FORMULÁRIO DE ESTATÍSTICA II

• VALOR ESPERADO, MOMENTOS E PARÂMETROS

- $\text{Var}(X) = E(X-\mu)^2 = E(X^2) - \mu^2$
- $\text{Cov}(X,Y) = E\{(X-\mu_X)(Y-\mu_Y)\} = E(XY) - E(X)E(Y)$; $\rho_{X,Y} = \dfrac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$
- $E(aX+bY) = aE(X) + bE(Y)$; $\text{Var}(aX+bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(X,Y)$ com a, b constantes

• DISTRIBUIÇÕES TEÓRICAS

• UNIFORME (DISCRETA)

- Caso $x = 1,2,\ldots,n$: $f(x) = \dfrac{1}{n}$; $E(X) = \dfrac{n+1}{2}$; $\text{Var}(X) = \dfrac{n^2-1}{12}$
- Caso $x = 0,1,2,\ldots,m$: $f(x) = \dfrac{1}{m+1}$; $E(X) = \dfrac{m}{2}$; $\text{Var}(X) = \dfrac{m(m+2)}{12}$

• BERNOULLI $X \sim B(1,\theta)$

$f(x\,|\,\theta) = \theta^x(1-\theta)^{1-x}$, $\quad x = 0,1 \quad (0<\theta<1)$; $E(X) = \theta$; $\text{Var}(X) = \theta(1-\theta)$

• BINOMIAL $X \sim B(n,\theta)$

$f(x\,|\,\theta) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$, $x = 0,1,2,\ldots,n$, $(0<\theta<1)$;

$E(X) = n\theta$; $\text{Var}(X) = n\theta(1-\theta)$; $\Im(\theta) = \dfrac{n}{\theta(1-\theta)}$, n conhecido

Propriedades:

- $X \sim B(n,\theta) \Leftrightarrow (n-X) \sim B(n,1-\theta)$
- $X_i \sim B(n_i,\theta)$, independentes $(i = 1,2,\ldots,k)$ $\Rightarrow \sum_{i=1}^{k} X_i \sim B(n,\theta)$, $n = \sum_{i=1}^{k} n_i$

• POISSON $X \sim \text{Po}(\lambda)$

$f(x\,|\,\lambda) = \dfrac{e^{-\lambda}\lambda^x}{x!}$, $x = 0,1,2,\ldots$, $(\lambda>0)$; $E(X) = \lambda$; $\text{Var}(X) = \lambda$; $\Im(\lambda) = \dfrac{1}{\lambda}$

Propriedades:

- $X_i \sim \text{Po}(\lambda_i)$ independentes $(i = 1,2,\ldots,k)$ $\Rightarrow X = \sum_{i=1}^{k} X_i \sim \text{Po}\left(\sum_{i=1}^{k}\lambda_i\right)$
- $X \sim B(n,\theta)$ com n grande e θ pequeno, então $X \overset{a}{\sim} \text{Po}(n\theta)$

• UNIFORME (CONTÍNUA) $X \sim \text{U}(\alpha,\beta)$

$f(x\,|\,\alpha,\beta) = \dfrac{1}{\beta-\alpha} \quad \alpha < x < \beta$; $\text{E}(X) = \dfrac{\alpha+\beta}{2}$; $\text{Var}(X) = \dfrac{(\beta-\alpha)^2}{12}$

• NORMAL $X \sim N(\mu,\sigma^2)$

$f(x\,|\,\mu,\sigma^2) = \dfrac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\dfrac{1}{2\sigma^2}(x-\mu)^2\right\}$, $\quad -\infty < x < +\infty$, $-\infty < \mu < +\infty$, $0 < \sigma < +\infty$;

$E(X) = \mu$; $\text{Var}(X) = \sigma^2$; $\Im(\mu) = \dfrac{1}{\sigma^2}$ σ^2 conhecido ; $\Im(\sigma^2) = \dfrac{1}{2\sigma^4}$ μ conhecido

Propriedades:

- Normal estandardizada $Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$; $\phi(z) = \phi(-z)$; $\Phi(z) = 1 - \Phi(-z)$
- $X_i \sim N(\mu,\sigma^2)$ independentes $(i = 1,2,\ldots,n)$ $\Rightarrow Y = \sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$; $\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \dfrac{\sigma^2}{n}\right)$
- $X_i \sim N(\mu_i,\sigma_i^2)$ independentes $(i = 1,2,\ldots,k)$ $\Rightarrow Y = \sum_{i=1}^{k}\alpha_i X_i \sim N(\mu_Y,\sigma_Y^2)$; $\mu_Y = \sum_{i=1}^{k}\alpha_i\mu_i$; $\sigma_Y^2 = \sum_{i=1}^{k}\alpha_i^2\sigma_i^2$

