ESTADÍSTICA II

Profesor: Pablo Fdez Gallardo
pablo.fernandez@uam.es

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

$cov(X,Y) = E((X-E(X))(Y-E(Y))) = E(XY) - E(X)E(Y)$

$cov(\mathbb{X}) = E((\mathbb{X}-E(\mathbb{X}))(\mathbb{X}-E(\mathbb{X}))^T) = \begin{pmatrix} V(X_1) & cov(X_1,X_2) & \vdots \\ cov(X_2,X_1) & V(X_2) & \vdots \\ \cdots & \cdots & \end{pmatrix}$

$cov(A\mathbb{X}+b) = A\,cov(\mathbb{X})A^T$

$\rho(X,Y) = \frac{cov(X,Y)}{\sqrt{V(X)V(Y)}} \in [-1,1]$ $\quad \rho(\mathbb{X}) = \begin{pmatrix} 1 & \rho(X_1X_2) & \vdots \\ \rho(X_2,X_1) & 1 & \vdots \\ \cdots & \cdots & \end{pmatrix}$

$cov(\mathbb{X}) = \sqrt{D(\mathbb{X})}\,\rho(\mathbb{X})\sqrt{D(\mathbb{X})}$ $\quad \boxed{V(X+Y) = V(X)+V(Y)+2cov(X,Y)}$

V sim y def. pos. $\Rightarrow V = UU^T$, U^{-1} raíz cuad. V^{-1}

$f_{\mathbb{X}}(x) = \frac{1}{(2\pi)^{n/2}} \cdot \frac{1}{\sqrt{\det(V)}} \exp\left(\frac{-1}{2}(x-\mu)^T V^{-1}(x-\mu)\right)$

$f_{\mathbb{X}}(x) = \frac{1}{(2\pi)^{n/2}} \cdot \frac{1}{|\det(U)|} \exp\left(\frac{-1}{2}\|U^{-1}(x-\mu)\|^2\right)$

$\mathbb{X} \sim N_n(\mu, V) \Leftrightarrow \mathbb{X} = \mu + U\mathbb{Y}$ con $\mathbb{Y} \sim N(\vec{0}, I_n)$

e.d. $U^{-1}(\mathbb{X}-\mu) \sim N(\vec{0}, I_n)$

$\mathbb{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \\ \hline x_{p+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \mathbb{X}_1 \\ \hline \mathbb{X}_2 \end{pmatrix}$; $\mu = \begin{pmatrix} \mu_1 \\ \hline \mu_2 \end{pmatrix}$; $V = \left(\begin{array}{c|c} V_{11} & V_{12} \\ \hline V_{21} & V_{22} \end{array}\right)$

$\mathbb{X}_1, \mathbb{X}_2$ indep. $\Leftrightarrow V_{12} = V_{21}^T = \vec{\vec{0}}$ (todo ceros)

$\mathbb{X}_1 | \mathbb{X}_2 = a \sim N_p(\tilde{\mu}, \tilde{V})$ con:

$\tilde{\mu} = \mu_1 + V_{12}V_{22}^{-1}(a-\mu_2)$

$\tilde{V} = V_{11} - V_{12}V_{22}^{-1}V_{21}$

Si $\mathbb{X} \sim N_n(\mu, V)$, $a \in \mathbb{R}^n$, $A \in \mathcal{M}_{p\times n}$ $p<n$:

$a^T\mathbb{X} \sim N(a^T\mu, a^TVa)$, $A\mathbb{X} \sim N(A\mu, AVA^T)$

$b \in \mathbb{R}^n$, $A\mathbb{X}+b \sim N(A\mu+b, AVA^T)$

$Z \sim \chi_n^2 \Rightarrow E(Z) = n$, $V(Z) = 2n$

Si $\mathbb{X} \sim N_n(0, I_n) \Rightarrow \mathbb{X}^T\mathbb{X} \sim \chi_n^2$

Si $\mathbb{X} \sim N_n(\mu, V) \Rightarrow (\mathbb{X}-\mu)^T V^{-1}(\mathbb{X}-\mu) \sim \chi_n^2$

Tma 1: $\mathbb{X} \sim N(0, I)$, B sim + idemp. $\Rightarrow \mathbb{X}^TB\mathbb{X} \sim \chi^2_{tr(B)}$

