

Introducción a la Inferencia Estadística

Inferencia Estadística

Población: Es una variable aleatoria con una cierta distribución de probabilidad con sus correspondientes parámetros.

Ejemplo: Se desea estudiar la población de todos los pesos de los recién nacidos en Tucumán durante el mes de Setiembre de 2019.

En muchos casos es imposible acceder a todos los valores de una población, por las siguientes razones:

1. Costo económico
2. Costo en tiempo
3. Muestreo destructivo

Inferencia Estadística

Por esta razón se debe tomar una muestra:

- **Definición: Muestra aleatoria (m.a.):**
- Sea X una v. a. una m.a. de X es un conjunto de vs. as. X_1, X_2, \dots, X_n . independientes e idénticamente distribuidas, iid, como X

IMPORTANTE:

Dada una m.a, esto es: X_1, X_2, \dots, X_n . iid con media μ y varianza σ^2 , se tiene

que: $E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} E\left(\sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \sum_{i=1}^n \mu = \frac{1}{n} n\mu = \mu$

$$V(\bar{X}) = V\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} V\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{1}{n^2} \sum_{i=1}^n \sigma^2 = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

Ahora veamos para la suma:

$$E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = \sum_{i=1}^n \mu = n\mu$$

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) = \sum_{i=1}^n \sigma^2 = n\sigma^2$$

Ejemplo

Suponga que se quiere hacer un estudio sobre la edad de los ingresantes a una cierta facultad en el año 1997. La siguiente tabla muestra la composición de las edades de los ingresantes:

Edad	17	18	19	20	21	22
%	12	30	30	15	10	3

La población es el conjunto de todas las edades o sea $\{17, 18, 19, 20, 21, 22\}$ con sus respectivas probabilidades 0.12, 0.30, 0.30, 0.15, 0.10, 0.03.

Se tiene:

X = edad de los ingresantes

Ejemplo

Si se desea tomar una muestra de n estudiantes con reposición se tiene:

X_1 = edad del primer estudiante

X_2 = edad del segundo estudiante

\vdots
 X_n = edad de n -ésimo estudiante

$$P(X_1=17)=0.12, P(X_1=18)=0.30, \dots, P(X_1=22)=0.03$$

$$P(X_2=17)=0.12, P(X_2=18)=0.30, \dots, P(X_2=22)=0.03$$

$$\vdots \quad \vdots \quad \vdots \\ P(X_n=17)=0.12, P(X_n=18)=0.30, \dots, P(X_n=22)=0.03$$

X	P _X (X)
17	0.12
18	0.30
19	0.30
20	0.15
21	0.10
22	0.03

Entonces X_1, X_2, \dots, X_n son réplicas independientes de un experimento aleatorio y al llevarlo a cabo producen la muestra aleatoria X_1, X_2, \dots, X_n (son variables aleatorias iid como X).

¿Qué quiero?

Dada una v. a. de interés, digamos X , y dada una muestra observada de la misma:

- ❖ Se desea conocer de qué distribución proviene la muestra
- ❖ Cuales son sus parámetros
- ❖ ¿Si disponemos de la posibilidad de tomar muestras, utilizamos esa muestra, sino LA SIMULAMOS.