

## FORMULÁRIO DE ESTATÍSTICA II

• **MODELO REGRESSÃO LINEAR (MRL)**

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + u_i, \quad i = 1, 2, \dots, n \Leftrightarrow y_i = \mathbf{x}_i \boldsymbol{\beta} + u_i, \quad i = 1, 2, \dots, n \Leftrightarrow \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

• **OLS (estimadores dos mínimos quadrados)**

Caso geral	Caso particular: $y_i = \beta_0 + \beta_1 x_i + u_i$
$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$	$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$
$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \dots + \hat{\beta}_k x_{ik}$	$\hat{Var}(\hat{\beta}_0   X) = \frac{\hat{\sigma}^2 \sum x_i^2}{n \sum (x_i - \bar{x})^2}$
$\hat{u}_i = y_i - \hat{y}_i$	$\hat{\beta}_1 = \frac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2}$
$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{(n - k - 1)}$	$\hat{Var}(\hat{\beta}_1   \mathbf{X}) = \frac{\hat{\sigma}^2}{\sum (x_i - \bar{x})^2}$
$\hat{Var}(\hat{\boldsymbol{\beta}}   \mathbf{X}) = \hat{\sigma}^2 (\mathbf{X}^T \mathbf{X})^{-1}$	$\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n - 2}$
$\hat{Var}(\hat{\beta}_j   \mathbf{X}) = \frac{\hat{\sigma}^2}{SST_j (1 - R_j^2)}$	

Nota:  $R_j^2$  -  $R^2$  da regressão de  $x_j$  sobre todos os outros regressores;  $SST_j = \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2$ .

**Propriedades:**

$$\sum_{i=1}^n \hat{u}_i = 0 \quad (\text{com termo independente}); \quad \sum_{i=1}^n x_{ij} \hat{u}_i = 0 \quad (j = 1, 2, \dots, k); \quad \sum_{i=1}^n \hat{y}_i \hat{u}_i = 0;$$

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n \hat{u}_i^2 \quad (\text{com termo independente})$$

**Coefficiente de Determinação:**

$$R^2 = \frac{\left( \sum_i (y_i - \bar{y})(\hat{y}_i - \bar{y}) \right)^2}{\sum_i (y_i - \bar{y})^2 \sum_i (\hat{y}_i - \bar{y})^2} = r_{y, \hat{y}}^2 \quad \text{com } r_{y, \hat{y}} \text{ o coeficiente de correlação entre } y \text{ e } \hat{y}$$

**Coefficiente de Determinação no modelo com termo independente:**

$$R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}; \quad SST = \sum_{i=1}^n (y_i - \bar{y})^2; \quad SSE = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2; \quad SSR = \sum_{i=1}^n \hat{u}_i^2$$

**Coefficiente de Determinação ajustado:**

$$\bar{R}^2 = 1 - \frac{SSR / (n - k - 1)}{SST / (n - 1)} = 1 - (1 - R^2) \frac{n - 1}{n - k - 1}$$

• **INFERÊNCIA ESTATÍSTICA DO MRL:**

$$\bullet \quad y_i \sim N(\mathbf{x}_i \boldsymbol{\beta}, \sigma^2) \quad ; \quad u \sim N(0, \sigma^2); \quad \hat{\beta}_j \sim N(\beta_j, \text{Var}(\hat{\beta}_j)); \quad \frac{\sum_{i=1}^n \hat{u}_i^2}{\sigma^2} = \frac{(n - k - 1) \hat{\sigma}^2}{\sigma^2} \sim \chi^2(n - k - 1)$$

$$\bullet \quad t_j = \frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \sim t(n - k - 1) \quad \text{ou} \quad F_j = t_j^2 = \frac{(\hat{\beta}_j - \beta_j)^2}{se(\hat{\beta}_j)^2} \sim F(1, n - k - 1)$$

**Casos particulares:**

$H_0 : \beta_j = 0$	$H_0 : \beta_j = \beta_j^0$
$t_j = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \sim t(n - k - 1)$	$t_j = \frac{\hat{\beta}_j - \beta_j^0}{se(\hat{\beta}_j)} \sim t(n - k - 1)$

- Testes de  $q$  restrições lineares sobre os coeficientes de regressão,  $H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ ,

$$F = \frac{SSR_r - SSR_{ur}}{SSR_{ur}} \times \frac{n-k-1}{q} \sim F(q, n-k-1)$$

$$F = \frac{R_r^2 - R_{ur}^2}{1 - R_{ur}^2} \times \frac{n-k-1}{q} \sim F(q, n-k-1) \text{ (restrições homogêneas)}$$

**Nota:**  $SSR_r$  = soma dos quadrados dos resíduos do modelo com as  $q$  restrições lineares;

$SSR_{ur}$  = soma dos quadrados dos resíduos do modelo sem restrições.

