

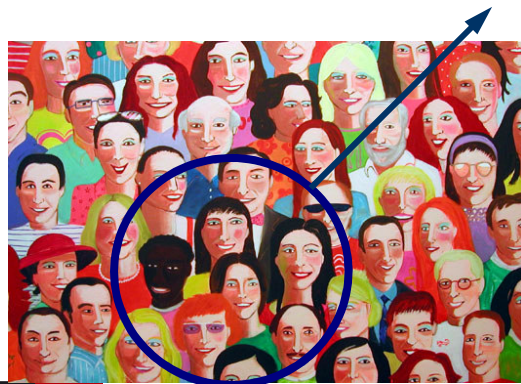
Estadística

Inferencia Estadística

Polación y muestra

Objetivos de la Estadística:

- Aprender de la observación.
- Generalizar lo que aprendemos de una **muestra** a toda la **población**



Inferencia

Muestra de n observaciones: $X_1, X_2, \dots, X_n$ de una población X

Muestra aleatoria simple:

Las X_i tienen las mismas características que X .

Son independientes entre sí.



Cada muestra nos da una información:

Histograma, media
muestral varianza
muestral ,...

**La información
depende de la muestra
elegida**

Estadístico

Un **Estadístico** es una función de la muestra, por lo tanto:

- Su valor depende de los valores de la muestra.
- Es una **variable aleatoria**
- Su distribución es la **distribución muestral**

Ejemplo: La media muestral:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Distribución de la media muestral

Sea X una población con $E[X] = \mu$ $V[X] = \sigma^2$

¿Cuál es la distribución

de la media muestral?

$$E[\bar{X}] = E\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] = \frac{E[X_1] + E[X_2] + \dots + E[X_n]}{n} = \frac{n\mu}{n} = \mu$$

$$V[\bar{X}] = V\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] = \frac{V[X_1] + V[X_2] + \dots + V[X_n]}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

Son
independientes

Distribución de la media muestral

La media muestral es una suma de variables aleatorias

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{X_1}{n} + \frac{X_2}{n} + \dots + \frac{X_n}{n}$$

Por el Teorema Central del Límite

$$\bar{X} \approx N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Problemas en Inferencia Estadística

POBLACIÓN $\rightarrow X \approx F(\theta)$ desconocido

A partir de una M.A.S. $X_1, X_2, \dots, X_n$ queremos **estimar** el valor de θ

Estimar : Asignar un valor a algo que desconocemos

¿Cómo?

A partir del modelo F y de la muestra

- ➡ Estimación puntual: Un solo valor
- ➡ Estimación por intervalos: Un rango de valores
- ➡ Contraste de hipótesis: Rechazo o aceptación de ciertas hipótesis sobre el parámetro

Estimador

Un **estimador** es un estadístico con la misión de acercarse lo más posible al verdadero valor del parámetro

Valor del parámetro θ  Estimador $\hat{\theta}$

¿Cómo seleccionamos el estimador más adecuado?

Propiedades deseables de los estimadores

- Error cuadrático medio

$$E\left[(\hat{\theta} - \theta)^2\right] = V[\hat{\theta}] + Sesgo^2$$

donde $Sesgo = E[\hat{\theta}] - \theta$

Buscamos el estimador con menor error cuadrático medio:



Menor sesgo

Menor varianza

Estimador insesgado

Un estimador es **insesgado** o **centrado** si su **sesgo** es cero

Entonces

$$E[\hat{\theta}] - \theta = 0 \Rightarrow E[\hat{\theta}] = \theta$$

Buscamos estimadores insesgados

Ejemplo: La media muestral es insesgada para estimar la media poblacional

$$E[\bar{X}] = \mu$$

Ejemplo

Sea una población X con media desconocida y varianza 1. Se toma una muestra de tamaño 3. Se consideran los siguientes estimadores de la media poblacional

$$\hat{\mu}_1 = \frac{2}{5}X_1 + \frac{1}{5}X_2 + \frac{3}{5}X_3$$

$$\hat{\mu}_2 = \frac{1}{8}X_1 + \frac{5}{8}X_2 + \frac{2}{8}X_3$$

$$\hat{\mu}_3 = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$$

¿Son insesgados?

Solución

- Calculamos la esperanza de cada uno de ellos

$$\begin{aligned}
 E[\hat{\mu}_1] &= E\left[\frac{2}{5}X_1 + \frac{1}{5}X_2 + \frac{3}{5}X_3\right] = E\left[\frac{2}{5}X_1\right] + E\left[\frac{1}{5}X_2\right] + E\left[\frac{3}{5}X_3\right] = \\
 &= \frac{2}{5}E[X_1] + \frac{1}{5}E[X_2] + \frac{3}{5}E[X_3] = \frac{2}{5}\mu + \frac{1}{5}\mu + \frac{3}{5}\mu = \mu \\
 E[\hat{\mu}_2] &= E\left[\frac{1}{8}X_1 + \frac{5}{8}X_2 + \frac{2}{8}X_3\right] = E\left[\frac{1}{8}X_1\right] + E\left[\frac{5}{8}X_2\right] + E\left[\frac{2}{8}X_3\right] = \\
 &= \frac{1}{8}E[X_1] + \frac{5}{8}E[X_2] + \frac{2}{8}E[X_3] = \frac{1}{8}\mu + \frac{5}{8}\mu + \frac{2}{8}\mu = \mu \\
 E[\hat{\mu}_3] &= E\left[\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3\right] = E\left[\frac{1}{3}X_1\right] + E\left[\frac{1}{3}X_2\right] + E\left[\frac{1}{3}X_3\right] = \\
 &= \frac{1}{3}E[X_1] + \frac{1}{3}E[X_2] + \frac{1}{3}E[X_3] = \frac{1}{3}\mu + \frac{1}{3}\mu + \frac{1}{3}\mu = \mu
 \end{aligned}$$

