Distribución por Muestreo:

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

Parámetro:

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:

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

Ejemplos:

Estimador	Simbología
Media Muestral	$\bar{x}$
Varianza Muestral	s^2
Desviación Muestral	s
Proporción Muestral	p

Cuando no es posible calcular el valor de un parámetro, entonces recurrimos a una muestra y en base a dicha muestra calculamos el valor del estimador, el cual nos da una idea respecto al verdadero valor del parámetro que hubiéramos obtenido en base a la población, ese valor particular del estimador obtenido en base a una muestra se denomina **Estimación Puntual**

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

Error de Muestreo:

Es la diferencia entre la estimación y el parámetro.

El valor del estimador varía según de que muestra provenga, por lo tanto es una variable aleatoria y en consecuencia podemos 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 y sus respectivas probabilidades.

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

a) Calcular el valor de los siguientes parámetros (valores poblacionales): Media poblacional, Varianza
poblacional Desviación poblacional.

b) Encontrar la distribución por muestreo del estimador correspondiente al parámetro calculado en el punto a) en base a todas las muestras de magnitud 2 que surjan seleccionando los elementos con reposición)

c) Calcular la Esperanza Matemática del estimador

d) Calcular la Varianza del estimador

e) Calcular la Desviación estándar del estimador

f) Comprobar la relación de la esperanza matemática y de la desviación estándar con los parámetros dela población.

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

Varianza poblacional: σ^2

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

= 21,50 – 20,25

σ^2 = 1,25

Desviación 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 contamos con toda la información referida a la población entonces no es posible calcular los Parámetros poblacionales.

Por ejemplo para calcular la media poblacional necesitamos conocer todos los valores de la variable que se presentaron en la población, si no disponemos de dicha información entonces no será posible calcularla.

¿De qué manera procedemos ante esta situación?

En tal caso extraemos una muestra y en partir de dicha muestra calculamos la media, ese valor hallado nos da una idea del verdadero valor que hubiésemos obtenido en la población.

Esa medida que se obtiene a partir de la muestra se denomina Estimador, por ejemplo la Media Muestral
($\overline{X}$) 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 equivale a calcular la Media en la muestra seleccionada.

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

Antes de extraer la muestra no tenemos la certeza con respecto a que elementos serán seleccionados, por lo tanto tampoco tenemos 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 la media muestral (Distribución por muestreo de $\overline{x}$): se plantea de la siguiente manera:

$\overline{x}$	$P(\overline{x})$

En la primera columna indicamos los posibles valores del Estimador y en la segunda columna indicamos las respectivas probabilidades.

Para obtener todos los posibles valores del Estimador necesitamos extraer todas las muestras posibles de tamaño “n” a partir de la población bajo estudio.

La cantidad total de muestras de tamaño n que podemos extraer 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 podemos extraer 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 podemos extraer C_N^n muestras de tamaño n a partir de una población de tamaño N seleccionando los elementos sin reposición.

b) Distribución por muestreo de $\bar{x}$:

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, calculamos la cantidad de muestras que obtendremos, para ello debemos identificar 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 podemos obtener a partir de una población de tamaño N surge de la siguiente fórmula: N^n

$N^n = 4^2 = 16$

Significado:

Podemos obtener 16 muestras de tamaño 2 seleccionando los elementos con reposición (cada elemento se devuelve ó 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

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** $\overline{X} = \dfrac{x_1 + x_2}{2}$

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

$\overline{X}$	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, estamos en condiciones de formar la **Distribución por muestreo**, para ello:

- En la primera columna de la tabla indicamos una vez cada uno de los diferentes valores que puede asumir el estimador $\overline{x}$
- En la segunda columna indicamos 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

$$P(\overline{x}) = \frac{\text{cantidad de veces que se repite cada valor de } \overline{x}}{\text{total de muestras}}$$

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

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

c) **Esperanza matemática de la Media Muestral (Esperanza matemática de $\overline{x}$)**

La fórmula de cálculo es la siguiente:

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

Para calcular la esperanza matemática, obtenemos el producto de cada valor de la variable por su respectiva probabilidad y luego sumamos los productos obtenidos

$\bar{x}$	$P(\bar{x})$	$\bar{X} * P(\bar{X})$
3	1/16	3 * (1/16)
3,5	2/16	3,5 * (2/16)
4	3/16	4 * (3/16)
4,5	4/16	4,5 * (4/16)
5	3/16	5 * (3/16)
5,5	2/16	5,5 * (2/16)
6	1/16	6 * (1/16)
		E($\bar{x}$) = 72/16 = **4,5**

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

La fórmula de cálculo es la siguiente:

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

Para calcular la varianza, obtenemos el producto del cuadrado del valor de la variable por su respectiva probabilidad y a continuación sumamos los productos obtenidos. A la sumatoria resultante (20,875) le restamos el cuadrado de la esperanza matemática (20,25)

