

Distribución Muestral

Consideremos todas las posibles muestras de tamaño n en una población. Para cada muestra

podemos calcular un estadístico muestral la media muestral (\bar{x}), la desviación estándar

muestral s , la proporción muestral (\bar{p}), entre otros. Que variará de una a otra muestra. Así obtenemos una distribución de los estadísticos resultantes de cada muestra que se llama distribución muestral, por lo tanto se puede considerar que la distribución de probabilidad de todas las muestras de un determinado tamaño de muestra de la población.

Las muestras aleatorias obtenidas de una población son, por naturaleza propia, **impredecibles**. No se esperaría que dos muestras aleatorias del mismo tamaño y tomadas de la misma población tenga la misma media muestral o que sean completamente parecidas; puede esperarse que cualquier estadístico, como la media muestral, calculado a partir de las medias en una muestra aleatoria, cambie su valor de una muestra a otra, por ello, se quiere estudiar la distribución de todos los valores posibles de un estadístico. Tales distribuciones serán muy importantes en el estudio de la estadística inferencial, porque las inferencias sobre las poblaciones se harán usando estadísticos muestrales. Conocer esta distribución muestral y sus propiedades permitirá juzgar la confiabilidad de un estadístico muestral como un instrumento para hacer inferencias sobre un parámetro poblacional desconocido.

Hay dos tipos de distribuciones muestrales:

- La distribución muestral de medias.
- Distribución muestral de proporciones.

Pero en este trabajo haremos referencia en específica a la distribución muestral de medias.

Distribución muestral de la media

Es la distribución de probabilidad de todos los valores de la media muestral (\bar{x}). La distribución muestral de medias como otras distribuciones de probabilidad tiene un valor esperado, una desviación estándar y una forma característica.

Valor esperado:

Cuando se tienen las medias de varias muestras sacadas de una población, la media de todos esos valores, se le conocerá como **valor esperado** de la media muestral. Por lo tanto:

Valor esperado de la media muestral

$$E(\bar{x}) = \text{valor esperado de la media muestral } (\bar{x})$$

μ = media poblacional

$$E(\bar{x}) = \mu$$

De una forma más explícita, El valor esperado de la media muestral $E(\bar{x})$, es igual a la media de la población (μ) de la que se tomó la muestra. Cuando el valor esperado de un estimador puntual es igual al parámetro poblacional, se dice que el estimador puntual es insesgado.

Desviación estándar de la media muestral.

Es posible demostrar que usando el muestreo aleatorio simple, la desviación estándar de la media muestral depende de si la población es finita o infinita. Las dos fórmulas para la desviación estándar son las siguientes:

$$\sigma_{\bar{x}} = \text{desviación estándar de la media muestral.}$$

σ = desviación estándar poblacional.

n = tamaño de la muestra.

N = tamaño de la población.

Población Finita

$$\sigma_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \left(\frac{\sigma}{\sqrt{n}} \right)$$

Población infinita

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

Al observar las dos fórmulas se puede notar que en la de población finita tiene un factor extra y en la finita no, a este factor se le conoce como factor de corrección para una población finita. Y esto se ve sí y solo sí al momento de efectuar un muestro a una población finita esta es grande, mientras que si el tamaño de la muestra es pequeño, en estos casos el factor de corrección para una población finita es igual a 1. Por lo tanto, la

diferencia entre el valor de la desviación estándar de la media muestral en el caso de la

$$\sigma_x = \frac{\sigma}{\sqrt{n}}$$

población finita o infinita se vuelve despreciable. Entonces es una buena aproximación a la desviación estándar de la media muestral, aun cuando la población sea finita. Esta observación lleva a la siguiente regla general, para calcular la desviación estándar de la media muestral:

Uso de la expresión siguiente para calcular así la desviación estándar de la media muestral.

$$\sigma_x = \frac{\sigma}{\sqrt{n}}$$

Siempre que:

1. La población sea infinita.
2. La población sea finita y el tamaño de la muestra sea menor o igual al 5% del

tamaño total de la población es decir que $n/N = \leq 0.05$.

En los casos en que $n/N < 0.05$, para calcular σ_x deberá usarse la versión para poblaciones finitas.

Para calcular σ_x se necesita conocer (La desviación estándar de la población) y para

diferenciar σ_x de σ , a la desviación estándar de la media muestral se le llama (error estándar de la media). Este término se refiere a la desviación estándar de un estimador puntual, a través de este valor se puede determinar qué tan lejos puede estar la media muestral de la media poblacional.

Forma de la distribución muestral

El paso final en la identificación de las características de la distribución muestral de la media es determinar la forma de la distribución muestral. Y se consideran dos casos:

1. La población tiene distribución normal. Cuando la población tiene distribución normal, la distribución muestral de la media está distribuida normalmente sea cual sea el tamaño de la muestra.

2. La población no tiene distribución normal. Y cuando no tienen distribución normal el teorema del límite central ayuda a determinar la forma de la distribución de la media, el teorema expone lo siguiente:

Teorema del Límite central

Cuando se seleccionan muestras aleatorias simples de tamaño n de una población, la

distribución muestral de la media muestral \bar{x} puede aproximarse mediante una distribución normal a medida que el tamaño de la muestra se hace grande.

Ejercicios:

- 1) La media de una población es de 200 y su desviación estándar es de 50, se tomara una muestra aleatoria simple de tamaño 100 y se usara la media muestral para estimar la media poblacional. Determine:
- ¿Cuál es el valor esperado de la media muestral?
 - ¿Cuál es la desviación estándar de la media muestral?
 - Muestre la distribución muestral de la media.

