

Resumen distribuciones de muestreo

Distribuciones de muestreo

$$N = 5$$

1. De una población finita de tamaño N , se extraen de manera aleatoria todas las muestras posibles de tamaño n
2. Se calcula la estadística de interés para cada muestra.
3. Listar en una columna los distintos valores observados de la estadística y en otra columna las frecuencias correspondientes de cada valor observado.

en donde la variable x es la edad de un niño

$$x_1 = 6, x_2 = 8, x_3 = 10, x_4 = 12, x_5 = 14$$

la media se puede calcular como:

$$\mu = \frac{\sum_{i=1}^n x_i}{N} = \frac{6 + 8 + 10 + 12 + 14}{5} = 10$$

Y la varianza:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{N} = \frac{1}{5} (6-10)^2 + (8-10)^2 + (10-10)^2 + (12-10)^2 + (14-10)^2 = 16$$

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{N} = \frac{1}{5} (16 + 4 + 0 + 4 + 16) = \frac{40}{5} = 8$$

Si queremos hallar la distribución de muestreo, debemos recolectar todas las posibles muestras de una población de tamaño 5.

para tomar todas las posibles muestras, las tomamos tanto con orden como con reemplazo, el número de muestras posibles es:

$$N^n = 5^2 = 25$$

Podemos entonces, sacar todas las posibles muestras de tamaño 2, calcular el estadístico de interés, en este caso \bar{x} , listarlo y determinar su frecuencia relativa:

Muestreo a partir de una distribución normal

Cuando el muestreo se realiza a partir de una población que sigue una distribución normal, la distribución de la media muestral tiene las siguientes propiedades:

1. la distribución de \bar{x} será normal.
2. La media, $\mu_{\bar{x}}$ de la distribución de \bar{x} será igual a la media de la población de la cual se extrajo.
3. La varianza $\sigma_{\bar{x}}^2$ de la distribución de \bar{x} será igual a la varianza de la población dividida entre el tamaño de muestra.

Teorema del límite central

Si \bar{x} es la media de una muestra aleatoria de tamaño n que se toma de una población con media μ y varianza finita σ^2 , entonces la variable $z_n = \frac{\sqrt{n}(\bar{x} - \mu)}{\sigma}$ tiende a distribuirse como una normal estándar a medida que n tiende a infinito.

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

Ejemplo: distribuciones de muestreo

Para una población de tamaño 5

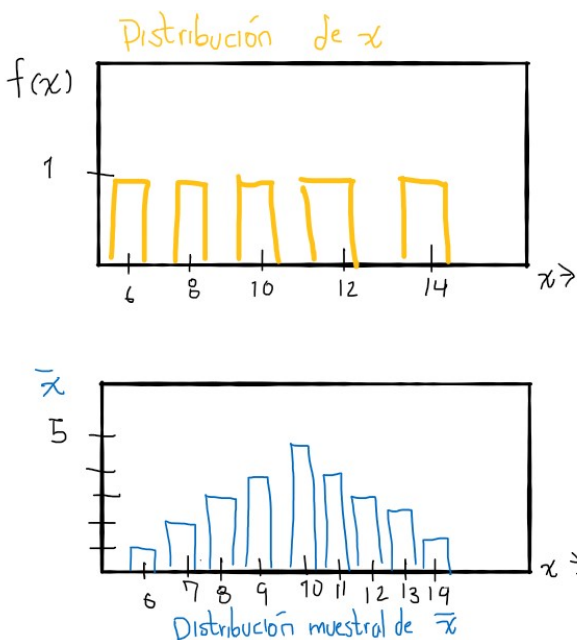
todas las posibles muestras de $n=2$
 Estadístico de interés: \bar{x}

	Segundo muestreo					\bar{x}	f_i	f_r
	6	8	10	12	14	6	1	1/25
						7	2	2/25
						8	3	3/25
						9	4	4/25
						10	5	5/25
						11	4	4/25
						12	3	3/25
						13	2	2/25
						14	1	1/25
							25	

Primer muestreo

6	6,6	6,8	6,10	6,12	6,14
8	8,6	8,8	8,10	8,12	8,14
10	10,6	10,8	10,10	10,12	10,14
12	12,6	12,8	12,10	12,12	12,14
14	14,6	14,8	14,10	14,12	14,14

y graficar las frecuencias para evaluar las distribuciones de la variable y del estadístico \bar{x} de las muestras.



