Práctica 8

miércoles, 26 de junio de 2024 18:11

A partir de una variable aleatoria cuya distribución no conocemos (nos dan valores medidos de la variable) tratar de inferir algo sobre la distrtibución de la variable

Población: conjunto de todos los objetos de interés
Característica que me interesa es variable aleatoria, ej: x altura de un estudiante tomado al azar

Si tengo x1, x2... x20 (20 muestras distintas)

$$\mu = E(X) \approx \frac{1}{20}\sum_{i=1}^{20} x_i$$

Siempre que tomo una muestra para estimar un valor desconocido consigo un aproximado, no el verdadero. Tendría que tomar toda la muestra para que sea el posta

Estimación puntual
$\overline{X}$ es el estimador puntual de mu, el valor que toma es la estimación
Si x1,x2..x3.. Se miden bajo las mismas condiciones puedo asumir que tienen la misma distribución de probabilidad (IDENTICAMENTE DISTRIBUIDAS)
SI O SI TIENEN QUE SER INDEPENDIENTES
Entonces es una muestra aleatoria de la variable aleatoria original x
También si se la distribución pero no las propiedades
A partir de las variables aleatorias genero un estadístico: una función de la muestra aleatoria

$$h(X_1, X_2 \ldots X_n)$$

Función de n variables aplicada a la muestra aplicada que da como resultado otra variable aleatoria que es el estadístico.
Estimador puntual cuando uso un estadístico para estimar un parámetro desconocido

$\hat{\mu} = \overline{x}$ es los de arriba.

El estadístico más común es el promedio muestral x barra

Varianza también es desconocido lo puedo estimar como varianza muestral.

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

Puedo también (por ejemplo), estimar a $\hat{\mu}$ con $\frac{X_1 + X_2}{2}$

Cuál es mejor??

Estimador insesgado:

Sea θ un parámetro desconocido de var aleatoria X

Sea $\hat{\theta}$ estimador puntual

$\hat{\theta}$ es insesgado si $E(\hat{\theta}) = \theta$ para todo θ

con $X_1, X_2 \ldots X_n$ muestra aleatoria de var aleatoria X

$\mu = E(X)$

$$\hat{\mu} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

prop de linealidad. $E(A \cdot X) = A \cdot E(X) \to$ linealidad.

$E(\hat{\mu}) = E(\overline{X}) = E\left(\frac{1}{n}\sum x_i\right) \to$ esperanza de una suma es

$$E(\hat{\mu}) = E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)$$ → esperanza de una suma es suma de esperanzas.

$$= \frac{1}{n} E\left(\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu$$ → porque me lo dicen

→ suma de constante μ

$$= \frac{1}{\cancel{n}} \cancel{n}\mu = \mu.$$ → promedio muestral es estimador insesgado.

Si tengo dos estimadores y ambos son insesgados, me fijo las varianzas y el que tenga más chica es mejor

$$V(\hat{\mu}_1) = V\left(\frac{1}{n-1}\sum_{i=1}^{n-1} x_1\right) =$$

$V(aX) = a^2 \cdot V(X)$ → linealidad.

$$= \frac{1}{(n-1)^2} V\left(\sum_{i=1}^{n-1} x_1\right)$$ → varianza de una suma, si asumo independencia es suma de varianzas

$$= \frac{1}{(n-1)^2}\sum_{i=1}^{n-1} V(X_1) = \frac{1}{(n-1)^{\cancel{2}}} \cdot \cancel{n-1} \cdot \sigma^2$$

$$V(\hat{\mu}_1) = \underbrace{\frac{\sigma^2}{n-1}}_{\text{varianza}}$$

$$V(\hat{\mu}_2) = V\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n^2} V\left(\sum_{i=1}^{n} x_i\right)$$

$$= \frac{1}{n^2} \cdot \sum_{i=1}^{n} V(X_i) = \frac{1}{n^{\cancel{2}}} \cdot \cancel{n} \cdot \sigma^2 = \boxed{\frac{\sigma^2}{n}}$$

suma las independientes.

$n \geq 2.$

La varianza + chica es $\hat{\mu}_2$ por lo que $\hat{\mu}_2$ es mejor

Si uno es insesgado y uno no, calcular igual la varianza de ambos y ver error cuadrático medio

$$ECM(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$$ → variable. → estimador

cuánto se aleja de θ

usar este.

Teorema: $ECM(\hat{\theta}) = V(\hat{\theta}) + (b(\hat{\theta}))^2$

$b(\hat{\theta}) = E(\hat{\theta}) - \theta$ → sesgo de $\hat{\theta}$

si $\hat{\theta}_1$ y $\hat{\theta}_2$ son 2 est. puntuales de θ, es mejor el de menor ECM.

si dos insesgados, elijo $\hat{\theta}$ y comparar varianzas.

Estimador consistente:

$\hat{\theta}$ un estimador puntual de θ.

