

Resumen Primer Parcial Inferencia Estadística

Ignacio Pardo

Resumen Primer Parcial Inferencia Estadística

Introducción

Resumen de la materia de inferencia estadística de la Licenciatura en Tecnología Digital en la Universidad Torcuato Di Tella.

Contenido

- Resumen Primer Parcial Inferencia Estadística
 - Introducción
 - Contenido
 - Esperanza
 - Varianza
 - * Desvío Estándar
 - * Covarianza
 - * Correlación
 - Continuas
 - * Distribución Normal
 - Función acumulada
 - * Distribución Uniforme
 - * Distribución Exponencial
 - Discretas
 - * Distribución Bernoulli
 - * Distribución Binomial
 - * Distribución Poisson
 - Convergencia en Probabilidad
 - * Propiedades
 - Estimación por LGN
 - * Estimación Esperanza
 - * Estimación Varianza
 - * Estimación Proporción
 - * Estimación de Probabilidad
 - Formulas Consistencia
 - * Sesgo
 - Asintóticamente Insesgado
 - * Error Estándar
 - * Error Cuadrático Medio
 - Desigualdad de Chebyshev
 - Desigualdad de Markov
 - Momentos
 - * Momentos de una variable aleatoria
 - Discreta
 - Continua

- Estimación por Máxima Verosimilitud (Likelihood)
 - * Log-likelihood
 - Optimización
- Intervalos de Confianza
 - * Intervalo de Confianza para μ
 - * T-Student

Esperanza

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

$$E(\bar{X}_n) = \frac{1}{n} \times \sum_{i=1}^n E(X_i)$$

Si X y Y son variables aleatorias con esperanza finita y $a, b, c \in \mathbb{R}$ son constantes entonces

- $E[c] = c$
- $E[cX] = cE[X]$
- $E[X + Y] = E[X] + E[Y]$
- Si $X \geq 0$ entonces $E[X] \geq 0$
- Si $X \leq Y$ entonces $E[X] \leq E[Y]$
- Si X está delimitada por dos números reales, a y b , esto es $a < X < b$ entonces también lo está su media, es decir, $a < E[X] < b$
- Si $Y = a + bX$, entonces $E[Y] = E[a + bX] = a + bE[X]$
- Si X y Y son variables aleatorias independientes entonces $E[XY] = E[X]E[Y]$

Varianza

$$\text{Var}[X] = E[X^2] - E[X]^2 \implies E[X^2] = \text{Var}[X] + E[X]^2$$

Sean X y Y dos variables aleatorias con varianza finita y $a \in \mathbb{R}$

- $\text{Var}(X) \geq 0$
- $\text{Var}(a) = 0$
- $\text{Var}(aX) = a^2 \text{Var}(X)$
- $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$, donde $\text{Cov}(X, Y)$ denota la covarianza de X e Y
- $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ si X y Y son variables aleatorias independientes.
- $\text{Var}(Y) = E(\text{Var}(Y|X)) + \text{Var}(E(Y|X))$ cálculo de la Varianza por Pitágoras, donde $Y|X$ es la variable aleatoria condicional Y dado X .

Desvío Estándar

$$\text{SD}(X) = \sigma = \sqrt{\text{Var}(X)} \implies \sigma^2 = \text{Var}(X)$$

Covarianza

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

Correlación

$$\rho_{xy} = \frac{\text{cov}_{xy}}{\sigma_x \sigma_y} = \frac{\text{cov}_{xy}}{\text{SD}(x) \text{SD}(y)}$$

Continuas

$$F_x(x) = \int_{-\infty}^x f_X(u) du \quad f_X(x) = \frac{d}{dx} F_x(x)$$

$$P(a < X < b) = \int_a^b f(x)dx = F_x(b) - F_x(a)$$

$$P(a < X < b) = P(a < X \leq b) = P(a \leq X < b) = P(a \leq X \leq b)$$

Distribución Normal

Si $X \sim N(\mu, \sigma^2)$ y $a, b \in \mathbb{R}$, entonces $aX + b \sim N(a\mu + b, a^2\sigma^2)$

Si $X \sim N(\mu, \sigma^2)$, entonces $Z = \frac{X-\mu}{\sigma}$ es una variable aleatoria normal estándar: $Z \sim N(0, 1)$.