- **EXPONENCIAL** $X \sim \text{Ex}(\lambda)$; $X \sim \text{Ex}(\lambda) \Leftrightarrow X \sim G(1,\lambda)$

  $f(x \mid \lambda) = \lambda e^{-\lambda x}$ $x > 0,\ \lambda > 0$; $F(x \mid \lambda) = 1 - e^{-\lambda x}$; $E(X) = \dfrac{1}{\lambda}$; $\text{Var}(X) = \dfrac{1}{\lambda^2}$; $\Im(\lambda) = \dfrac{1}{\lambda^2}$

  Propriedades:

  - $X_i \sim \text{Ex}(\lambda)$ independentes $(i = 1,2,\ldots,k)$ $\Rightarrow \sum_{i=1}^{k} X_i \sim G(k;\lambda)$ e $\min_i X_i \sim \text{Ex}(k\lambda)$

- **GAMA** $X \sim G(\alpha,\lambda)$

  Função gama: $\Gamma(\alpha) = \int_0^{+\infty} e^{-x} x^{\alpha-1} dx$ $(\alpha > 0)$ com $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1), \alpha > 1$; $\Gamma(n) = (n-1)!$; $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

  $f(x \mid \alpha,\lambda) = \dfrac{\lambda^{\alpha} e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)}$, $x > 0$, $\alpha,\lambda > 0$; $E(X) = \dfrac{\alpha}{\lambda}$; $\text{Var}(X) = \dfrac{\alpha}{\lambda^2}$; $\Im(\lambda) = \dfrac{\alpha}{\lambda^2}$ α conhecido

  Propriedades:

  - $X_i \sim G(\alpha_i,\lambda)$, $(i = 1,2,\ldots,k)$ independentes $\Rightarrow \sum_{i=1}^{k} X_i \sim \text{G}(\alpha,\lambda)$; $\alpha = \sum_{i=1}^{k} \alpha_{\text{i}}$
  - $X \sim G(\alpha,\lambda) \Rightarrow cX \sim G\left(\alpha, \dfrac{\lambda}{c}\right)$ $c > 0$ constante

- **QUI-QUADRADO** $X \sim \chi^2(n)$

  $f(x \mid n) = \dfrac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)}$, $x > 0$, $n > 0$ inteiro ; $E(X) = n$; $\text{Var}(X) = 2n$;

  Propriedades:

  - $X \sim \chi^2(n) \Leftrightarrow X \sim \text{G}\left(\dfrac{n}{2},\dfrac{1}{2}\right)$
  - $\sqrt{2\chi^2(n)} - \sqrt{2n-1} \overset{a}{\sim} N(0,1)$
  - $X \sim \text{G}(n;\lambda) \Leftrightarrow 2\lambda X \sim \chi^2(2n)$
  - $X_i \sim \chi^2(n_i)$, independentes $(i = 1,2,\ldots,k) \Rightarrow \sum_{i=1}^{k} X_i \sim \chi^2(n)$, $n = \sum_{i=1}^{k} n_i$
  - $X_i \sim N(0,1)$ independentes $(i = 1,2,\ldots,k) \Rightarrow \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$; $X \sim N(0,1) \Rightarrow X^2 \sim \chi^2(1)$

- ***t*-"STUDENT"** $T \sim t(n)$

  $T = \dfrac{U}{\sqrt{V/n}} \sim t(n)$ com $U \sim N(0,1)$ e $V \sim \chi^2(n)$ independentes; $E(T) = 0$; $\text{Var}(T) = \dfrac{n}{n-2}$ $(n > 2)$

- ***F*-SNEDCOR** $X \sim F(m,n)$

  $F = \dfrac{U/m}{V/n} \sim F(m,n)$ com $U \sim \chi^2(m)$, $V \sim \chi^2(n)$ independentes

  $E(X) = \dfrac{n}{n-2}$ $(n > 2)$; $\text{Var}(X) = \dfrac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ $(n > 4)$

  Propriedades: - $X \sim F(m,n) \Rightarrow \dfrac{1}{X} \sim F(n,m)$ - $T \sim t(n) \Rightarrow T^2 \sim F(1,n)$

- **TEOREMA DO LIMITE CENTRAL E COROLÁRIOS**

  X_i iid, com $E(X_i) = \mu$ e $\text{Var}(X_i) = \sigma^2 \Rightarrow \dfrac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1)$

  Corolário: $X_i \sim B(1;\theta)$, independentes então $\dfrac{\sum_{i=1}^{n} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \overset{a}{\sim} N(0,1)$

  Correcção de continuidade: $P(a \le X \le b) \approx \Phi\left(\dfrac{b + \frac{1}{2} - n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \Phi\left(\dfrac{a - \frac{1}{2} - n\theta}{\sqrt{n\theta(1-\theta)}}\right)$ com a e b inteiros