Coro 2: $\mathbb{X} \sim N(\mu, \sigma^2 I)$, B sim + idemp.,

si $\mu^TB\mu = 0 \Rightarrow \frac{1}{\sigma^2}\mathbb{X}^TB\mathbb{X} \sim \chi^2_{traza(B)}$

Tma 3: $\mathbb{X} \sim N_n(\mu, \sigma^2 I_n)$, $A \in \mathcal{M}_{p\times n}$ $(p<n)$,

$B, C \in \mathcal{M}_{n\times n}$ sim. e idemp. $\Rightarrow$

$\Rightarrow \begin{cases} \text{Si } AB = 0 \Rightarrow A\mathbb{X}, \mathbb{X}^TB\mathbb{X} \text{ indep.} \\ \text{Si } BC = 0 \Rightarrow \mathbb{X}^TB\mathbb{X}, \mathbb{X}^TC\mathbb{X} \text{ indep.} \end{cases}$

Tma 4: $\mathbb{X} \sim N_n(\mu, \sigma^2 I_n)$, $A \in \mathcal{M}_{p\times n}$ rang p,

$B \in \mathcal{M}_{q\times n}$ rang q, $p+q \leq n$.

Si $AB^T = 0_{p\times q} \Rightarrow A\mathbb{X}$ y $B\mathbb{X}$ indep.

K-S 1 MUESTRA: X cont. $\Rightarrow \Delta_n^X = \Delta_n^{UNIF[0,1]}$

$\Delta_1 = \max\{|F(x_1)|, |1-F(x_1)|\}$, $Z = \max\{U, 1-U\}$

$z_0 \in (0,1)$: $\mathbb{P}(Z \leq z_0) = \mathbb{P}(\max\{U, 1-U\} \leq z_0) =$

$= \mathbb{P}(U \leq z_0, 1-U \leq z_0) = \mathbb{P}(1-z_0 \leq U \leq z_0) =$

$= \begin{cases} 2z-1, & \text{si } z \in [\frac{1}{2}, 1] \\ 0 & \text{otherwise} \end{cases} \Rightarrow Z \sim UNIF[\frac{1}{2}, 1]$

Sabemos que $\sqrt{n}\Delta_n^+ \xrightarrow{d} Z$, $F_Z(z) = 1-e^{-2z^2}$ $(z>0)$

$W_n = 4n(\Delta_n^+)^2$. $\omega > 0$, $\mathbb{P}(W_n \leq \omega) =$

$= \mathbb{P}(4n(\Delta_n^+)^2 \leq \omega) = \mathbb{P}(2\sqrt{n}\,\Delta_n^+ \leq \sqrt{\omega}) =$

$= \mathbb{P}(\sqrt{n}\Delta_n^+ \leq \frac{\sqrt{\omega}}{2}) \xrightarrow{d} 1-e^{-2\omega/4} = 1-e^{-\omega/2}$

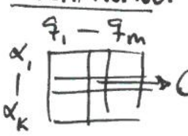
TEST χ^2 PEARSON $\begin{cases} \alpha_1, \dots, \alpha_K \to \text{valores} \\ p_1, \dots, p_K \to \text{probs} \end{cases}$ observados: $O_1, \dots, O_K$

$n = \sum O_j$

$b = \sum \frac{(O_i - E_i)^2}{E_i}$ $\quad R = \{b > \chi^2_{\{K-1;\alpha\}}\}$ esperados: $E_1, \dots, E_K = np_1, \dots, np_K$

INDEP./HOMOG. $X \to \begin{cases} \alpha_1, \dots, \alpha_K \to \text{vals} \\ p_1, \dots, p_K \to \text{probs} \end{cases}$ $\quad Y \to \begin{cases} \beta_1, \dots, \beta_m \to \text{vals} \\ q_1, \dots, q_m \to \text{probs} \end{cases}$

$\hookrightarrow$ pueden conocerse o no



$E_{ij} = np_iq_j$ $\quad R = \{b > \chi^2_{\{Km-1-(K-1)-(m-1)\}}$

$n = \sum\sum O_{ij}$ $\quad$ OJO SI SE ESTIMAN PARAMS $\quad$ estim.