TODOS SON
INSESADOS

Ejemplos

- ¿Es la varianza muestral un estimador insesgado para estimar la varianza poblacional?

$$\hat{\sigma}^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$$

$$\begin{aligned}
 E[\hat{\sigma}^2] &= E\left[\frac{1}{n} \sum (x_i - \bar{x})^2\right] = E\left[\frac{1}{n} \sum x_i^2 - \bar{x}^2\right] = \frac{1}{n} \sum E[x_i^2] - E[\bar{x}^2] = \\
 &= \frac{1}{n} \sum (\underbrace{\sigma^2 - \mu^2}_{\hat{\sigma}^2 = E[x_i^2] - \mu^2}) - \left(\underbrace{\frac{\sigma^2}{n} - \mu^2}_{\frac{\hat{\sigma}^2}{n} = E[\bar{x}^2] - \mu^2}\right) = \frac{1}{n} n(\sigma^2 - \mu^2) - \left(\frac{\sigma^2}{n} - \mu^2\right) = \frac{n-1}{n} \sigma^2
 \end{aligned}$$

NO es insesgado

Cuasi-varianza

- Como la varianza muestral no es insesgado para estimar la varianza poblacional, buscamos un estimador que sí lo sea:

Cuasi-varianza muestral

$$\hat{S}^2 = \frac{1}{n-1} \sum_i (x_i - \bar{x})^2 = \frac{n}{n-1} \hat{\sigma}^2$$

$$E[\hat{S}^2] = E\left[\frac{n}{n-1} \hat{\sigma}^2\right] = \frac{n}{n-1} E[\hat{\sigma}^2] = \frac{n}{n-1} \left(\frac{n-1}{n} \sigma^2\right) = \sigma^2$$

Recordemos que por el
Teorema de Fisher

$$\frac{n \hat{\sigma}^2}{\sigma^2} \approx \chi_{n-1}^2 \Leftrightarrow \frac{(n-1) \hat{S}^2}{\sigma^2} \approx \chi_{n-1}^2$$

Estimador Eficiente

Un estimador es **eficiente** si tiene varianza mínima

Para buscar el estimador eficiente calculamos la varianza de los estimadores posibles y nos quedamos con el que tenga la menor

¿Cuál de los estimadores anteriores tiene menor varianza?

Solución

Calculamos sus varianzas

$$\begin{aligned} V[\hat{\mu}_1] &= V\left[\frac{2}{5}X_1 + \frac{1}{5}X_2 + \frac{3}{5}X_3\right] = V\left[\frac{2}{5}X_1\right] + V\left[\frac{1}{5}X_2\right] + V\left[\frac{3}{5}X_3\right] = \\ &= \left(\frac{2}{5}\right)^2 V[X_1] + \left(\frac{1}{5}\right)^2 V[X_2] + \left(\frac{3}{5}\right)^2 V[X_3] = \frac{4}{25} \cdot 1 + \frac{1}{25} \cdot 1 + \frac{9}{25} \cdot 1 = \frac{13}{25} \end{aligned}$$

$$V[\hat{\mu}_2] = V\left[\frac{1}{8}X_1 + \frac{5}{8}X_2 + \frac{2}{8}X_3\right] = \frac{1}{64} + \frac{25}{64} + \frac{4}{64} = \frac{30}{64}$$

$$V[\hat{\mu}_3] = V\left[\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3\right] = \frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{1}{3}$$

Métodos de construcción de estimadores

- Método de los momentos



- Fáciles de calcular
- NO siempre tiene buenas propiedades

- Método de máxima verosimilitud



- Buenas propiedades
- Más difíciles de calcular

Método de los momentos

JUSTIFICACIÓN: Los momentos muestrales convergen a los poblacionales

Para estimar la media

poblacional μ



la media muestral $\bar{X}$

Para estimar la varianza

poblacional σ^2



la varianza muestral $\hat{\sigma}^2$

Para estimar la proporción

poblacional p



la proporción muestral $\hat{p}$

Método de los momentos

Procedimiento: Para calcular el estimador de un parámetro debemos igualar los momentos muestrales a los poblacionales. (tantos como parámetros queramos estimar) y resolver el sistema

$$E[X] = \frac{1}{n} \sum_{i=1}^n X_i$$

...

$$E[X^k] = \frac{1}{n} \sum_{i=1}^n X_i^k$$

Ejemplos

1. Calcular el estimador mediante el método de los momentos del parámetro p de una población X con distribución $B(m,p)$ con m conocido.

