Quando uma variável resposta Y é função de duas ou mais variáveis regressoras $(X_1, X_2, \dots, X_K)$ temos uma REGRESSÃO LINEAR MÚLTIPLA.

O modelo estatístico é dado por:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_K X_{iK} + \varepsilon_i$$

com: $i = 1, \dots, n$ e $j = 1, \dots, K$.

Considerando cada variável X_{ij} centrada em sua média $\bar{X}_j$, temos: $x_{ij} = X_{ij} - \bar{X}_j \Rightarrow X_{ij} = x_{ij} + \bar{X}_j$

Desta forma,

$$Y_i = \beta_0 + \beta_1(x_{i1} + \bar{X}_1) + \beta_2(x_{i2} + \bar{X}_2) + \dots + \beta_K(x_{iK} + \bar{X}_K) + \varepsilon_i$$

$$Y_i = \underbrace{(\beta_0 + \beta_1 \bar{X}_1 + \dots + \beta_K \bar{X}_K)}_{\alpha} + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_K x_{iK} + \varepsilon_i$$

Matricialmente, o modelo RLM é dado por:

$$\underset{\sim}{Y} = \mathbb{X}\underset{\sim}{\theta} + \underset{\sim}{\varepsilon}$$

$\underset{\sim}{Y}$ é um vetor $n \times 1$ da variável aleatória Y.

$$\underset{\sim}{Y} = {}_{n}\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{1}$$

$\mathbb{X}$ é a matriz $n \times p$ de delineamento, $p = K+1$

$$\mathbb{X} = {}_{n}\begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1K} \\ 1 & X_{21} & X_{22} & \cdots & X_{2K} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nK} \end{bmatrix}_{p=K+1}$$

$\underset{\sim}{\theta}$ é o vetor de parâmetros $p \times 1$, $p = k+1$

$$\underset{\sim}{\theta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}_{p \times 1}$$

$\underset{\sim}{\varepsilon}$ é o vetor $n \times 1$ de erros

$$\underset{\sim}{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{n \times 1}$$

Suposições:

i) A variável y é uma função LINEAR das k variáveis regressoras;

ii) As variáveis x_j são de natureza fixa;

iii) $E(\varepsilon_i) = 0 \Rightarrow E(\underset{\sim}{\varepsilon}) = \phi = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}_{n \times 1}$

iv) Erros são homocedásticos e independentes

$$\left.\begin{array}{l} Var(\varepsilon_i) = \sigma^2 \\ Cov(\varepsilon_i, \varepsilon_{i'}) = 0 \end{array}\right\} Var(\underset{\sim}{\varepsilon}) = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}_{n \times n} = I_n \sigma^2 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{n \times n} \sigma^2$$

v) Erros têm distribuição NORMAL

$$\varepsilon_i \sim N(0, \sigma^2) \Rightarrow \underset{\sim}{\varepsilon} \sim N(\phi, I\sigma^2)$$

$$y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2) \Rightarrow \underset{\sim}{Y} \sim N(X\underset{\sim}{\theta}, I\sigma^2)$$

ESTIMAÇÃO DOS PARÂMETROS – MMQ $(n > p)$

Considerando o modelo: $\underset{\sim}{Y} = \mathbb{X}\underset{\sim}{\theta} + \underset{\sim}{\varepsilon}$, temos:

$$\underset{\sim}{\varepsilon} = \underset{\sim}{Y} - \mathbb{X}\underset{\sim}{\theta} \Rightarrow \text{minimizar } \sum_{i=1}^{n} \varepsilon_i = \underset{\sim}{\varepsilon}^T \underset{\sim}{\varepsilon} = [\varepsilon_1, \dots \varepsilon_n]\begin{bmatrix}\varepsilon_1\\ \vdots \\ \varepsilon_n\end{bmatrix}$$

$$\underset{\sim}{\varepsilon}^T\underset{\sim}{\varepsilon} = [\underset{\sim}{Y} - \mathbb{X}\underset{\sim}{\theta}]^T[\underset{\sim}{Y} - \mathbb{X}\underset{\sim}{\theta}]$$