K-S: $F_n(t) = \frac{1}{n}\#\{1 \leq i \leq n : x_i \leq t\}$ $\quad x_1, \dots, x_n$ muestra ordenada

$\Delta_n = \sup_{t\in\mathbb{R}}|F_n(t) - F_X(t)| = \max_{1\leq i\leq n}\left[\max\left\{\left|F_X(x_i) - \frac{i-1}{n}\right|, \left|F_X(x_i) - \frac{i}{n}\right|\right\}\right]$

para todo $n \geq 1$, Δ_n se distribuye igual $\forall X$ v.a. continua

Si X cont. $\Rightarrow \Delta_n^X \overset{d}{=} \Delta_n^{unif(0,1)}$.

Glivenko-Cantelli: $\forall X$ v.a. $\Delta_n \longrightarrow 0$ c.s $n\to\infty$ ($\Rightarrow$ en prob.)

K-S: $\sqrt{n}\Delta_n \xrightarrow[n\to\infty]{d} Z$ (cierta v.a. que no depende de X)

Obs: $\Delta_n^+ = \sup_{t\in\mathbb{R}}(F_n(t) - F_X(t))^+ \to \lim_{n\to\infty}\mathbb{P}(\sqrt{n}\Delta_n^+ \leq x) = 1-e^{-2x^2}$ $(x>0)$

$R = \{\sqrt{n}\,\delta_n > \text{percentil } \alpha \text{ de } Z \text{ de K-S}\}$

RL-SIMPLE-EXTRA $\quad |R| = \frac{|cov_{X,Y}|}{\sqrt{V_XV_Y}}$; $\frac{|\hat{\beta}_1|}{S_R\sqrt{1/nV_X}} = \frac{\sqrt{n-2}\cdot|R|}{\sqrt{1-R^2}}$

$EMC = \frac{1}{n}\sum e_i^2 = \frac{1}{n}\sum(y_i - \hat{y}_i)^2$

$a \in \mathbb{R}^K$: $f(a) = \frac{|a^T(\mu_0-\mu_1)|^2}{a^T\Sigma a}$ sea máximo $\Rightarrow a_m = \Sigma^{-1}(\mu_0-\mu_1) \Rightarrow$

$\Rightarrow \hat{a}_m = S_p^{-1}(\bar{x}_0 - \bar{x}_1)$; $S_p = \frac{(n_0-1)S_0 + (n_1-1)S_1}{n_0+n_1-2}$; $S_0 = \frac{1}{n_0-1}\sum_{j=1}^{n_0}(x_j - \bar{x}_0)(x_j - \bar{x}_0)^T$ fila

PROBIT: $p(x) = h(\beta^T\tilde{x}) = \Phi(\beta^T\tilde{x}) \Rightarrow \beta^T\tilde{x} = \Phi^{-1}(p(x))$

LOGIT: $p(x) = h(\beta^T\tilde{x}) = \frac{1}{1+\exp(-\beta^T\tilde{x})}$; $\mathcal{O}(x) = \frac{p(x)}{1-p(x)} = \exp(\beta^T\tilde{x})$

Sea $\hat{p}(x) = h(\hat{\beta}^T\tilde{x})$, $\hat{\beta} \sim N(\beta, (X^T\hat{W}X)^{-1})$ $\quad \begin{pmatrix} \hat{p}(x_1)(1-\hat{p}(x_1)) & & 0 \\ & \ddots & \\ 0 & & \hat{p}(x_n)(1-\hat{p}(x_n)) \end{pmatrix}$

$IC_{1-\alpha}(\beta_j) = \hat{\beta}_j \pm Z_{\alpha/2}\,SE(\hat{\beta}_j)$

$\sqrt{\text{elem}(j+1)}$ diagonal

REGLAS ÓPTIMAS CLASIF.

$R_1 = \{x \in \Omega : P_0f_0 - P_1f_1 \leq 0\} = \{x \in \Omega : P_1f_1 \geq P_0f_0\}$

$\mathbb{P}(\text{"mala clasif"}) = P_1\int_{R_0} f_1(\vec{x})d\vec{x} + P_0\int_{R_1} f_0(\vec{x})d\vec{x}$

Coste medio mala clasif. $= C_{0|1}P_1\int_{R_0} f_1 + C_{1|0}P_0\int_{R_1} f_0 \Rightarrow R_1 = \{x \in \Omega : C_{0|1}P_1f_1 \geq C_{1|0}P_0f_0\}$

$\mathbb{Y} = \mathbb{X}\beta + \varepsilon$, n obs, K var.