$$X \sim N(\mu, \sigma^2) \implies Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$$

$$Z \sim N(0, 1) \implies X = \sigma Z + \mu \sim N(\mu, \sigma^2)$$

$$X \sim N(\mu, \sigma^2)$$

$$X + b \sim N(\mu + b, \sigma^2)$$

$$aX \sim N(a \times \mu, a^2 \times \sigma^2)$$

$$\frac{X-\mu}{\sigma} \sim N(0, 1)$$

$$\bar{X}_n \sim N(\mu, \sigma^2/n), \text{ si } X_i \text{ son i.i.d}$$

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0, 1), \text{ si } X_i \text{ son i.i.d}$$

Función acumulada

$$f(x) = \phi(x) = \frac{1}{\sqrt{2 \times \pi \times \sigma^2}} \times e^{-\frac{(x-\mu)^2}{2 \times \sigma^2}}$$

Distribución Uniforme

$$P(a < X < b) = \frac{1}{b-a}$$

$$E(X) = \frac{a+b}{2}$$

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

Distribución Exponencial

$$f_X(X) = \lambda e^{-\lambda X}, \text{ para } X \geq 0$$

$$F_X(x) = P(X > x) = 1 - e^{-\lambda x}$$

$$E(X) = \frac{1}{\lambda}$$

$$\text{Var}(X) = \frac{1}{\lambda^2}$$

Discretas

Distribución Bernoulli

$$\begin{aligned}P(X = 1) &= p \\P(X = 0) &= 1 - p \\E(X) &= p \\Var(X) &= p(1 - p)\end{aligned}$$

Distribución Binomial

Distribución Poisson

Convergencia en Probabilidad

Sean X_n una secuencia de variables aleatorias, $X_n \xrightarrow{p} X$ si $\forall \epsilon > 0$

$\lim_{n \rightarrow \infty} P(|\bar{X}_n - E(X)| > \epsilon) = 0$, por Ley de los Grandes Números.

Propiedades

Si $X_n \xrightarrow{p} a$ y $Y_n \xrightarrow{p} b$, entonces:

- $X_n + Y_n \xrightarrow{p} a + b$
- $X_n Y_n \xrightarrow{p} a \cdot b$
- $\frac{X_n}{Y_n} \xrightarrow{p} \frac{a}{b}$ si $b \neq 0$
- $g(X_n) \xrightarrow{p} g(a)$ si g es una función continua

Estimación por LGN

Sean X_1, X_2, \dots, X_n variables aleatorias independientes e idénticamente distribuidas (i.i.d.) con esperanza μ y varianza σ^2 . Entonces, la media muestral $\bar{X}_n \xrightarrow{p} \mu$ por LGN. El estimador $\hat{\mu}_n = \bar{X}_n$ es consistente.

Estimación Esperanza

parámetro de interés: $\mu = E(X)$

muestra aleatoria: $X_1, \dots, X_n \sim f$, i.i.d

estimador: $\hat{\mu}_n = \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$

estimador consistente: $\hat{\mu}_n = \bar{X}_n \xrightarrow{p} \mu$

Estimación Varianza

parámetro de interés: $\sigma^2 = Var(X)$

muestra aleatoria: $X_1, \dots, X_n \sim f$, i.i.d

estimador: $\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$

estimador: $\hat{s}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$

Estimación Proporción

parámetro de interés: $p = P(X = 1)$

muestra aleatoria: $X_1, \dots, X_n \sim f$, i.i.d

estimador: $\hat{p}_n = \frac{1}{n} \sum_{i=1}^n X_i$

estimador consistente: $\bar{X}_n \xrightarrow{p} p$, por LGN

Estimación de Probabilidad

parámetro de interés: $p = F(x) = P(X \leq x)$

muestra aleatoria: $X_1, \dots, X_n \sim f$, i.i.d

definimos $Y_i \sim \text{Bernoulli}(p)$

$Y_i = X_i \leq x = \{1 \text{ si } X_i \leq x, 0 \text{ si } X_i > x\}$

estimador: $\hat{F}_n(x) = \bar{Y}_n = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{X_i \leq x}$

estimador consistente: $\hat{F}_n(x) = \bar{Y}_n \xrightarrow{p} F(x)$

Formulas Consistencia

Sesgo

$$\text{Sesgo}(\hat{\theta}_n) = E(\hat{\theta}_n) - \theta$$

Si $\text{Sesgo}(\hat{\theta}_n) = 0 \implies E(\hat{\theta}_n) = \theta$ entonces $\hat{\theta}_n$ es insesgado.

