Parámetros Poblacionales

- Estadísticos muéstrales

Hemos estudiado diferentes medidas numéricas correspondientes a conjuntos de datos, entre otras, estudiamos la media, la desviación estándar etc. Ahora vamos a distinguir entre medidas numéricas calculadas con conjuntos de datos poblacionales y las calculadas con datos muestrales. Si la medida numérica se calcula para el conjunto de datos poblacionales le llamaremos valor del parámetro poblacional y si se calcula para el conjunto, de datos muestrales, le llamaremos valor del estadístico muestral.

Parámetros

En una población finita de tamaño *N* los parámetros poblacionales ***media***, ***varianza*** y ***proporción poblacional*** vienen dados por:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} X_i$$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2$$

$$p = \frac{X}{N} = \frac{\#ENP}{\#P}$$

Estadísticos

Para una muestra aleatoria simple de tamaño *n*, $(X_1, \ldots, X_n)$ los estadísticos: ***media***, ***varianza*** y ***proporción muestral*** se definen como:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{N}(X_i - \overline{X})^2$$

$$p = \frac{X}{n} = \frac{\#EnP}{\#P}$$

Si en vez de considerar las n variables aleatorias, independientes e idénticamente distribuidas ($X_1,\ldots,X_n$) que constituyen la muestra aleatoria simple, consideramos una muestra concreta ($x_1, \ldots, x_n$) entonces los valores de estos estadísticos muestrales tomarían la siguiente forma:

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

$$x^2 = \frac{1}{n-1} \left(x_i - \bar{x} \right)^2$$

$$\hat{p} = \frac{x}{n} = \frac{\# EnP}{\# P}$$

Veremos que el estadístico es una función de las observaciones muestrales, y en estos casos asigna a cada muestra observada: la media de los valores, la varianza o la proporción, respectivamente.

Función de distribución empírica

La función de distribución de una variable aleatoria X está definida como: $F(x) = P(X < x)$

y representa la proporción de valores que son menores o iguales que x. De manera similar podemos definir la función de distribución empírica para una muestra.

Definición

Consideremos una población con función de distribución $F(x)$ y sean $(x_1, \ldots, x_n)$ los valores observados correspondientes a una MAS procedente de esa población, designamos por $N(x)$ el número de valores observados que son menores o iguales que x. Entonces definimos la función de distribución empírica de la muestra, que la notaremos por $F(x)$