  Corolário: $X \sim \text{Po}(\lambda) \Rightarrow \dfrac{X - \lambda}{\sqrt{\lambda}} \overset{a}{\sim} N(0,1)$, quando $\lambda \to +\infty$

Correcção de continuidade: $P(a \le X \le b) \approx \Phi\left(\frac{b+\frac{1}{2}-\lambda}{\sqrt{\lambda}}\right) - \Phi\left(\frac{a-\frac{1}{2}-\lambda}{\sqrt{\lambda}}\right)$

- **AMOSTRAGEM. DISTRIBUIÇÕES POR AMOSTRAGEM**
  - **MÉDIA E VARIÂNCIA AMOSTRAIS**

$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$; $S^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2$; $S'^2 = \frac{n}{n-1}S^2$; $E(\overline{X}) = \mu$; $\text{Var}(\overline{X}) = \frac{\sigma^2}{n}$; $E(S^2) = \frac{n-1}{n}\sigma^2$; $E(S'^2) = \sigma^2$

  - **DISTRIBUIÇÕES POR AMOSTRAGEM**

POPULAÇÕES NORMAIS

Média	$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\frac{\overline{X}-\mu}{S'/\sqrt{n}} \sim t(n-1)$
Diferença de médias	Variâncias conhecidas $\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \sim N(0,1)$	
	Variâncias desconhecidas mas iguais $T = \frac{\overline{X}_1-\overline{X}_2-(\mu_1-\mu_2)}{\sqrt{\frac{1}{m}+\frac{1}{n}}\sqrt{\frac{(m-1)S_1'^2+(n-1)S_2'^2}{m+n-2}}} \sim t(m+n-2)$	
Variância	$\frac{nS^2}{\sigma^2} = \frac{(n-1)S'^2}{\sigma^2} \sim \chi^2(n-1)$	
Relação de variâncias	$\frac{S_1'^2}{S_2'^2}\frac{\sigma_2^2}{\sigma_1^2} \sim F(m-1,n-1)$ ou $\frac{S_2'^2}{S_1'^2}\frac{\sigma_1^2}{\sigma_2^2} \sim F(n-1,m-1)$	
Amostras emparelhadas	$\frac{\overline{Z}-\mu_X-\mu_Y}{S_Z'/\sqrt{n}} \sim t(n-1)$, (X_i, Y_i) - amostra emparelhada, $Z_i = X_i - Y_i$; $\overline{Z} = \overline{X} - \overline{Y}$	

GRANDES AMOSTRAS: CASO GERAL

Média	$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1)$	$\frac{\overline{X}-\mu}{S'/\sqrt{n}} \overset{a}{\sim} N(0,1)$
Diferença de médias	$\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \overset{a}{\sim} N(0,1)$	$\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1'^2}{m}+\frac{S_2'^2}{n}}} \overset{a}{\sim} N(0,1)$

GRANDES AMOSTRAS: POPULAÇÃO DE BERNOULLI

Proporção	$\frac{\overline{X}-\theta}{\sqrt{\frac{\theta(1-\theta)}{n}}} \overset{a}{\sim} N(0,1)$	$\frac{\overline{X}-\theta}{\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}} \overset{a}{\sim} N(0,1)$
Diferença de proporções	$\frac{\overline{X}_1-\overline{X}_2-(\theta_1-\theta_2)}{\sqrt{\frac{\theta_1(1-\theta_1)}{m}+\frac{\theta_2(1-\theta_2)}{n}}} \overset{a}{\sim} N(0,1)$	$\frac{\overline{X}_1-\overline{X}_2-(\theta_1-\theta_2)}{\sqrt{\frac{\overline{X}_1(1-\overline{X}_1)}{m}+\frac{\overline{X}_2(1-\overline{X}_2)}{n}}} \overset{a}{\sim} N(0,1)$
Igualdade de proporções	$\frac{\overline{X}_1-\overline{X}_2}{\sqrt{\left(\frac{1}{m}+\frac{1}{n}\right)\hat{\theta}(1-\hat{\theta})}} \overset{a}{\sim} N(0,1)$ com $\hat{\theta} = \frac{m\overline{X}_1+n\overline{X}_2}{m+n}$	

GRANDES AMOSTRAS: **POPULAÇÃO DE POISSON**

Média	$\frac{\overline{X}-\lambda}{\sqrt{\frac{\lambda}{n}}} \overset{a}{\sim} N(0,1)$	$\frac{\overline{X}-\lambda}{\sqrt{\frac{\overline{X}}{n}}} \overset{a}{\sim} N(0,1)$
Diferença de médias	$\frac{\overline{X}_1-\overline{X}_2-(\lambda_1-\lambda_2)}{\sqrt{\frac{\lambda_1}{m}+\frac{\lambda_2}{n}}} \overset{a}{\sim} N(0,1)$	$\frac{\overline{X}_1-\overline{X}_2-(\lambda_1-\lambda_2)}{\sqrt{\frac{\overline{X}_1}{m}+\frac{\overline{X}_2}{n}}} \overset{a}{\sim} N(0,1)$
Igualdade de médias	$\frac{\overline{X}_1-\overline{X}_2}{\sqrt{\left(\frac{1}{m}+\frac{1}{n}\right)\overline{X}}} \overset{a}{\sim} N(0,1)$ com $\overline{X} = \frac{m\overline{X}_1+n\overline{X}_2}{m+n}$	