Es importante realizar una muestra probabilística para tener en cuenta que cuando otro experimentador saque una muestra diferente tendrá resultados diferentes.

Podemos calcular la media de \bar{x} , $\mu_{\bar{x}}$:

$$\mu_{\bar{x}} = \frac{\sum_{i=1}^{25} x_i}{N^n} = \frac{\sum_{i=1}^{25} x_i}{25}$$

$$\mu_{\bar{x}} = \frac{250}{25} = 10$$

La media muestral, coincide con la media poblacional.

$$\mu_{\bar{x}} = \frac{250}{25} = 10$$

$$\sigma_{\bar{x}}^2 = \frac{\sum_{i=1}^{N^n} (\bar{x} - \mu_{\bar{x}})^2}{N^n}$$

$$\sigma_{\bar{x}}^2 = \frac{(6-10)^2 + 2 \cdot (7-10)^2 + \dots + (14-10)^2}{25}$$

$$\sigma_{\bar{x}}^2 = \frac{100}{25} = 4$$

El resultado anterior nos permite observar que la varianza de la muestra, es igual a la varianza de la población dividida por el tamaño de muestra.

Tenemos los resultados:

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

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

Error estándar:

el error estándar es la raíz de la varianza:

$$\text{error estándar} = \frac{\sigma}{\sqrt{n}}$$

El error estándar es útil en los intervalos de confianza

Cuando la población presenta una distribución normal

1. la distribución de \bar{x}
2. la media $\mu_{\bar{x}}$ de la distribución de \bar{x} será igual a la media de la población de la cual se extrajo.
3. La varianza $\sigma_{\bar{x}}^2$ de la distribución de \bar{x} será igual a la varianza de la población dividida entre el tamaño de muestra.

Muestras aleatorias y distribuciones de muestreo

Muestra aleatoria

Sea X una característica de interés y $f(x)$ su función de densidad de probabilidad.

Un conjunto de n variables aleatorias, x_1, x_2, \dots, x_n independientes e idénticamente distribuidas se denomina muestra aleatoria

Estadística

Cualquier función de las variables aleatorias que conforman una muestra aleatoria, de tal forma que al realizar la observación no contiene valores desconocidos. Se denota con T .

Ejemplos: Sumatoria, media, rango, desviación.

Distribución de muestreo de una estadística.

Sea T una estadística, la distribución de T es la función de densidad de probabilidad que se obtendría de un número infinito de muestras aleatorias independientes. cada una de tamaño n , provenientes de la población de interés.

Distribución de muestreo de \bar{x}

Teorema 1: Distribución de muestreo de \bar{x} varianza conocida, normalidad

Sea x una variable normalmente distribuida y x_1, x_2, \dots, x_n una muestra aleatoria, con promedio $E(x_i) = \mu$ y varianza $V(x_i) = \sigma^2$, $i = 1, 2, \dots, n$,

Entonces la distribución de \bar{x} es también *normal*, con promedio $E(\bar{x}) = \mu$ y varianza $\frac{\sigma^2}{n}$

Es decir, si $x \sim N(\mu, \sigma^2)$, Entonces $\bar{x} \sim N(\mu, \sigma^2/n)$ de donde

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma^2}{\sqrt{n}}} \sim N(0, 1)$$

Teorema 2 (Teorema del límite central):

Sea x una variable con distribución no especificada y x_1, x_2, \dots, x_n una muestra aleatoria, con promedio $E(x_i) = \mu$ y varianza $V(x_i) = \sigma^2$, $i = 1, 2, \dots, n$, teniendo μ y σ^2 finitas, entonces \bar{x} tiene una distribución que tiende a $\sim t(\mu, \sigma^2/n)$

Entonces:

$$T = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t \text{ student con } n-1 \text{ grados de libertad}$$

** obserbación **

Cuando el tamaño de muestra tiende a infinito:

cuando $n \rightarrow \infty$:

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim N(0, 1)$$

Distribucion de muestreo para varianza σ^2

Si $x \sim N(\mu, \sigma^2)$ entonces:

$$\chi^2 = \frac{(n-1) \cdot s^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

Problemas

Dado que χ^2 es asimétrica, las estimaciones no son tan buenas como cuando se trabaja con distribuciones simétricas. El intervalo de confianza va a ser amplio debido al sesgo de la distribución.