MA2- Clase 3- 12/09/2023

Distribución por Muestreo:

Antes de indicar qué es una distribución por muestreo, distinguiremos dos conceptos:

Parámetro (Parámetro Poblacional):

Es toda medida obtenida en base a datos provenientes de una población.

Ejemplos:

Parámetro	Simbología	Fórmula
Media Poblacional	**μ**	$\frac{\sum xi}{N}$
Varianza Poblacional	**σ^2**	$\frac{\sum x_i^2}{N} - \mu^2$
Desviación Poblacional	**σ**	$\sqrt{\frac{\sum x_i^2}{N} - \mu^2}$
Proporción Poblacional	**P**	$\frac{X}{N}$

Estimador (Medida muestral):

Es toda medida obtenida en base a datos provenientes de una muestra.

Ejemplos:

Estimador (Medida muestral)	Simbología
Media Muestral	$\bar{x}$
Varianza Muestral Corregida	$\hat{s}^2$
Desviación Muestral Corregida	$\hat{s}$
Proporción Muestral	$\hat{p}$

Cuando no es posible calcular el valor de un parámetro, entonces se recurre a una muestra y en base a dicha muestra debe calcularse el valor del estimador, el cual brinda una idea respecto al verdadero valor del parámetro que se hubiera obtenido en base a la población, ese valor particular del estimador obtenido en base a una muestra se denomina **Estimación Puntual.**

Estimador es la Medida. Ejemplo: Media Muestral

Estimación Puntual: Valor del Estimador obtenido a partir de la muestra seleccionada.

Ejemplo:

Realizar una Estimación Puntual de la media equivale a calcular la media en la muestra seleccionada.

Realizar una Estimación Puntual de la varianza equivale a calcular la varianza en la muestra seleccionada.

Generalmente el valor del estimador no coincide con el valor del parámetro, existe una diferencia que se denomina **Error de Muestreo.**

Error de Muestreo o Error de Estimación:

Es la diferencia entre la **Estimación Puntual** (Valor del estimador) y el **Parámetro poblacional (Verdadero valor del parámetro)**

El valor del estimador varía según de que muestra provenga, por lo tanto es una variable aleatoria y en consecuencia es posible obtener su distribución de probabilidades, la cual se denomina **Distribución por Muestreo.**

Distribución por Muestreo:

¿Qué es?

Es la distribución de probabilidades de un Estimador.

¿Qué indica?

Indica los valores que puede asumir el estimador (por ejemplo la media muestral) y sus respectivas probabilidades.

Cada posible valor del estimador tiene una determinada probabilidad de presentarse.

Distribución por Muestreo de la Media Muestral

Una población está compuesta por cuatro empleados que trabajan 3, 4, 5 y 6 (N=4 tamaño de la población) horas diarias. Sea x la cantidad de horas diarias que trabaja cada empleado, se pide:

a) Calcular el valor de los siguientes parámetros (valores poblacionales): **Media poblacional, Varianza poblacional y Desviación poblacional.**
b) Encontrar la **Distribución por muestreo del estimador** correspondiente al parámetro **Media poblacional** en base a todas las muestras de magnitud 2 que surjan por enumeración de todos los **arreglos** posibles (muestreo **con reposición**)
Parámetros de la Distribución por muestreo:
c) Calcular la **Esperanza Matemática del estimador**
d) Calcular la **Varianza del estimador**
e) Calcular la **Desviación Standar del estimador**
f) Comprobar la relación **Parámetros poblacionales** y los **Parámetros de la Distribución por muestreo**

Respuesta:

a) **Parámetros poblacionales**: Medidas que son calculadas a partir de una población.

Media poblacional: μ

$$\mu = \frac{\sum xi}{N} = \frac{3+4+5+6}{4}$$

μ = 4,5 En el caso planteado es la cantidad de horas promedio que trabajan los 4 empleados que forman la población, es decir en promedio cada empleado trabaja 4,5 horas diarias.