Respuesta:

a) $E(\bar{x}) = 200$

- b) Como el tamaño de la muestra representa más del 5% del tamaño de la población se utilizara la formula con el factor de corrección.

$$\sigma_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \left(\frac{\sigma}{\sqrt{n}} \right)$$

$$\sigma_{\bar{x}} = \sqrt{\frac{200-100}{200-1}} \left(\frac{50}{\sqrt{100}} \right) = 3.5444$$

- c) Esta tiene una distribución normal con un valor esperado de 200 y una desviación estándar muestral de la media de 3.5444.

- 2) El valor (en miles) de las ventas mensuales realizadas en una empresa manufacturera presenta una distribución normal la cual tiene una media de 300 y una desviación estándar de 40, de la cual se tomara una muestra aleatoria simple de

tamaño 10 y se usara la media muestral para estimar la media poblacional.

Determine:

- a) ¿Cuál es el valor esperado de la media muestral?
- b) ¿Cuál es la desviación estándar de la media muestral?
- c) Muestre la distribución muestral de la media.

Respuesta:

$$\bar{x}$$

a) $E(\bar{x}) = 300$

- b) Como el tamaño de la muestra representa menos del 5% del tamaño de la población se utilizara la formula sin el factor de corrección:

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

$$\sigma_{\bar{x}} = \frac{40}{\sqrt{10}} = 12.6491$$

- c) Esta tiene una distribución normal con un valor esperado de 300 y una desviación estándar muestral de la media de 12.6491.

- 3) El peso de un producto en kg sigue una distribución normal con una media de 30 y desviación típica de 3, un empresario decide aceptar un lote de 600 unidades que le envía el proveedor. Si al elegir 5 unidades de dicho producto al azar encuentra que su peso medio no es menor de 29. Calcular la probabilidad de que rechace el lote.

Respuesta:

Datos:

$$X \rightarrow N(30,3)$$

$$n = 5$$

$$\bar{X} \rightarrow N(\mu_{\bar{x}} = \mu = 30)$$

$$\sigma_{\bar{x}} = \frac{3}{\sqrt{5}}$$

$$P(\bar{X} < 29) = P(Z < Z_{\bar{x}}) = P(Z < -0.7453) = 0.2734$$

$$Z_{\bar{x}} = \frac{29 - 30}{\frac{3}{\sqrt{5}}} = -0.7453$$

Existe una probabilidad de 27.34% de que rechace el producto.

4) Las estaturas de 1000 estudiantes están distribuidas aproximadamente en forma normal con una media de 174.5 cm y una desviación estándar de 6.9cm, si se extraen 200 muestras aleatorias de tamaño 25 determinar:

- a) El número de medias muestrales que caen entre 172.5 y 175.8 cm.
- b) El número de medias muestrales que caen por debajo de 172cm.

Respuesta:

Como se puede observar en este ejercicio se cuenta con una población finita y un muestreo que sobrepasa el 5% de la población general por lo tanto a la formula se le debe agregar el factor de corrección.

Datos:

$$X \rightarrow N(174.5; 6.9)$$

$$N = 1000$$

$$n = 25$$

$$\mu = 174.5$$

$$\sigma = 6.9$$

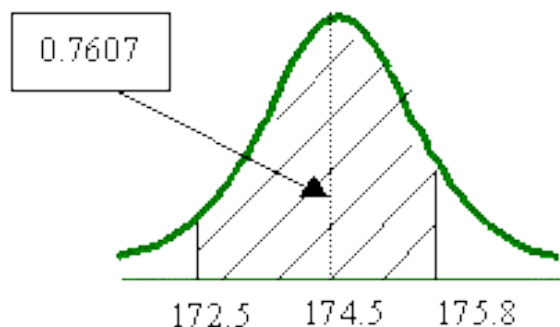
$$\sigma_{\bar{X}} = \sqrt{\frac{1000 - 25}{1000 - 1} \left(\frac{6.9}{\sqrt{25}} \right)} = 1.36$$

$$P(172.5 \leq \bar{X} \leq 175.8) = P(-1.4705 \leq Z_{\bar{X}} \leq 0.96) = 0.4292 + 0.3289 = 0.7607$$

a)

$$Z_{\bar{X}_1} = \frac{172.5 - 174.5}{1.36} = -1.47 = 0.4292$$

$$Z_{\bar{X}_2} = \frac{175.8 - 174.5}{1.36} = 0.96 = 0.3315$$



200 → 100%

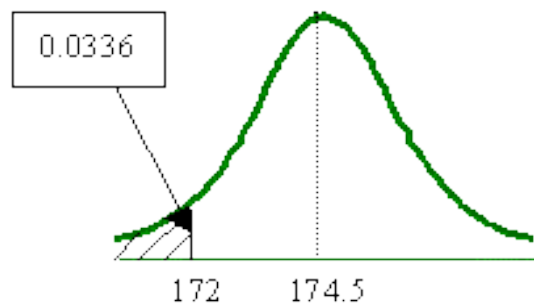
$X \rightarrow 76.07\%$

= 152 medias muestrales.

$$P(\bar{X} \leq 172) = P(Z_{\bar{X}} \leq -1.83) = 0.5 - 0.4664 = 0.0336$$

b)

$$Z_{\bar{X}} = \frac{172 - 174.5}{1.36} = -1.83 = 0.4664$$



200 → 100%

$X \rightarrow 3.36\%$

= 7 medidas muestrales.

Referencias bibliográficas

Anderson, D. R., D. J. Sweeney y T. A. Williams. (2008). *Estadística para la administración y la economía*. (10ª ed). México: CENGAGE Learning. 267- 283

Levine, D. M., T. C. Krehbiel y M. L. Berenson. (2006). *Estadística para la administración*. (4ª ed). México: Pearson Prentice Hall. 206

http://es.slideshare.net/chcluz/distribucin-muestral-3?from_action=save