$R_r^2$  = Coeficiente de determinação do modelo com as  $q$  restrições

$R_{ur}^2$  = Coeficiente de determinação do modelo sem restrições

### Casos particulares

- Uma única restrição ( $q=1$ )  $H_0 : \mathbf{R}\boldsymbol{\beta} = r \Leftrightarrow H_0 : \theta = 0$  com  $\theta = \mathbf{R}\boldsymbol{\beta} - r$ ,

$$t = \frac{R\hat{\boldsymbol{\beta}} - r}{\sqrt{\text{Var}(R\hat{\boldsymbol{\beta}})}} \sim t(n-k-1) \quad \text{ou} \quad t = \frac{\hat{\theta}}{se(\hat{\theta})} \sim t(n-k-1) \text{ com } \hat{\theta} = \mathbf{R}\hat{\boldsymbol{\beta}} - r$$

- $H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$

$$F = \frac{R^2}{1 - R^2} \times \frac{n-k-1}{k} \sim F(k, n-k-1)$$

### • HETEROCEDASTICIDADE:

$$\text{Var}(\hat{\boldsymbol{\beta}} | \mathbf{X}) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \text{Var}(\mathbf{u} | \mathbf{X}) \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} = (\mathbf{X}^T \mathbf{X})^{-1} \sum_{i=1}^n \sigma_i^2 \mathbf{x}_i^T \mathbf{x}_i (\mathbf{X}^T \mathbf{X})^{-1}$$

- **Estimador robusto de White** (estimador da variância consistente com heterocedasticidade)

$$\hat{\text{Var}}(\hat{\boldsymbol{\beta}} | \mathbf{X}) = (\mathbf{X}^T \mathbf{X})^{-1} \sum_{i=1}^n \hat{u}_i^2 \mathbf{x}_i^T \mathbf{x}_i (\mathbf{X}^T \mathbf{X})^{-1}$$

- **Inferência sobre  $\beta_j$ :**

$$t_j = \frac{\hat{\beta}_j - \beta_j}{se^*(\hat{\beta}_j)} \sim t(n-k-1) \quad \text{com } se^*(\hat{\beta}_j) = \text{erro-padrão consistente com heterocedasticidade}$$

- **Testes de Heterocedasticidade:**

Estatística-teste LM:  $LM = nR_{\hat{u}^2}^2 \sim \chi^2(p)$  com  $R_{\hat{u}^2}^2$  e  $p$  respectivamente o coeficiente de determinação e o n° de regressores (excluindo o termo independente) da regressão auxiliar de teste.

### • PREVISÃO

- **Previsão em Média:**  $E(y | x_1 = x_1^0, \dots, x_k = x_k^0) = \beta_0 + \beta_1 x_1^0 + \dots + \beta_k x_k^0 = \theta_0$ ;  $\hat{\theta}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_1^0 + \dots + \hat{\beta}_k x_k^0$

$$t_{\hat{\theta}} = \frac{\hat{\theta}_0 - \theta_0}{se(\hat{\theta}_0)} \sim t(n-k-1)$$

- **Previsão pontual:**  $y^0 = \beta_0 + \beta_1 x_1^0 + \dots + \beta_k x_k^0 + u^0$ ;  $\hat{y}^0 = \hat{\beta}_0 + \hat{\beta}_1 x_1^0 + \dots + \hat{\beta}_k x_k^0 = \hat{\theta}_0$

$$t = \frac{y^0 - \hat{y}^0}{\sqrt{se(\hat{\theta}_0)^2 + \hat{\sigma}^2}} \sim t(n-k-1)$$

### • TESTE RESET

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \delta \hat{y}^2 + u \quad \text{- testar } H_0 : \delta = 0$$

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \delta_1 \hat{y}^2 + \delta_2 \hat{y}^3 + u \quad \text{- testar } H_0 : \delta_1 = \delta_2 = 0$$