1. Models de regressió lineal

1.1. Model general (mínims quadrats)

$$y = X\beta + \varepsilon$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_{11} & \dots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{nm} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

- $\boldsymbol{y}$: Vector de resposta
- $\boldsymbol{X}$: Matriu de disseny
  - $x_{ij} \in \{0, 1\}$
  - $x_{i1} = 1$ (*intercept*)
- $\boldsymbol{\varepsilon}$: Errors aleatoris (*i.i.d.*)
  - $\mathbb{E}(\varepsilon) = 0$
  - $\mathbb{V}(\varepsilon) = \sigma^2 I_n$

» Incògnites β

Si $\text{Rang}(X) = m$:

$$\widehat{\beta} = \boxed{(X^T X)^{-1} X^T y}$$

- $\mathbb{E}(\widehat{\beta}) = \beta$
- $\mathbb{V}(\widehat{\beta}) = \sigma^2 \, (X^T X)^{-1}$

» Variància de l'error σ^2

$$\widehat{\sigma}^2 = \boxed{\frac{RSS}{n-m}}$$

1.2. Model clàssic

Tenim $\boxed{\varepsilon \sim N(0, \sigma^2 I)}$.

» Incògnites β

Si $\text{Rang}(X) = m$:

$$l^T \widehat{\beta} \sim \boxed{N(l^T\beta,\, \sigma^2\, l^T (X^T X)^{-1}\, l)}$$

» Variància de l'error σ^2

Si $\text{Rang}(X) = m$:

$$\boxed{\frac{RSS}{\sigma^2} \sim \chi^2_m}$$

Residus estandaritzats

$$rst_i = \frac{r_i}{\widehat{\sigma}\,\sqrt{1 - p_{ii}}} \sim N(0, 1)$$

- r_i element del residu $r = Y - \widehat{Y} = (1 - P)Y$
- p_{ii} element de $P = X\,(X^T X)^{-1} X^T$

Outliers

$|rst_i| > 2 \implies$ potencial *outlier*

1.3. Model amb errors dependents

Tenim $\boxed{\varepsilon_i = \rho\,\varepsilon_i + e_i}$ amb ρ conegut tal que:

- $\mathrm{E}(e_i) = 0$
- $\mathrm{Var}(e_i) = \sigma^2$

Mètode

Construir el nou model (clàssic):

$$\boxed{y_i^* = \beta_0(1-\rho) + \beta_1 x_i^* + e_i}$$

amb $\begin{cases} y_i^* = y_i - \rho\, y_{i-1} \\ x_i^* = x_i - \rho\, x_{i-1} \end{cases}$

1.4. Model dels **mínims quadrats generalitzat**

Tenim $\boxed{\mathrm{Var}(\varepsilon) = \sigma^2\Sigma = \sigma^2 P^2}$ amb Σ conegut.

Mètode

Construir el nou model (clàssic):

$$\boxed{\begin{gathered} P^{-1}Y = P^{-1}X\beta + P^{-1}\varepsilon \\ y^* = X^*\beta + \varepsilon^* \end{gathered}}$$

- $\widehat{\beta} = \boxed{(X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}y}$

2. Bondat d'ajustament

Coeficient de determinació

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \in (0,1)$$

Coeficient de correlació

> V.A.s A, B amb mitjanes μ_A, μ_B i variàncies σ_A^2, σ_B^2

$$\rho_{AB} = \boxed{\frac{\mathbb{E}((A-\mu_A)(B-\mu_B))}{\sigma_A\,\sigma_B}} \in (-1,1)$$

Coeficient de correlació mostral

> Observacions $a_1, \ldots, a_n$ i $b_1, \ldots, b_n$ amb mitjanes mostrals $\overline{a}, \overline{b}$, i variàncies mostrals s_a^2, s_b^2

$$r_{AB} = \boxed{\frac{\sum_{i=1}^n (a_i - \overline{a})(b_i - \overline{b})}{s_a\,s_b}} \in (-1,1)$$

3. Tests d'hipòtesis

Test t generalitzat (1 coeficient)

- Hipòtesis nul·la $\boldsymbol{H_0} : l^T\beta = 0$
- Estadístic de contrast $\boxed{\boldsymbol{T} = \frac{\varphi^T y}{\widehat{\sigma}\sqrt{\varphi^T\varphi}} \sim t_{n-m}}$ amb

  $\boldsymbol{\varphi^T} = l^T(X^TX)^{-1}X^T$
- Rebutjem H_0 si $|T| > t_{1-\frac{\alpha}{2},\,n-m}$

Interval de confiança

$$IC_\alpha(l^T\beta) = \left(l^T\widehat{\beta} \pm t_{1-\frac{\alpha}{2},\,n-m}\,\widehat{\sigma}\,\sqrt{l^T\,(X^TX)^{-1}\,l}\right)$$

