Resumen Inferencia Estadística

Introducción

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

Contenido

Esperanza

$\mathrm{E}(X) = \sum_{i=1}^{n} x_i \, \mathrm{F_x}(x_i)$
$\mathrm{E}(\overline{X}_n) = \frac{1}{n} \times \sum_{i=1}^{n} \mathrm{E}(X_i)$

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

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

Varianza

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

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

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

Desvío Estándar

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

Covarianza

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

Correlación

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

Continuas

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

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

$$\mathrm{P}(a < X < b) = \mathrm{P}(a < X \leq b) = \mathrm{P}(a \leq X < b) = \mathrm{P}(a \leq X \leq b)$$

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: $Z \sim N(0,1)$.

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

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

$$\begin{gathered} X \sim N(\mu, \sigma^2) \\ X + b \sim N(\mu + b, \sigma^2) \\ aX \sim N(a \times \mu, a^2 \times \sigma^2) \\ \frac{X - \mu}{\sigma} \sim N(0,1) \\ \overline{X}_n \sim N(\mu, \sigma^2/n), \text{ si } X_i \text{ son i.i.d} \\ \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1), \text{ si } X_i \text{ son i.i.d} \end{gathered}$$

Función acumulada

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

Distribución Uniforme

$\mathrm{P}(a < X < b) = \frac{1}{b-a}$
$\mathrm{E}(X) = \frac{a+b}{2}$
$\mathrm{Var}(X) = \frac{(b-a)^2}{12}$

Distribución Exponencial

$\mathrm{f_X}(X) = \lambda e^{-\lambda X}$, para $X \geq 0$
$\mathrm{F_X}(x) = \mathrm{P}(X > x) = 1 - e^{-\lambda x}$
$\mathrm{E}(X) = \frac{1}{\lambda}$
$\mathrm{Var}(X) = \frac{1}{\lambda^2}$

Discretas

Distribución Bernoulli

$\mathrm{P}(X = 1) = p$
$\mathrm{P}(X = 0) = 1 - p$
$\mathrm{E}(X) = p$
$\mathrm{Var}(X) = p(1 - p)$

Distribución Binomial

$\mathrm{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$
$\mathrm{E}(X) = np$
$\mathrm{Var}(X) = np(1-p)$

Distribución Poisson

$\mathrm{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$
$\mathrm{E}(X) = \lambda$
$\mathrm{Var}(X) = \lambda$

Convergencia en Probabilidad

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

$\lim_{n\to\infty} \mathrm{P}\left(|\overline{X}_n - \mathrm{E}(X)| > \epsilon\right) = 0$, 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 $b \neq 0$
- $\mathrm{g}(X_n) \xrightarrow{p} \mathrm{g}(a)$ si g es una función continua

Estimación por LGN

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

Estimación Esperanza

parámetro de interés: $\mu = \mathrm{E}(X)$
muestra aleatoria: $X_1, \ldots, X_n \sim f$, i.i.d
estimador: $\hat{\mu}_n = \overline{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$
estimador consistente: $\hat{\mu}_n = \overline{X}_n \xrightarrow{p} \mu$

Estimación Varianza

parámetro de interés: $\sigma^2 = \mathrm{Var}(X)$
muestra aleatoria: $X_1, \ldots, X_n \sim f$, i.i.d
estimador: $\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$
estimador: $\hat{s}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$

Estimación Proporción

parámetro de interés: $p = \mathrm{P}(X = 1)$
muestra aleatoria: $X_1, \ldots, X_n \sim f$, i.i.d
estimador: $\hat{p}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$
estimador consistente: $\overline{X}_n \xrightarrow{p} p$, por LGN

Estimación de Probabilidad

parámetro de interés: $p = \mathrm{F}(x) = \mathrm{P}(X \leq x)$
muestra aleatoria: $X_1, \ldots, X_n \sim f$, i.i.d
definimos $Y_i \sim \mathrm{Bernoulli}(p)$
$Y_i = X_i \leq x = \{1 \text{ si } X_i \leq x, 0 \text{ si } X_i > x$
estimador: $\hat{F}_n(x) = \overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} 1_{X_i \leq x}$
estimador consistente: $\hat{F}_n(x) = \overline{Y}_n \xrightarrow{p} \mathrm{F}(x)$

Formulas Consistencia

Sesgo

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

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

Asintóticamente Insesgado

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

Error Estándar

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

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

Error Cuadrático Medio

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

Desigualdad de Chebyshev

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

Desigualdad de Markov

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

Momentos

Momentos de una variable aleatoria

Discreta

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

Continua

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

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

Estimación por Máxima Verosimilitud (Likelihood)

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

Log-likelihood

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

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

Intervalos de Confianza

Intervalo de Confianza para μ

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

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

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

$\mathrm{P}(-z \leq Z \leq z) = 1-\alpha$
$\phi(z) = \mathrm{P}(Z \leq z) = 1-\alpha/2$
$\mathrm{P}(-z_{\alpha/2} \leq \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2}) = 1-\alpha$
$P(\hat{\mu}_n - z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}} \leq \mu \leq \hat{\mu}_n + z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}) = 1-\alpha$
$\mathrm{IC} = (\hat{\mu}_n - z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}, \hat{\mu}_n + z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}})$
$\overline{X}_n \pm z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{n}}$

T-Student

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

Sea $T = \frac{\overline{X}_n - \mu}{\sqrt{\frac{s^2}{n}}} = \frac{\overline{X}_n - \mu}{s/\sqrt{n}} \sim \mathrm{t}_{n-1}$, $\hat{\mu}_n = \overline{X}_n$

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

$\mathrm{P}(-t_{n-1,\alpha/2} \leq T \leq t_{n-1,\alpha/2}) = 1-\alpha$
$\mathrm{P}(\hat{\mu}_n - t_{n-1,\alpha/2}\sqrt{\frac{s^2}{n}} \leq \mu \leq \hat{\mu}_n + t_{n-1,\alpha/2}\sqrt{\frac{s^2}{n}}) = 1-\alpha$
$\mathrm{P}(\hat{\mu}_n - t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \hat{\mu}_n + t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}) = 1-\alpha$
$\mathrm{IC} = (\hat{\mu}_n - t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}, \hat{\mu}_n + t_{n-1,\alpha/2}\frac{s}{\sqrt{n}})$
$\overline{X}_n \pm t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}$