OBS: $[a_{11}\ a_{12}\ a_{13}]\begin{bmatrix}a_{11}\\ a_{21}\\ a_{31}\end{bmatrix} = a_{11}.a_{11} + a_{12}.a_{21} + a_{13}.a_{31}$

$$Z = [\underset{\sim}{Y}^T - \underset{\sim}{\theta}^T \mathbb{X}^T][\underset{\sim}{Y} - \mathbb{X}\underset{\sim}{\theta}]$$

$(A+B)^T = A^T + B^T$

$(AB)^T = B^T A^T$

$$Z = \underset{\sim}{Y}^T\underset{\sim}{Y} - \underbrace{\underset{\sim}{Y}^T\mathbb{X}}_{A^T}\underset{\sim}{\theta} - \underset{\sim}{\theta}^T\underbrace{\mathbb{X}^T\underset{\sim}{Y}}_{A} + \underset{\sim}{\theta}^T\underbrace{\mathbb{X}^T\mathbb{X}}_{B}\underset{\sim}{\theta}$$

$$\frac{\partial A^T\underset{\sim}{\theta}}{\partial\underset{\sim}{\theta}} = \frac{\partial\underset{\sim}{\theta}^T A}{\partial\underset{\sim}{\theta}} = A$$

$$\frac{\partial\underset{\sim}{\theta}^T B\underset{\sim}{\theta}}{\partial\underset{\sim}{\theta}} = 2B\underset{\sim}{\theta}$$

$$\frac{\partial Z}{\partial\underset{\sim}{\theta}} = 0$$

$\hookrightarrow -A - A + 2B\underset{\sim}{\theta} = 0 \quad \begin{cases} A = X^T\underset{\sim}{Y} \\ B = X^TX \end{cases}$

$$-\mathbb{X}^T\underset{\sim}{Y} - \mathbb{X}^T\underset{\sim}{Y} + 2\mathbb{X}^T\mathbb{X}\hat{\underset{\sim}{\theta}} = 0$$

$$-2\mathbb{X}^T\underset{\sim}{Y} + 2\mathbb{X}^T\mathbb{X}\hat{\underset{\sim}{\theta}} = 0$$

$$\mathbb{X}^T\mathbb{X}\hat{\underset{\sim}{\theta}} = \mathbb{X}^T\underset{\sim}{Y} \quad \text{(SEN)}$$

$$\underbrace{(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\mathbb{X}}_{I}\hat{\underset{\sim}{\theta}} = (\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\underset{\sim}{Y}$$

$$\boxed{\hat{\underset{\sim}{\theta}} = (\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\underset{\sim}{Y}}$$

QUEM SÃO AS MATRIZES ?

Considerando n observações e duas variáveis X_1 e X_2, temos:

$$\underset{\sim}{Y} = {}_n\begin{bmatrix}Y_1\\ \vdots\\ Y_n\end{bmatrix}_1 \qquad \mathbb{X} = {}_n\begin{bmatrix}1 & X_{11} & X_{12}\\ \vdots & \vdots & \vdots\\ 1 & X_{n1} & X_{n2}\end{bmatrix}_{p=3} \qquad \mathbb{X}^T = {}_p\begin{bmatrix}1 & \dots & 1\\ X_{11} & \dots & X_{n1}\\ X_{12} & \dots & X_{n2}\end{bmatrix}_n$$

$$\mathbb{X}^T\mathbb{X} = {}_p\begin{bmatrix}1 & \dots & 1\\ X_{11} & \dots & X_{n1}\\ X_{12} & \dots & X_{n2}\end{bmatrix}_n {}_n\begin{bmatrix}1 & X_{11} & X_{12}\\ \vdots & \vdots & \vdots\\ 1 & X_{n1} & X_{n2}\end{bmatrix}_p = {}_p\begin{bmatrix}n & \sum X_{i1} & \sum X_{i2}\\ & \sum X_{i1}^2 & \sum X_{i1}X_{i2}\\ \text{sim} & & \sum X_{i2}^2\end{bmatrix}_p$$