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

e) Desviación Estándar de $\bar{x}$

La fórmula de cálculo es la siguiente:

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

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

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

$\sigma(\bar{x})$ = 0,79056 ≈ **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 $\bar{x}$**

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

2) $\sigma^2(\bar{x}) = \sigma^2 / n$
   = **1,25** / 2
   0,625 = 0,625

3) $\sigma(\bar{x}) = \sigma / \sqrt{n}$
   = **1,118** / $\sqrt{2}$
   0,7906 = 0,7906

1) **La Esperanza matemática de la Media Muestral es igual a la Media Poblacional**
2) **La Varianza de la Media Muestral es igual al cociente entre la Varianza Poblacional y el tamaño de cada muestra.**
3) **La Desviación de la Media Muestral es igual al cociente entre la Desviación Poblacional y la raíz cuadrada del tamaño de cada muestra.**

Resolver el siguiente ejercicio:

Una población está compuesta por tres vendedores de cierto comercio que la última semana realizaron 1, 3 y 5 ventas online. Sea x la cantidad ventas online que realizó cada vendedor, 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 calculado en el punto a) en base a todas las muestras de magnitud 2 que surjan seleccionando los elementos con reposición)
c) Calcular la Esperanza Matemática del estimador
d) Calcular la Varianza del estimador
e) Calcular la Desviación estándar del estimador
f) Comprobar la relación de la esperanza matemática y de la desviación estándar con los parámetros de la población.

Respuesta:

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

Media poblacional: μ

$$\mu = \frac{\sum xi}{N} = \frac{1+3+5}{3}$$

μ = 3

Varianza poblacional: σ^2

$$\sigma^2 = \frac{\sum x_i^2}{N} - \mu^2$$

=

σ^2 =

Desviación poblacional:

$$\sigma = \sqrt{\frac{\sum x_i^2}{N} - \mu^2} =$$

σ =

b) Distribución por muestreo de $\bar{x}$:

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 b.1. **Calcular la cantidad de muestras posibles**

En primer lugar, calculamos la cantidad de muestras que obtendremos, para ello debemos identificar 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 podemos obtener a partir de una población de tamaño N surge de la siguiente fórmula: N^n

$N^n = 3^2 = 9$ muestras

Significado:

Podemos obtener 9 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 3

Paso b.2. **Obtener todas las muestras posibles**

x_1 = cantidad de ventas online realizadas por el primer vendedor seleccionado para formar la muestra
x_2 = cantidad de ventas online realizadas por el segundo vendedor seleccionado para formar la muestra

x_1 \ x_2	1	3	5
1			
3			
5			

Paso b.3.

Calcular en cada muestra el valor del estimador $\bar{X} = \frac{x_1 + x_2}{2}$

$\bar{X}$ = cantidad promedio de ventas online realizadas por cada empleado que forma la muestra

$\bar{X}$	1	3	5
1			
3			
5			

Una vez cumplidos estos dos pasos, estamos en condiciones de formar la **Distribución por muestreo**, para ello:

- En la primera columna de la tabla indicamos una vez cada uno de los diferentes valores que puede asumir el estimador $\bar{x}$
- En la segunda columna indicamos 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.3. dividido el total de muestras

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

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**.**

$\bar{x}$	$P(\bar{x})$

c) **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, obtenemos el producto de cada valor de la variable por su respectiva probabilidad y luego sumamos los productos obtenidos:

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

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

La fórmula de cálculo es la siguiente:

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

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

Para calcular la varianza, obtenemos el producto del cuadrado del valor de la variable por su respectiva probabilidad y a continuación sumamos los productos obtenidos. A la sumatoria resultante $\sum \bar{x}^2 P(\bar{x})$ le restamos el cuadrado de la esperanza matemática $[E(\bar{x})]^2$

$\bar{x}$	$P(\bar{x})$	$X^2 * P(\bar{X})$
		$\sum \bar{x}^2 * P(\bar{x})$ =

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

La fórmula de cálculo es la siguiente:

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

f) Para el **muestreo con reposición** (MCR) se verifican las siguientes relaciones entre los **Parámetros Poblacionales** y los **Parámetros de la Distribución de $\bar{x}$**

1) $E(\bar{x}) = \mu$ 2) $\sigma^2(\bar{x}) = \sigma^2 / n$ 3) $\sigma(\bar{x}) = \sigma / \sqrt{n}$

La Esperanza matemática de la Media Muestral es igual a la Media Poblacional (μ)

La Varianza de la Media Muestral es igual al cociente entre la Varianza Poblacional (σ^2) y el tamaño de cada muestra (n).
La Desviación de la Media Muestral es igual al cociente entre la Desviación Poblacional (σ) y la raíz cuadrada del tamaño de cada muestra.