$$F_n(x) = \frac{N(x)}{n}$$

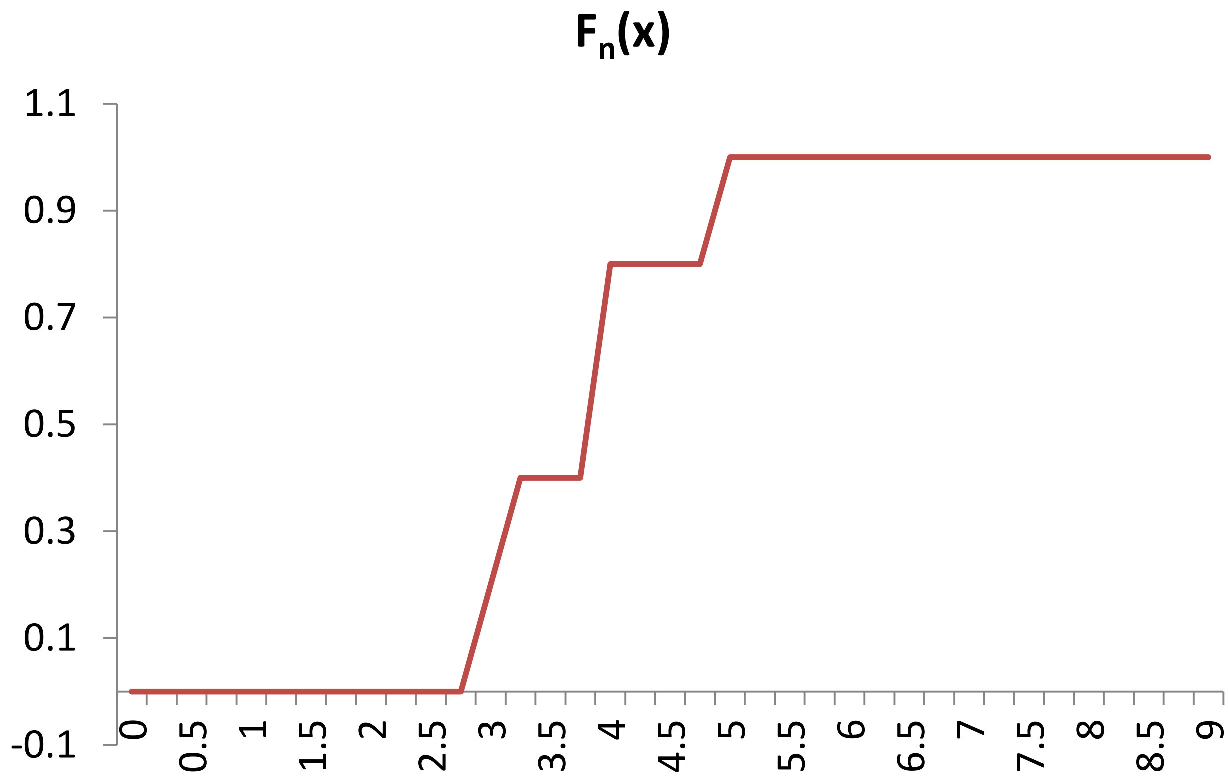
Ejemplo 2

Dada una muestra aleatoria formada por las observaciones muestrales (3, 8, 5, 4, 5). Obtener la función de distribución empírica y su correspondiente representación gráfica.

solución

observaciones muestrales	x		N(x)	$F_n(x)$
3	0	<3	0	0
8	3	<=3	1	0.2
5	4	<=4	2	0.4
4	5	<=5	4	0.8
5	8	<=8	5	1

$$F_n(x) = \begin{cases} 0 & x < 3 \\ 0.2 & x \leq 3 \\ 0.4 & x \leq 4 \\ 0.8 & x \leq 5 \\ 1 & x \leq 8 \end{cases}$$



La función de distribución empírica tiene las mismas propiedades que la función de distribución de la variable aleatoria, y se puede demostrar, utilizando el teorema de Glivenko-Cantelli, que $F_n(x)$ converge en probabilidad a $F(x)$. Para efectos prácticos, implica que cuando el tamaño de la muestra crece, la gráfica de la función de distribución empírica se aproxima bastante a la de la función de distribución de la población, se puede entonces utilizar como estimador de la población. De todo esto se deduce que la función de distribución empírica o su gráfica puede utilizarse para determinar la forma general de la distribución poblacional. También es fácil y muy frecuente el reconocer la forma de la distribución observando el histograma correspondiente que nos daría idea de la función de densidades.

DISTRIBUCIÓN MUESTRAL DE ESTADISTICOS

Los estadísticos muestrales (proporción media y varianza muestral) se pueden utilizar para estimar los correspondientes parámetros poblacionales. Así pues, para estudiar propiedades de estos estadísticos como estimadores poblacionales, será necesario estudiar las características de la distribución de estos estadísticos.

Los estadísticos muestrales se calculan a partir de los valores (X_1,..., X_n) de una muestra aleatoria. Estos estadísticos son también variables aleatorias, como tales tienen asignada una ~~distribución de probabilidades~~. Entonces los estadísticos muestrales: Proporción, media, varianza, etc., tienen su correspondiente distribución de probabilidad. Si estas distribuciones de probabilidad se pueden obtener, será entonces posible establecer afirmaciones basadas en probabilidades acerca de estos estadísticos.

Nos interesamos en determinar las distribuciones de probabilidad de algunos estadísticos muestrales, en concreto, para la media y varianza muestral, que serán de utilidad en distintas aplicaciones estadísticas.