- **ESTATÍSTICA-TESTE DO χ^2**
  - **TESTE DE AJUSTAMENTO**

    $Q = \sum_{j=1}^{m} \frac{(N_j - fe_j)^2}{fe_j} \overset{a}{\sim} \chi^2(m-1)$; $fe_j = np_{\circ j}$ - frequência esperada da classe j ;

    Com estimação de k parâmetros para obter as estimativas $\hat{p}_{\circ j}$: $Q \overset{a}{\sim} \chi^2(m-k-1)$
  - **TESTE DE INDEPENDÊNCIA**:

    $Q = \sum_{i=1}^{r}\sum_{j=1}^{s} \frac{(N_{ij} - fe_{ij})^2}{fe_{ij}} \overset{a}{\sim} \chi^2((r-1)(s-1))$; $fe_{ij} = \frac{N_{i\circ}N_{\circ j}}{n}$ - frequência esperada da classe ij

- **MODELO REGRESSÃO LINEAR (MRL)**

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

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

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

<u>Nota</u>: 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:

1. $\sum_{i=1}^{n}\hat{u}_i = 0$ (com termo independente)
2. $\sum_{i=1}^{n} x_{ij}\hat{u}_i = 0$ $(j = 1,2,\ldots,k)$;
3. $\sum_{i=1}^{n}\hat{y}_i\hat{u}_i = 0$;
4. $\sum_{i=1}^{n} y_i^2 = \sum_{i=1}^{n}\hat{y}_i^2 + \sum_{i=1}^{n}\hat{u}_i^2$ (com termo independente)

Coeficiente de Determinação:

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

Coeficiente de Determinação no modelo com termo independente:

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

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

- **INFERÊNCIA ESTATÍSTICA DO MRL**:
- $y_i \sim N(\mathbf{x}_i \boldsymbol{\beta}, \sigma^2)$; $u \sim N(0, \sigma^2)$
- $\hat{\beta}_j \sim N\left(\beta_j, Var(\hat{\beta}_j)\right)$
- $\frac{\sum_{i=1}^n \hat{u}_i^2}{\sigma^2} = \frac{(n-k-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-k-1)$
- $t_j = \frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \sim t(n-k-1)$ ou $F_j = t_j^2 = \frac{\left(\hat{\beta}_j - \beta_j\right)^2}{se(\hat{\beta}_j)^2} \sim F(1, n-k-1)$

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, $H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{r}$,

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

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.

Casos particulares

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

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

- Nulidade conjunta dos coeficientes de declive

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

- Nulidade conjunta de um subconjunto de ***q*** coeficientes

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

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

Notas: R_r^2 = Coeficiente de determinação do modelo com as *q* restrições
R_{ur}^2 = Coeficiente de determinação do modelo sem restrições

- **HETEROCEDASTICIDADE**:

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

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

    $\hat{V}ar(\hat{\boldsymbol{\beta}} \mid \mathbf{X}) = (\mathbf{X}^T\mathbf{X})^{-1}\sum_{i=1}^{n}\hat{u}_i^2\mathbf{x}_i^T\mathbf{x}_i\,(\mathbf{X}^T\mathbf{X})^{-1}$.

  - **Inferência sobre** β_j :

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

  - **Testes de Heterocedasticidade:**

    Estatística-teste LM: $\text{LM} = nR_{\hat{u}^2}^2 \overset{a}{\sim} \chi^2(p)$ com $R_{\hat{u}^2}^2$ e p respectivamente o coeficiente de determinação e o nº de regressores (excluindo o termo independente) da regressão auxiliar de teste.

- **PREVISÃO**

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

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

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

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

  - **Variável dependente** $\log(y)$:

    $\widehat{\log} y^0 = \hat{\beta}_0 + \hat{\beta}_1 x_1^0 + \ldots + \hat{\beta}_k x_k^0$

    - se $u \sim N(0,\sigma^2) \Rightarrow \hat{y}^0 = \exp(\hat{\sigma}^2/2)\exp(\widehat{\log} y^0)$

    - outras situações $\Rightarrow \hat{y}^0 = \hat{\alpha}_0 \exp(\widehat{\log} y^0)$ com $\hat{\alpha}_0$ obtido numa regressão auxiliar

- **TESTE RESET**

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

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