$\hookrightarrow$ simétrica

$$\underset{\sim}{X}^T \underset{\sim}{y} = {}_p\begin{bmatrix} 1 & \cdots & 1 \\ x_{11} & \cdots & x_{n1} \\ x_{12} & \cdots & x_{n2} \end{bmatrix}_n {}_n\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}_1 = {}_p\begin{bmatrix} \sum y_i \\ \sum x_{i1} y_i \\ \sum x_{i2} y_i \end{bmatrix}_1$$

Considerando o estimador $\hat{\underset{\sim}{\theta}} = (X^T X)^{-1} X^T \underset{\sim}{y}$, precisamos da inversa da matriz $X^T X$.

Como calcular as inversas?

- Matriz 2x2

$$Q = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad Q^{-1} = \frac{1}{\det(Q)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$ em que $\det(Q) = ad - bc$

- Matriz 3x3

$$Q = \begin{bmatrix} a & b & c \\ b & e & f \\ c & f & h \end{bmatrix} \qquad Q^{-1} = \begin{bmatrix} A & B & C \\ B & E & F \\ C & F & H \end{bmatrix}$$

em que:

$$A = (eh - f^2)/z \qquad E = (ah - c^2)/z$$

$$B = -(bh - cf)/z \qquad F = -(af - bc)/z$$

$$C = (bf - ce)/z \qquad H = (ae - b^2)/z$$

$$z = aeh + 2bcf - af^2 - b^2h - c^2e \Rightarrow \det(Q)$$

Vamos verificar as matrizes e cálculos com o RLS!

Considerando o RLS

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad i = 1, \ldots, n$$

$$\underset{\sim}{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \underset{\sim}{\mathbb{X}} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad \underset{\sim}{\theta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \quad \underset{\sim}{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\mathbb{X}^T\mathbb{X} = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{bmatrix} \qquad \mathbb{X}^T y = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \Sigma y_i \\ \Sigma x_i y_i \end{bmatrix}$$

$$(\mathbb{X}^T\mathbb{X})^{-1} = \frac{1}{n\Sigma x_i^2 - (\Sigma x_i)^2} \begin{bmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & n \end{bmatrix}$$

$$\underset{\sim}{\hat{\theta}} = (\mathbb{X}^T\mathbb{X})^{-1} \mathbb{X}^T \underset{\sim}{y} = \frac{1}{n\Sigma x_i^2 - (\Sigma x_i)^2} \begin{bmatrix} \Sigma x_i^2 & -\Sigma x_i \\ -\Sigma x_i & n \end{bmatrix} \begin{bmatrix} \Sigma y_i \\ \Sigma x_i y_i \end{bmatrix}$$

$$\underset{\sim}{\hat{\theta}} = \frac{1}{n\Sigma x_i^2 - (\Sigma x_i)^2} \begin{bmatrix} \Sigma x_i^2 . \Sigma y_i - \Sigma x_i \Sigma x_i y_i \\ -\Sigma x_i \Sigma y_i + n\Sigma x_i y_i \end{bmatrix}$$

$$\underset{\sim}{\hat{\theta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = \begin{bmatrix} \dfrac{\Sigma x_i^2 \Sigma y_i - \Sigma x_i \Sigma x_i y_i}{n\Sigma x_i^2 - (\Sigma x_i)^2} \\[2ex] \dfrac{n\Sigma x_i y_i - \Sigma x_i \Sigma y_i}{n\Sigma x_i^2 - (\Sigma x_i)^2} \end{bmatrix}$$

→ equivalente $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

→ equivalente $\hat{\beta}_1 = \dfrac{SPXY}{SQX}$

Propriedades:

i) Os elementos de $\hat{\underset{\sim}{\theta}}$ são combinações lineares dos y_i

$$\hat{\underset{\sim}{\theta}} = \begin{bmatrix} \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} \Rightarrow \hat{\beta}_j = \sum_{i=1}^{n} c_{ij} y_i$$

$$\hat{\underset{\sim}{\theta}} = \underbrace{(X^T X)^{-1} X^T}_{{}_p C_n} \underset{\sim}{y}$$