Ejemplo 3

	Tienda	Horas de servicio
T1	1	12
T2	2	10
T3	3	14
T4	4	9
T5	5	10

Supongamos una población formada por cinco tiendas existentes en cierto municipio del Edo. de México. La característica a investigar será el número de horas que diariamente permanecen abiertas esas tiendas y que representaremos por la variable aleatoria X.

X= Horas de servicio de las tiendas

Parámetro media poblacional

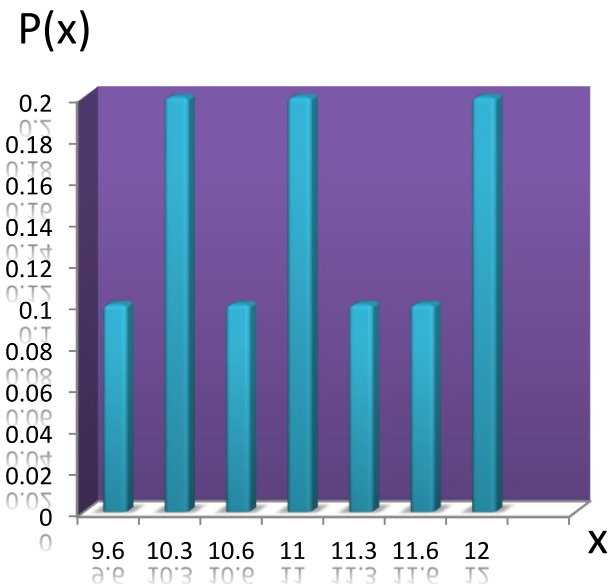
$$\mu = \frac{1}{N}\sum_{i=1}^{N} X_i = \frac{1}{5}(12+10+14+9+10) = \frac{55}{5} = 11$$

Parámetro varianza poblacional

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i-\mu) = \frac{1}{5}\begin{pmatrix}[12-11]^2+[10-11]^2+ \\ [14-11]^2+[9-11]^2+[10-11]^2\end{pmatrix} = \frac{16}{5}$$

Existen diez posibles muestras aleatorias simples de tamaño 3 que se pueden tomar y el valor del estadístico media muestral La distribución de probabilidad del estadístico media muestral 2 viene
dada por la Tabla 1.8.

Distribución muestral del estadístico media muestral $\overline{X}$



Muestras	observaciones muestrales	Estadístico media muestral
T1,T2,T3	12, 10 14	12
T1,T2,T4	12, 10, 9	10.3
T1,T2,T5	12, 10, 10	10.6
T1,T3,T4	12, 14, 9	11.6
T1,T3,T5	12, 14, 10	12
T1,T4,T5	12, 9, 10	10.3
T2,T3,T4	10, 14, 9	11
T2,T3,T5	10, 14, 10	11.3
T2,T4,T5	10, 9, 10	9.6
T3,T4,T5	14, 9, 10	11

Valor del Estadístico Muestral	Número de repeticiones	Probabilidad
9.6	1	0.1
10.3	2	0.2
10.6	1	0.1
11	2	0.2
11.3	1	0.1
11.6	1	0.1
12	2	0.2

Ejemplo 3

Una empresa dedicada al transporte y distribución de mercancías, tiene una plantilla de *50* trabajadores. Durante el último año se ha observado que *25* trabajadores han faltado un solo día al trabajo, 20 trabajadores han faltado dos días y 5 trabajadores han faltado tres días. Si se toma una muestra aleatoria, con reemplazamiento, de tamaño dos (X_1, X_2) del total de la plantilla, obtener:

1. La distribución de probabilidad del número de días que ha faltado al trabajo un empleado, su media y su varianza.
2. Distribución de probabilidad del estadístico media muestral.
3. La distribución de probabilidad del estadístico varianza muestral.
4. La media y varianza del estadístico media muestral.
5. La probabilidad de que el estadístico media muestral, sea menor que 2.
6. La media y varianza del estadístico varianza muestral.
7. La probabilidad de que el estadístico varianza muestral, sea menor o igual que 0.5.