas, y dada la alta probabilidad de sobrecarga, lo más pertinente sería limitar el número de pasajeros a menos de 20. La capacidad de 20 pasajeros no es suficientemente segura.

Referencia:
Triola, M., 2013, Estadística, Pearson Educación.