Varianza poblacional: σ²

$$\sigma^2 = \frac{\sum x_i^2}{N} - \mu^2 = \frac{3^2+4^2+5^2+6^2}{4} - 4,5^2 = \frac{86}{4} - 20,25$$

= 21,50 – 20,25

σ² = 1,25

Desviación poblacional: σ

Es la Raíz cuadrada la Varianza Poblacional

$$\sigma = \sqrt{\frac{\sum x_i^2}{N} - \mu^2 = \frac{3^2+4^2+5^2+6}{4} - 4,5^2} = \sqrt{1,25}$$

σ =1,118

Cuando no se dispone de toda la información referida a la población entonces no es posible calcular los Parámetros poblacionales.

Por ejemplo para calcular la media poblacional, es necesario conocer todos los valores de la variable que se presentaron en la población, si dicha información entonces no se encuentra disponible entonces no será posible calcular el verdadero valor de la media.

¿De qué manera se procede ante esta situación?

En tal caso se extrae una muestra y en partir de dicha muestra se obtiene la media, ese valor hallado brinda una idea del verdadero valor que se hubiera obtenido en la población.

Esa medida que se obtiene a partir de la muestra se denomina Estimador, por ejemplo la **Media Muestral ($\bar{x}$)** es un estimador de la **Media poblacional**.

El valor particular que asume un Estimador en la muestra seleccionada se denomina Estimación Puntual.

Por ejemplo, realizar una Estimación Puntual de la Media consiste en calcular la Media en la muestra seleccionada.

El valor de un Estimador varía según de cual muestra provenga.

Antes de extraer la muestra no se tiene la certeza con respecto a que elementos serán seleccionados, es por ello que tampoco se tiene la certeza con respecto a que valor asumirá el estimador en la muestra, en consecuencia todo estimador es una variable aleatoria.

Si un Estimador es una variable aleatoria entonces tiene una distribución de probabilidades.

La distribución de probabilidades de un Estimador se denomina Distribución por Muestreo e indica los posibles valores del estimador y sus respectivas probabilidades.
Por ejemplo, la **Distribución por Muestreo de la Media Muestral (Distribución por Muestreo de $\overline{x}$)** se plantea de la siguiente manera:

$\overline{x}$ **(Media Muestral)**	$P(\overline{x})$

Media Muestral: Media o promedio que es calculado a partir de una muestra.

Primera columna: Todos los posibles valores de la Media Muestral (se obtienen a partir de todas las muestras que podemos extraer de la población dada)
Segunda columna: Respectivas probabilidades (Probabilidad de que la Media Muestral asuma determinado valor)

Para obtener todos los posibles **valores del estimador** (que en este caso es la **media muestral**) deben extraerse todas las muestras a partir de la población dada.
En la primera columna se indican los posibles valores del Estimador y en la segunda columna se indican las respectivas probabilidades.

Para obtener todos los posibles valores del Estimador es necesario extraer todas las muestras posibles de tamaño "n" a partir de la población bajo estudio.

La cantidad total de muestras de tamaño n que pueden extraerse de una población de tamaño N varía según el tipo de muestreo: MCR (Muestreo Con Reposición) o MSR (Muestreo Sin Reposición)

MCR:
Si cada muestra se extrae seleccionando los elementos con reposición entonces la cantidad total de muestras se obtiene a través de la siguiente fórmula: N^n
Es decir que pueden extraerse N^n muestras de tamaño n a partir de una población de tamaño N seleccionando los elementos **con reposición**.

MSR:

Si cada muestra se extrae seleccionando los elementos **sin reposición** entonces la cantidad total de muestras se obtiene a través de la siguiente fórmula: C_N^n (combinatoria de N elementos tomados de n en n)

Es decir que pueden extraerse C_N^n muestras de tamaño n a partir de una población de tamaño N seleccionando los elementos sin reposición.

A partir de la **Distribución por Muestreo de un estimador** pueden calcularse los siguientes **parámetros:**

Esperanza Matemática
Varianza
Desviación Estándar

Parámetros de la Distribución por Muestreo de la Media Muestral (Distribución por Muestreo de $\bar{x}$):

Esperanza Matemática de la media muestral (Esperanza Matemática de $\bar{x}$):

$$E(\bar{x}) = \sum \bar{x} * P(\bar{x})$$

Varianza de la media muestral (Varianza de $\bar{x}$):

$$\sigma^2(\bar{x}) = \sum \bar{x}^2 P(\bar{x}) - [E(\bar{x})]^2$$

Desviación Estándar de la media muestral (Desviación Estándar de $\bar{x}$):

$$\sigma(\bar{x}) = \sqrt{\sigma^2(\bar{x})}$$

Para obtener la **Distribución por muestreo de un estimador,** previamente deben seguirse los siguientes pasos:

Paso 1- **Calcular la cantidad de muestras posibles**
Paso 2- **Obtener todas las muestras posibles**
Paso 3- **Calcular en cada muestra el valor del estimador**

Paso 1- **Calcular la cantidad de muestras posibles**

En primer lugar, se determina la cantidad total de muestras que pueden extraerse, para ello debe identificarse en qué condiciones son seleccionados los elementos para constituir la muestra, de la siguiente manera:

Si las muestras se obtienen por enumeración de todos los arreglos posibles, significa que los elementos a incluir en cada muestra son seleccionados con reposición, entonces la cantidad total de muestras de tamaño n que pueden obtenerse a partir de una población de tamaño N surge de la siguiente fórmula: N^n

MCR ___ $N^n = 4^2 = 16$

Significado:

Pueden obtenerse 16 muestras de tamaño 2 seleccionando los elementos **con reposición** (cada elemento se devuelve o repone a la población antes de extraer el siguiente) a partir de una población de tamaño 4

Paso 2- **Obtener todas las muestras posibles**

x_1 = cantidad de horas diarias trabajadas por el primer empleado seleccionado para formar la muestra
x_2 = cantidad de horas diarias trabajadas por el segundo empleado seleccionado para formar la muestra

16 muestras de 2 empleados cada una:

x_1 \ x_2	3	4	5	6
3	3 3	3 4	3 5	3 6
4	4 3	4 4	4 5	4 6
5	5 3	5 4	5 5	5 6
6	6 3	6 4	6 5	6 6

Paso 3- **Calcular en cada muestra el valor del estimador**

En este caso el estimador es la **media muestral**

$\bar{x}$= cantidad promedio de horas diarias por cada empleado que forma la muestra

El valor del estimador obtenido en una muestra seleccionada se denomina Estimación Puntual
Hacer una estimación puntual de la media muestral es calcular el valor de la media en la muestra seleccionada

$$\bar{x} = \frac{x_1 + x_2}{2}$$

x_1: cantidad de horas diarias que trabaja el primer empleado incluido en la muestra
x_2: cantidad de horas diarias que trabaja el segundi empleado incluido en la muestra

Media en cada una de las 16 muestras de tamaño de 2

x_1 \ x_2	3	4	5	6
3	3 3	3 4	3 5	3 6
4	4 3	4 4	4 5	4 6
5	5 3	5 4	5 5	5 6
6	6 3	6 4	6 5	6 6

$\bar{x} = \frac{x_1 + x_2}{2}$	3	4	5	6
3	**3**	**3,5**	**4**	**4,5**
4	**3,5**	**4**	**4,5**	**5**
5	**4**	**4,5**	**5**	**5,5**
6	**4,5**	**5**	**5,5**	**6**

Una vez cumplidos estos dos pasos, es posible formar la **Distribución por muestreo de la media muestral ($\bar{x}$)**, para ello:
En la primera columna de la tabla se indica una vez cada uno de los diferentes valores que puede asumir el estimador $\bar{x}$
En la segunda columna se indican las probabilidades asociadas a cada valor, las cuales se obtienen mediante el cociente entre la cantidad de veces que se repite cada valor (que obtenemos contando en la tabla obtenida en el punto b.2) dividido el total de muestras

$P(\bar{x})$ = $\frac{\text{cantidad de veces que se presenta cada valor de } \bar{x}}{\text{cantidad total de muestras} = N^n}$

$\bar{x} = \frac{x_1 + x_2}{2}$	3	4	5	6
3	3	3,5	4	4,5
4	3,5	4	4,5	5
5	4	4,5	5	5,5
6	4,5	5	5,5	6

Distribución por muestreo de la media muestral (Distribución por muestreo de $\bar{x}$): Indica los posibles valores de la media muestral y sus respectivas probabilidades.

$\bar{x}$	$P(\bar{x})$
3	1/16
3,5	2/16
4	3/16
4,5	4/16
5	3/16
5,5	2/16
6	1/16

$$\sum P(\bar{x}) = \frac{16}{16}$$

La suma de todas las probabilidades debe ser igual a 1.

