Distribución Muestral de Medias

Considera que se obtienen todas las **muestras posibles sin reemplazamiento**, de tamaño *N,* de una población finita de tamaño $N_p > N$. Si se simbolizan la media y la desviación estándar de la distribución muestral de medias por μ_x y σ_x y la media y desviación estándar de la población por μ y σ, respectivamente, entonces

$$\mu_x = \mu \quad \text{y} \quad \sigma_x = \frac{\sigma}{\sqrt{N}} \sqrt{\frac{N_p - N}{N_p - 1}}$$

si la población es infinita o **el muestreo se hace con reemplazamiento**, el resultado anterior se deduce a

$$\mu_x = \mu \quad \text{y} \quad \sigma_x = \frac{\sigma}{\sqrt{N}}$$

Para valores grandes de N ($N \geq 30$), la distribución muestral de medias es en forma aproximada a una distribución normal, con media μ_x y desviación estándar σ_x, independientemente de la población (siempre y cuando la media y la varianza de la población sean finitas y el tamaño sea al menos el doble que el de la muestra). Este resultado para una población infinita es un caso especial del *teorema del límite central* de la teoría avanzada de probabilidad, el cual demuestra que la precisión de la aproximación se incrementa conforme *N* se hace más grande. En ocasiones esto se indica diciendo que la distribución muestral es *asintóticamente normal.*

En caso de que la población esté normalmente distribuida, la distribución muestral de medias también lo está, aún con valores pequeños de *N* (es decir, *N<30*).

Ejemplo:

Una población consiste en cinco números 2, 3, 6, 8 y 11. Considere todas las muestras de tamaño igual a 2 que pueden obtenerse, con reemplazamiento, a partir de esta población. Calcule *a)* la media de la población, *b)* la desviación estándar de la población y *c)* la media de la distribución muestral de medias.

Solución

a) $\mu = \frac{2+3+6+8+11}{5} = \frac{30}{5} = 6.0$

b) $\sigma^2 = \frac{(2-6)^2(3-6)^2(6-6)^2(8-6)^2(11-6)^2}{5} = \frac{16+9+0+4+25}{5} = 10.8$

y σ= 3.29

c) Existen 5 muestras de tamaño igual a 2 que pueden obtenerse con reemplazamiento (ya que cualquiera de los 5 números de la primera muestra llegan a asociarse con cualquiera de los 5 números de la segunda muestra).
Estos son:

(2,2)	(2,3)	(2,6)	(2,8)	(2,11)
(3,2)	(3,3)	(3,6)	(3,8)	(3,11)
(6,2)	(6,3)	(6,6)	(6,8)	(6,11)
(8,2)	(8,3)	(8,6)	(8,8)	(8,11)
(11,2)	(11,3)	(11,6)	(11,8)	(11,11)

Las medias muestrales correspondientes son:

2.0	2.5	4.0	5.0	6.5
2.5	3.0	4.5	5.5	7.0
4.0	4.5	6.0	7.0	8.5
5.0	5.5	7.0	8.0	9.5
6.5	7.0	8.5	9.5	11.0

y la media de la distribución muestral de medias es

$$\mu_{\times} = \frac{suma\ de\ todas\ las\ medias\ muestrales}{25} = \frac{150}{25} = 6.0$$

Ilustrando el hecho de que $\mu_x = \mu$

Ejemplo:

Supongamos que la estatura de 3,000 estudiantes universitarios hombres se distribuye normalmente, con una media de 68.0 pulgadas y una desviación estándar de 3.0 pulgadas. Si se obtienen 80 muestras de 25 estudiantes cada una. ¿Cuáles serían la media y la desviación estándar esperadas de la distribución muestral de medias resultante si los muestreos se hubieran hecho con a) reemplazamiento y b) sin reemplazamiento?

Solución:

El número de muestras de tamaño de 25 que podrían obtenerse teóricamente de un grupo de 3,000 estudiantes con y sin reemplazamiento es de $(3\ 000)^{25}$ y $\frac{3\ 000}{25}$, que son mucho mayores que 80. Por lo tanto, no se obtiene una verdadera distribución muestral de medias, sino solo una distribución muestral experimental. Sin embargo, ya que el número de muestras es grande, debe haber grande concordancia entre las dos distribuciones muestrales. Entonces, la media y la desviación esperadas estarían cerca de las de la distribución teórica. Por lo tanto:

a) $\mu_x = \mu = 68.0\ pulgadas$ y $\sigma_x = \frac{\sigma}{\sqrt{N}} = \frac{3}{\sqrt{25}} = 0.6\ pulgadas$

b) $\mu_x = 68.0\ pulgadas$ y $\sigma_x = \frac{\sigma}{\sqrt{N}} \sqrt{\frac{N_p - N}{N_p - 1}} = \frac{3}{\sqrt{25}} \sqrt{\frac{3\ 000 - 25}{3\ 000 - 1}}$

que solo es ligeramente menor que 0.6 pulgadas y puede; por lo tanto, para propósitos prácticos, considerarse igual que en el muestreo con reemplazamiento.

Distribución Muestral de Medias

Por lo tanto, se esperaría que la distribución muestral experimental de medias esté distribuida aproximadamente de manera normal, con media 68.0 pulgadas y desviación estándar 0.6 pulgadas.

Referencia:

Spiegel M. y Stephens L. (2002). Estadística. México: McGraw-Hill