Asintóticamente Insesgado

$$\text{Sesgo}(\hat{\theta}_n) = E(\hat{\theta}_n) - \theta \xrightarrow{p} 0$$

Error Estándar

$$SE(\hat{\theta}_n) = \sqrt{\text{Var}(\hat{\theta}_n)}$$

$$\text{Var}(\hat{\theta}_n) = E[(\hat{\theta}_n - E(\hat{\theta}_n))^2]$$

Error Cuadrático Medio

$$ECME(\hat{\theta}_n) = \text{Sesgo}(\hat{\theta}_n)^2 + \text{Var}(\hat{\theta}_n)$$

$$ECME(\hat{\theta}_n) = E[(\hat{\theta}_n - \theta)^2]$$

Desigualdad de Chebyshev

$$P(|X - \mu| \geq \epsilon) \leq \frac{\text{Var}(X)}{\epsilon^2}$$

Desigualdad de Markov

$$P(X > \epsilon) \leq \frac{E(X)}{\epsilon}$$

Momentos

Momentos de una variable aleatoria

Discreta

$$m_k = E(X^k) = \sum_{i=1}^n (x_i - \bar{x})^k$$

Continua

$$m_k = \int_{\mathbb{R}} x^k f(x) dx$$

$$E(X) = \mu$$

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

$$E(X^3) = \mu^3 + 3\mu\sigma^2$$

$$E(X^4) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$$

Estimación por Máxima Verosimilitud (Likelihood)

$$\mathcal{L}(\theta; \underline{X}) = \prod_{i=1}^n f(\theta; x_i)$$

Log-likelihood

$$\ell(\theta; \underline{X}) = \ln(\mathcal{L}(\theta; \underline{X})) = \sum_{i=1}^n \ln(f(\theta; x_i))$$

Optimización

$$\frac{d\ell(\theta; \underline{X})}{d\theta} = \frac{d \ln(\mathcal{L}(\theta; \underline{X}))}{d\theta} = \sum_{i=1}^n \frac{d \ln(f(\theta; x_i))}{d\theta} = 0$$

Intervalos de Confianza

Intervalo de Confianza para μ

Sea la Variable Aleatoria $X_i \sim N(\mu, \sigma^2)$, nuestro parámetro de interés es μ . $\bar{X}_n \sim N(\mu, \sigma^2/n)$. Si σ^2 es conocido, entonces \bar{X}_n es insesgado y su varianza es σ^2/n .

Sea $Z = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$, $n = \bar{X}_n$

El intervalo de confianza $1 - \alpha$ es:

$$P(-z \leq Z \leq z) = 1 - \alpha$$

$$\phi(z) = P(Z \leq z) = 1 - \alpha/2$$

$$P(-z_{\alpha/2} \leq \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}) = 1 - \alpha$$

$$P(\hat{\mu}_n - z_{\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}} \leq \mu \leq \hat{\mu}_n + z_{\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}}) = 1 - \alpha$$

$$IC = (\hat{\mu}_n - z_{\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}}, \hat{\mu}_n + z_{\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}})$$

$$\bar{X}_n \pm z_{\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}}$$

T-Student

Si σ^2 es desconocido, entonces \bar{X}_n es insesgado y su varianza es s^2/n

Sea $T = \frac{\bar{X}_n - \mu}{\sqrt{\frac{s^2}{n}}} = \frac{\bar{X}_n - \mu}{s/\sqrt{n}} \sim t_{n-1}$, $n = \bar{X}_n$

El intervalo de confianza $1 - \alpha$ es:

$$\begin{aligned} P(-t_{n-1, \alpha/2} \leq T \leq t_{n-1, \alpha/2}) &= 1 - \alpha \\ P(\hat{\mu}_n - t_{n-1, \alpha/2} \sqrt{\frac{s^2}{n}} \leq \mu \leq \hat{\mu}_n + t_{n-1, \alpha/2} \sqrt{\frac{s^2}{n}}) &= 1 - \alpha \\ P(\hat{\mu}_n - t_{n-1, \alpha/2} \frac{s}{\sqrt{n}} \leq \mu \leq \hat{\mu}_n + t_{n-1, \alpha/2} \frac{s}{\sqrt{n}}) &= 1 - \alpha \\ \text{IC} &= (\hat{\mu}_n - t_{n-1, \alpha/2} \frac{s}{\sqrt{n}}, \hat{\mu}_n + t_{n-1, \alpha/2} \frac{s}{\sqrt{n}}) \\ &\quad \bar{X}_n \pm t_{n-1, \alpha/2} \frac{s}{\sqrt{n}} \end{aligned}$$