Distr. $\hat{\mathbb{Y}} = \mathbb{X}\hat{\beta}$.

Si $\mathbb{Y} \sim N(\mathbb{X}\beta, \sigma^2I_n) \Rightarrow$

$\Rightarrow \hat{\beta} \sim N(\beta, \sigma^2(X^TX)^{-1})$.

$\mathbb{E}(\hat{\mathbb{Y}}) = \mathbb{X}E(\hat{\beta}) = \mathbb{X}\beta$

$\mathbb{V}(\hat{\mathbb{Y}}) = \mathbb{X}\mathbb{V}(\hat{\beta})\mathbb{X} = \sigma^2H$

$\Rightarrow \hat{\mathbb{Y}} \sim N(\mathbb{X}\beta, \sigma^2H)$

$\sum V(\hat{Y}_i) = \sigma^2 tr(H) = \sigma^2(K+1)$

$\underbrace{\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}}_{\mathbb{Y}} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}}_{\mathbb{X}} \underbrace{\begin{pmatrix} \mu \\ \lambda \end{pmatrix}}_{\beta} + \underbrace{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}}_{\varepsilon \sim N(\vec{0}, \sigma^2 I_3)}$

$\hat{\beta} = (X^TX)^{-1}X^T\mathbb{Y} = \frac{1}{3}\begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix}$

$\hat{\beta} = A\mathbb{Y} \Rightarrow \hat{\beta} \sim N(A\mathbb{E}(\mathbb{Y}), A\,cov(\mathbb{Y})A^T)$

$\Pi_0 \to f_0 = 1-|x|$, $|x| \leq 1$

$\Pi_1 \to f_1 = 1-|x-\frac{1}{2}|$, $x \in [\frac{-1}{2}, \frac{3}{2}]$

$P_1 = 0'8$, $P_0 = 0'2$

$R_1 = \{x \in \mathbb{R} : P_1f_1 \geq P_0f_0\}$

$4f_1 \geq f_0$. Miramos $4f_1 = f_0$

$\Rightarrow R_0 = [-1, \frac{1}{3})$, $R_1 = [\frac{1}{3}, \frac{3}{2}]$

$P_0 = P_1 = 0'5$: prob. mala clf $=$

$= P_1\int_{R_0} f_1 + P_0\int_{R_1} f_0$

Ej. Fisher recuerdo:

$S_p = \frac{(n_0-1)S_0 + (n_1-1)S_1}{n_0+n_1-2}$

$\hat{a}_m = S_p^{-1}(\bar{x}_0 - \bar{x}_1)$

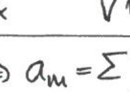
1) Proy $\bar{x}_0$: $\hat{a}_m^T\bar{x}_0$

2) Proy $\bar{x}_1$: $\hat{a}_m^T\bar{x}_1$

3) Proy x : $\hat{a}_m^Tx$

...iar más cercano

$H_0: \beta_j \geq 0 \Rightarrow R = \{\hat{\beta}_j ...$



$\frac{\alpha}{2} = 1 - F_X(b)$

p-valor $\equiv \alpha$

$H = h_{ij}$; H sim. def. pos. idemp.

1) $\sum_i^n h_{ii} = K+1$ (rango)

2) $0 < h_{ii} \leq 1$

3) $h_{ii} = h_{ii}^2 + \sum_{i\neq j} h_{ij}^2$

4) $h_{ij}^2 \leq \frac{1}{4}$ $i \neq j$

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

$\Rightarrow \mathbb{P}\left(\frac{X-\mu}{\sigma} \leq a\right) = \Phi(a)$

$\mathbb{X} \sim N_n(\vec{\mu}, \Sigma) \Rightarrow$

$\Rightarrow \mathbb{P}(U^{-1}(\mathbb{X}-\vec{\mu}) \leq a) =$

$= \Phi_n(a)$