${}_p C_n \Rightarrow$ matriz de coeficientes $\quad \mathbb{C} = {}_p\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{p1} & c_{p2} & & c_{pn} \end{bmatrix}_n$

$$\hat{\underset{\sim}{\theta}} = C\underset{\sim}{y} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & & & \\ c_{p1} & c_{p2} & \cdots & c_{pn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum c_{1i} y_i \\ \sum c_{2i} y_i \\ \vdots \\ \sum c_{pi} y_i \end{bmatrix}$$

ii) $\hat{\underset{\sim}{\theta}}$ é um estimador não viesado para $\underset{\sim}{\theta}$.

$E(A\underset{\sim}{y}) = A\,E(\underset{\sim}{y})$

$$E(\hat{\underset{\sim}{\theta}}) = E\Big[\underbrace{(X^T X)^{-1} X^T}_{A} \underset{\sim}{y}\Big] = (X^T X)^{-1} X^T E(\underset{\sim}{y}) = \underbrace{(X^T X)^{-1} X^T X}_{I} \underset{\sim}{\theta} = \underset{\sim}{\theta}$$

iii) A matriz de covariâncias de $\hat{\underset{\sim}{\theta}}$ é dada por:

$Var(A\underset{\sim}{y}) = A\,Var(\underset{\sim}{y})\,A^T$

$$Var(\hat{\underset{\sim}{\theta}}) = Var\Big[\underbrace{(X^T X)^{-1} X^T}_{A} y\Big] = (X^T X)^{-1} X^T\, Var(\underset{\sim}{y})\, X (X^T X)^{-1}$$

$$= (X^T X)^{-1} X^T I \sigma^2 X (X^T X)^{-1}$$

$$= \underbrace{(X^T X)^{-1} X^T X}_{I} (X^T X)^{-1} \sigma^2$$

$$= (X^T X)^{-1} \sigma^2$$

Para o modelo RLS, temos:

$$Var[\underset{\sim}{\hat{\theta}}] = \underbrace{\frac{1}{\underbrace{n\sum X_i^2 - (\sum X_i)^2}_{n\sum x_i^2}}\begin{bmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & n \end{bmatrix}}_{(X^TX)^{-1}} \cdot \sigma^2$$

$$= \frac{1}{n\sum x_i^2}\begin{bmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & n \end{bmatrix}\sigma^2$$

$$= \begin{bmatrix} \dfrac{\sum X_i^2}{n\sum x_i^2}^{*} & -\dfrac{\sum X_i}{n\sum x_i^2} \\ -\dfrac{\sum X_i}{n\sum x_i^2} & \dfrac{n}{n\sum x_i^2} \end{bmatrix}\sigma^2$$

$$(*)\ \sum X_i^2 = \sum X_i^2 - \frac{2(\sum X_i)^2}{n} + \frac{2(\sum X_i)^2}{n}$$

$$= \sum X_i^2 - 2\bar{X}(\sum X_i) + 2\bar{X}(\sum X_i).\frac{n}{n}$$

$$= \sum X_i^2 - 2\bar{X}(\sum X_i) + 2n\bar{X}^2$$

$$= \underbrace{\sum X_i^2 - 2\bar{X}(\sum X_i) + n\bar{X}} + n\bar{X}$$

$$= \sum (X_i - \bar{X})^2 + n\bar{X}$$

$$= \sum x_i^2 + n\bar{X}$$

$$= \begin{bmatrix} \dfrac{\sum x_i^2 + n\bar{X}}{n\sum x_i^2} & -\dfrac{\sum X_i/n}{\sum x_i^2} \\ -\dfrac{\sum X_i/n}{\sum x_i^2} & \dfrac{1}{\sum x_i^2} \end{bmatrix}\sigma^2$$

$$= \begin{bmatrix} \dfrac{1}{n} + \dfrac{\bar{X}}{\sum x_i^2} & -\dfrac{\bar{X}}{\sum x_i^2} \\ -\dfrac{\bar{X}}{\sum x_i^2} & \dfrac{1}{\sum x_i^2} \end{bmatrix}\sigma^2$$