$\hat{\theta}$ obtenido a partir de n variables aleatorias,

$\hat{\theta}_n$ Es consistente si el estimador puntual cuando la muestra es grande coincide con el parámetro

Sea $\hat{\theta}_n$ un estimador puntual de θ

- si $E(\hat{\theta}_n) \xrightarrow[n \to \infty]{} \theta$ o $b(\hat{\theta}_n) \xrightarrow[n \to \infty]{} 0$
- y $V(\hat{\theta}_n) \xrightarrow[n \to \infty]{} 0$, $\hat{\theta}_n$ es est. consistente.

Si alguna condición no se da igual puede ser consistente, no lo sé

2 métodos para conseguir estimadores:
Ambos necesitan que yo sepa la distribución de la variable

Método de los momentos

Sea $\theta_1, \theta_2, \theta_n$ parámetros desconocidos de una v.a aleatoria X donde sé la distribución:

- k-ésimo momento poblacional es $E(X^k)$
- k-ésimo momento muestral $\frac{1}{n}\sum_{i=1}^{n} x_i^k$ si x_1, x_2, x_n es muestra aleatoria de tam n de X

planteo

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$E(X^2) = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\vdots$$

$$E(X^R) = \frac{1}{n}\sum_{i=1}^{n} x_i^R$$

Con x_1, x_2, x_r muestra aleatoria de una v.a X $X \sim N(\mu, \sigma^2)$, μ y σ^2 desconocidos

Encontrar estimadores puntuales por momentos.

$r = 2 \to 2$ incógnitas

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad E(X) = \mu$$

$$E(X^2) = \frac{1}{n}\sum x_i^2 \qquad V(X) = \sigma^2 \qquad \text{pero } \sigma$$

$i=1$ $\qquad E[X]=\mu$

$E(X^2)=\frac{1}{n}\sum x_i^3 \qquad V(X)=\sigma^2$

$$\begin{cases} \mu = \frac{1}{n}\sum_{i=1}^{n} x_i \\ \sigma^2+\mu^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 \end{cases} \Rightarrow \hat{\mu}=\bar{x} \rightarrow \text{estimador para } \sigma$$

$E(X^2)=$

$V(X)=E(X^2)-E(X)^2$

$\sigma^2=E(X^2)-\mu^2$

$E(X^2)=\sigma^2+\sigma^2$

$$\widehat{\sigma^2}=\frac{1}{n}\sum_{i=1}^{n} x_i^2-\bar{x}^2 \Rightarrow \text{estimador para } \sigma^2$$

otro ej: x_1, x_2, x_n con $x \sim Exp.(\lambda)$

$$E(X)=\frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\frac{1}{\lambda}=\bar{x} \Rightarrow \hat{\lambda}=\frac{1}{\bar{x}}$$

---

otro x_1, x_2, x_n muestra a. de va X

X continua con fdp.

$$f(x)=\begin{cases}\theta x^{\theta-1} & \text{si } 0<x<1 \\ 0 & cc\end{cases}$$

estimador de θ $\qquad E(X)=\int_0^1 x\cdot\theta x^{\theta-1}\,dx$

$$\left.\frac{\theta\, x^{\theta+1}}{\theta+1}\right|_0^1=\frac{\theta}{\theta+1}$$

$$E(X)=\frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\frac{\theta}{\theta+1}=\bar{x}$$

$$\theta=\bar{x}\cdot(\theta+1)$$

$$\theta=\bar{x}\,\theta+\bar{x}$$

$$\theta-\bar{x}\,\theta=\bar{x}$$

$$\theta\cdot(1-\bar{x})=\bar{x}$$

$$\hat{\theta}=\frac{\bar{x}}{1-\bar{x}}$$

Método de máxima verosimilitud

x_1, x_2, x_3, x_n muestra de var aleatoria X. θ es parámetro desconocido de la distribución de X

a estimar.

Se plantea func. de max verosimilitud

$$L(x_1, x_2, x_n) = P(X_1 = x_1) \cdot P(X_2 = x_2) \cdots P(X_n = x_n)$$

Se toma como estimador de θ a aquel valor de θ que maximiza $L(x_1, x_2, x_n)$

Ej x_1, x_2, x_n de muestra aleatoria, $X \sim P(\lambda)$

Hallar EMV de λ.

Tengo q' plantear estimador.

$$P(X = x_i) = \frac{e^{-\lambda} \lambda^{x_i}}{x_i!}$$

$$L(x_1, x_2, x_n, \lambda) = P(X_1 = x_1) \cdot P(X_2 = x_2) \cdots P(X_3 = x_3)$$

$$= \frac{e^{-\lambda} \cdot \lambda^{x_1}}{x_1!} \cdot \frac{e^{-\lambda} \cdot \lambda^{x_2}}{x_2!} \cdots \frac{e^{-\lambda} \cdot \lambda^{x_n}}{x_n!}$$

$$= \frac{(e^{-\lambda})^n \cdot \lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!} = \frac{e^{-\lambda n} \lambda^{\sum x_i}}{\prod x_i!}$$