» Test F generalitzat (múltiples coeficients)

- Hipòtesis nul·la $H_0 : L\beta = 0$
- Estadístic de contrast

$$F = \boxed{\frac{(L\widehat{\beta})^T \, (L(X^T X)^{-1} \, L^T)^{-1} \, (L\widehat{\beta}) \, (n-m)}{q \cdot RSS}} \sim F_{q,\, n-m}$$

amb q nombre de files de L

- Rebutjem H_0 si $F > F_{1-\alpha,\, q,\, n-m}$

4. Observacions influents

» Mesura de la influència

Influència de l'element i-èssim en el model:

$$D_i = \boxed{\frac{(\widehat{\beta} - \widehat{\beta}_{\bar{i}})^T X^T X \, (\widehat{\beta} - \widehat{\beta}_{\bar{i}})}{(p+1) \, \widehat{\sigma}^2}}$$

- $\widehat{\beta}_{\bar{i}}$: estimador de β sense la i-èssima observació

Interpretació de D

Observació influent si:

- $D_i > 1$
- $D_i > \frac{4}{n}$

» *Leverage*

Leverage de la observació i-èssima del model:

element p_{ii} de $P = X\,(X^T X)^{-1} X^T$

- $p_{ii} > \frac{2m}{n} \implies$ possible influència excessiva

5. Estabilització de la variància

Tenim *V.A.* X amb:

- $\mathrm{E}(X) = \mu$
- $\mathrm{Var}(X) = k\,h(\mu)$

Transformacions estabilitzants de variància

Construir una nova *V.A.*:

$$\boxed{Y = g(X)}$$

- $\mathrm{Var}(Y) \sim k\,h(\mu)\,g'(\mu)^2$ (constant)
- $g(\mu) = \displaystyle\int \frac{C}{\sqrt{h(\mu)}}\,d\mu$

Distribució	Variància	Transformació
Poisson	$\propto \mu$	$g(X) = \sqrt{X}$ (*)
Binomial	$\propto \mu\,(1-\mu)$	$g(X) = \arcsin(\sqrt{X})$
Exponencial	$\propto \mu^2$	$g(X) = \ln(X)$
desconeguda (*Box-Cox*)	desconeguda	$g_\lambda(x_i) = \begin{cases} \dfrac{x_i^\lambda - 1}{\lambda} & \text{si } \lambda \neq 0 \\ \log(x_i) & \text{si } \lambda = 0 \end{cases}$

(*) Per n petita, podem fer servir $g(X) = \sqrt{X + \frac{3}{8}}$.

Apèndix

- Estimador $\widehat{\alpha}$ no esbiaixat si $\mathrm{E}(\widehat{\alpha}) = \alpha$

- **Covariància** $\mathrm{Cov}(a,b) = \boxed{\mathrm{E}(ab) - \mathrm{E}(a)\,\mathrm{E}(b)}$
- $\hat{y} = X\widehat{\beta}$
- $TSS = ESS + RSS$

Sum of Squares

- *Total Sum of Squares* (variabilitat de y)

$$\boldsymbol{TSS} = \sum_{i=1}^{n} (y_i - \overline{y})^2$$

- *Explained Sum of Squares* (variabilitat de y explicada pel model)

$$\boldsymbol{ESS} = \sum_{i=1}^{n} (\widehat{y}_i - \overline{y})^2$$

- *Residual Sum of Squares* (variabilitat de y **no** explicada pel model):

$$\boldsymbol{RSS} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Propietats de vectors aleatoris

> $\boldsymbol{y} = (\boldsymbol{y_1}, \dots, \boldsymbol{y_n})$ vector aleatori
> $\boldsymbol{b}$ vector de constants
> $\boldsymbol{A}, \boldsymbol{B}$ matrius de constants

Vector	Esperança ($\mathbb{E}$)	Variància ($\mathbb{V}$)
y	$\begin{pmatrix} \mathrm{E}(y_1) \\ \vdots \\ \mathrm{E}(y_n) \end{pmatrix}$	$\begin{pmatrix} \mathrm{Var}(y_1) & \mathrm{Cov}(y_1, y_2) & \dots & \mathrm{Cov}(y_1, y_n) \\ \mathrm{Cov}(y_2, y_1) & \mathrm{Var}(y_2) & \dots & \mathrm{Cov}(y_2, y_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(y_n, y_1) & \mathrm{Cov}(y_n, y_2) & \dots & \mathrm{Var}(y_n) \end{pmatrix}$
$A\,y + b$	$A \cdot \mathbb{E}(y) + b$	$A \cdot \mathbb{V}(y) \cdot A^T$

- $\operatorname{Cov}(Ay, By) = A \cdot \mathbb{V}(y) \cdot B^T$