Parámetros de la Distribución de Probabilidades de la Media Muestral ($\bar{x}$)

Esperanza matemática de la Media Muestral (Esperanza matemática de $\bar{x}$)
Varianza de la Media Muestral (Varianza de $\bar{x}$)
Desviación Estándar de la media muestral $\bar{x}$

a)Esperanza Matemática de la Media Muestral (Esperanza matemática de $\bar{x}$)

La fórmula de cálculo es la siguiente:

E($\bar{x}$) = $\sum \bar{x} * P(\bar{x})$

Para calcular la esperanza matemática, se obtiene el producto de cada valor de la variable por su respectiva probabilidad y luego se suman los productos calculados

$\bar{x}$(*media muestral*)	P($\bar{x}$)	$\bar{x}$* P ($\bar{x}$)
3,0	**1/16**	3 * (1/16) = 3/16
3,5	**2/16**	3,5 * (2/16) =7/16
4,0	**3/16**	4 * (3/16) = 12/16
4,5	**4/16**	4,5 * (4/16) = 18/16
5,0	**3/16**	5 * (3/16) =15/16
5,5	**2/16**	5,5 * (2/16) = 11/16
6,0	**1/16**	6 * (1/16) = 6/16
		E($\bar{x}$) = 72/16 = **4,5**

E($\bar{x}$) = ∑ $\bar{x}$* P ($\bar{x}$)
= **4,5**

d) Varianza de la Media Muestral (Varianza de $\bar{x}$)

La fórmula de cálculo es la siguiente:

$\sigma^2(\bar{x}) = \sum \bar{x}^2 P(\bar{x}) - [E(\bar{x})]^2$
= **20,875** – $(4,5)^2$
= **20,875** – 20,25
$\sigma^2(\bar{x})$ = **0,625**

Para calcular la varianza, se obtiene el producto del cuadrado del valor de la variable por su respectiva probabilidad y a continuación se suman los productos obtenidos. A la **sumatoria resultante (20,875)** se le resta el cuadrado de la Esperanza Matemática (20,25)

$\bar{x}$ **(media muestral)**	$P(\bar{x})$	$\bar{x}^2$* P ($\bar{x}$)
3,0	**1/16**	3^2 * (1/16) = 9/16
3,5	**2/16**	$(3,5)^2$ * (2/16) = 24,5/16
4,0	**3/16**	$(4)^2$ * (3/16) = 48/16
4,5	**4/16**	$(4,5)^2$ * (4/16) = 81/16
5,0	**3/16**	$(5)^2$ * (3/16) = 75/16
5,5	**2/16**	$(5,5)^2$ * (2/16) = 60,5/16
6,0	**1/16**	$(6)^2$ * (1/16) = 36/16
		$\sum \bar{x}^2 * P(\bar{x})$ = 334/16 =**20,875**

e) **Desviación Estándar de la media muestral $\bar{x}$**

La fórmula de cálculo es la siguiente:

$\sigma(\bar{x})= \sqrt{\sigma 2(\bar{x})}$

$=\sqrt{20{,}875 - (4{,}5)^2}$

$= \sqrt{0{,}625}$

= 0,79056

$\sigma(\bar{x}) \approx$ 0,7906

f) Para el Muestreo Con Reposición (MCR) se verifican las siguientes relaciones entre los **Parámetros Poblacionales** y los **Parámetros de la Distribución de $\bar{x}$:**

La Esperanza matemática de la Media Muestral es igual a la Media Poblacional:

$E(\bar{x}) = \mu$
4,5 = **4,5**

La Varianza de la Media Muestral es igual al cociente entre la Varianza Poblacional y el tamaño de cada muestra:

$$\sigma^2 (\bar{x}) = \frac{\sigma^2}{n}$$

$$= \frac{1{,}25}{2}$$

0,625 = 0,625

La Desviación de la Media Muestral es igual al cociente entre la Desviación Poblacional y la raíz cuadrada del tamaño de cada muestra:

$$\sigma (\bar{x}) = \frac{\sigma}{\sqrt{n}}$$

$$= \frac{1{,}118}{\sqrt{2}}$$

0,7906 = 0,7906