$Var(\hat{\beta}_0)$

$Var(\hat{\beta}_1)$

$Cov(\hat{\beta}_0, \hat{\beta}_1)$

Vimos que:

$$Var(\hat{\beta}_0) = \left(\frac{1}{n} + \frac{\bar{X}}{\sum x_i^2}\right)\sigma^2$$

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum x_i^2}$$

$$Cov(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\bar{X}\sigma^2}{\sum x_i^2}$$

iv) Como resultado dos itens i, ii e iii e que $\underset{\sim}{Y} \sim N(\mathbb{X}\underset{\sim}{\theta}, I\sigma^2)$

temos que: $\hat{\underset{\sim}{\theta}} \sim N(\underset{\sim}{\theta}, (\mathbb{X}^T\mathbb{X})^{-1}\sigma^2)$

v) Dada uma combinação linear dos parâmetros

$$\underset{\sim}{C}^T\underset{\sim}{\theta} = [c_0 \ c_1 \ \dots \ c_k]\begin{bmatrix}\theta_0 \\ \theta_1 \\ \vdots \\ \theta_k\end{bmatrix} = c_0\theta_0 + c_1\theta_1 + \dots + c_k\theta_k$$

Um estimador de mínimos quadrados, não viesado e de variância mínima é dado por: $\underset{\sim}{C}^T\hat{\underset{\sim}{\theta}}$

$$E(\underset{\sim}{C}^T\hat{\underset{\sim}{\theta}}) = \underset{\sim}{C}^T E(\hat{\underset{\sim}{\theta}}) = \underset{\sim}{C}^T\underset{\sim}{\theta}$$

$$Var(\underset{\sim}{C}^T\hat{\underset{\sim}{\theta}}) = \underset{\sim}{C}^T var(\hat{\underset{\sim}{\theta}})\underset{\sim}{C} = \underset{\sim}{C}^T (\mathbb{X}^T\mathbb{X})^{-1}\underset{\sim}{C}.\sigma^2$$

Desta forma, $\underset{\sim}{C}^T\hat{\underset{\sim}{\theta}} \sim N(\underset{\sim}{C}^T\underset{\sim}{\theta}, \underset{\sim}{C}^T(\mathbb{X}^T\mathbb{X})^{-1}\underset{\sim}{C}\,\sigma^2)$

vi) A aproximação de mínimos quadrados $(\hat{\underset{\sim}{Y}})$ para $\underset{\sim}{Y}$ é dada por:

$$\hat{\underset{\sim}{Y}} = \mathbb{X}\hat{\underset{\sim}{\theta}} = \underbrace{\mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T}_{\text{Hat matrix (matriz de projeção)}}\underset{\sim}{Y} = \overset{\to \text{H ou P}}{H}\underset{\sim}{Y}$$

$$E(\hat{\underset{\sim}{Y}}) = E(H\underset{\sim}{Y}) = H\,E(\underset{\sim}{Y}) = H\mathbb{X}\underset{\sim}{\theta} = \mathbb{X}\underbrace{(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\mathbb{X}}_{I}\underset{\sim}{\theta} = \mathbb{X}\underset{\sim}{\theta} = \underset{\sim}{Y}$$

$$\begin{aligned} Var(\hat{\underset{\sim}{Y}}) = Var(H\underset{\sim}{Y}) &= H\,Var(\underset{\sim}{Y})\,H^T \\ &= \mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T I\sigma^2\,\mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T \\ &= \mathbb{X}\underbrace{(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\mathbb{X}}_{I}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\sigma^2 \\ &= \underbrace{\mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T}_{H}\sigma^2 = H\sigma^2 \end{aligned}$$

$$\hat{\underset{\sim}{Y}} \sim N(\underset{\sim}{Y}, H\sigma^2) \Rightarrow \hat{\underset{\sim}{Y}} \sim N(\mathbb{X}\underset{\sim}{\theta}, \mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\sigma^2)$$