Establecemos la ecuación (sólo hay que estimar 1 parámetro $\rightarrow$ 1 ecuación).

$$E[X] = \frac{1}{n} \sum X_i \Rightarrow mp = \bar{x} \quad \Rightarrow \quad \hat{p} = \frac{\bar{x}}{m}$$

Ejemplos

2. Sea una variable aleatoria X con distribución de Poisson de parámetro desconocido. Se toma una muestra de tamaño 5 con los siguientes resultados: 2,3,0,2,1.

Obtener el estimador por el método de los momentos y calcularlo

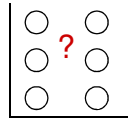
Establecemos la ecuación:

$$E[X] = \frac{1}{n} \sum X_i \Rightarrow \lambda = \bar{x} \quad \Rightarrow \quad \hat{\lambda} = \bar{x}$$

$$\hat{\lambda} = \bar{x} = \frac{2+3+0+2+1}{5} = \frac{8}{5}$$

Método de Máxima Verosimilitud

El principio de máxima verosimilitud está basado en un principio que se ilustra con el siguiente ejemplo:



Supongamos que tenemos una urna con 6 bolas de color blanco y negro, no todas del mismo color, pero que no sabemos en que proporción.

Para adivinar la composición de la urna se pueden extraer 2 bolas con reposición

B \ p	1/6	2/6	3/6	4/6	5/6
0	$(5/6)^2$	$(4/6)^2$	$(3/6)^2$	$(2/6)^2$	$(1/6)^2$
1	$2 \frac{1}{6} \frac{5}{6}$	$2 \frac{2}{6} \frac{4}{6}$	$2 \frac{3}{6} \frac{3}{6}$	$2 \frac{4}{6} \frac{2}{6}$	$2 \frac{5}{6} \frac{1}{6}$
2	$(1/6)^2$	$(2/6)^2$	$(3/6)^2$	$(4/6)^2$	$(5/6)^2$

Método de máxima verosimilitud

¿Cuántas bolas blancas hay en la urna?

Depende del número de bolas blancas extraídas :

- Ninguna bola blanca: $B=0$ $\rightarrow$ $p=1/6$ Hay una bola blanca
- Una bola blanca: $B=1$ $\rightarrow$ $p=3/6$ Hay 3 bolas blancas
- Dos bolas blancas: $B=2$ $\rightarrow$ $p=5/6$ Hay 5 bolas blancas

Buscamos el valor que maximiza la probabilidad

Método de máxima verosimilitud

Función de verosimilitud:

$$L_{\theta}(x_1, \dots, x_n) = \begin{cases} p(x_1) \cdots p(x_n) & \text{caso discreto} \\ f(x_1) \cdots f(x_n) & \text{caso continuo} \end{cases}$$

El **estimador máximo verosímil** es el que maximiza el logaritmo de la función de verosimilitud

$$\max_{\theta} \ln L_{\theta}(x_1, \dots, x_n) = L(\hat{\theta}, X_1, \dots, X_n)$$

Observaciones

- ❑ Siempre es un valor del espacio paramétrico.
- ❑ Si $\hat{\theta}$ es el estimador de máxima verosimilitud de θ , entonces $g(\hat{\theta})$ es el estimador de máxima verosimilitud de $g(\theta)$.
- ❑ No es único.
- ❑ No tiene por qué ser insesgado.
- ❑ Es asintóticamente $N(\hat{\theta}, \sqrt{V(\hat{\theta})})$

Ejemplo

Calcular el estimador máximo verosímil del parámetro p de la distribución Binomial $B(m;p)$, en una muestra aleatoria simple de tamaño n .

Tomamos una mas de la población y calculamos su función de verosimilitud (caso discreto)

$$L_p(X_1, \dots, X_n) = \binom{m}{x_1} p^{x_1} (1-p)^{m-x_1} \dots \binom{m}{x_n} p^{x_n} (1-p)^{m-x_n} =$$
$$\prod_{i=1}^n \left[\binom{m}{x_i} p^{x_i} (1-p)^{m-x_i} \right] = \prod_{i=1}^n \binom{m}{x_i} p^{\sum x_i} (1-p)^{nm - \sum x_i}$$

Para calcular el estimador, tomamos logaritmos y calculamos el máximo:

$$\ln L_p(X_1, \dots, X_n) = \sum \ln \binom{m}{x_i} + \sum x_i \ln p + (nm - \sum x_i) \ln(1-p)$$

Ejemplo

Calculamos el máximo, derivando e igualando a 0, por ser la función derivable:

$$\frac{\partial \ln L_p(X_1, \dots, X_n)}{\partial p} = \sum x_i \frac{1}{p} + (nm - \sum x_i) \frac{-1}{1-p} = 0$$

$$\frac{\sum x_i}{p} - \frac{(nm - \sum x_i)}{1-p} = 0 \Rightarrow \hat{p} = \frac{\sum x_i}{nm} = \frac{\bar{x}}{m}$$