Teste de Hipótese em Regressão Normal Linear Múltipla

Análise do Teste $H_0 : \beta_1 = \mathbf{0}_p$

Curso de Inferência Estatística - PPGEST/UFPE

17 de novembro de 2025

1 Modelo e Fundamentos Teóricos

A regressão linear múltipla modela a relação entre uma variável resposta Y e múltiplas variáveis explicativas $X_1, X_2, \ldots, X_p$. Os testes de hipótese permitem avaliar a significância estatística dos parâmetros e a relevância das variáveis explicativas. Este relatório apresenta o teste de hipótese $H_0 : \beta_1 = \mathbf{0}_p$ no modelo de regressão normal linear múltipla, analisando sua fundamentação teórica.

1.1 Especificação do Modelo

O modelo de regressão linear múltipla é especificado como:

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon \tag{1}$$

onde $\mu_i(\beta) = x_i^T\beta = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}$, $\mathbf{y}^T = (y_1, \ldots, y_n)$ é o vetor de variáveis resposta, $\beta = (\beta_0, \beta_1, \ldots, \beta_p)^T$ são parâmetros desconhecidos, $\mathbf{X}$ é a matriz modelo $n \times (p+1)$ de planejamento com primeira coluna de uns, e $\varepsilon^T = (\varepsilon_1, \ldots, \varepsilon_n)$ é o vetor de erros aleatórios.

1.2 Pressupostos Clássicos

Definição 1.1 (Pressupostos do Modelo Normal Linear).

1. **Linearidade**: $\mu_i(\beta) = x_i^T\beta$.
2. **Normalidade**: $\varepsilon_i \sim N(0, \sigma^2)$, $i = 1, \ldots, n$, independentes.
3. **Homocedasticidade**: $\mathrm{Var}(Y_i \mid x_i) = \sigma^2$ constante.
4. **Independência**: $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$, $\forall i \neq j$.
5. **Não-colinearidade**: $\mathbf{X}^T\mathbf{X}$ é inversível.

Sob esses pressupostos, $\mathbf{y} \sim N(\mathbf{X}\beta, \sigma^2\mathbf{I}_n)$.

1.3 Estimadores de Mínimos Quadrados

O estimador de mínimos quadrados (E.M.Q.) para β minimiza $S(\beta) = (\mathbf{y}-\mathbf{X}\beta)^T(\mathbf{y}-\mathbf{X}\beta)$, resultando em:

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} \tag{2}$$

Teorema 1.1 (Propriedades do Estimador). *Sob os pressupostos clássicos:*

(i) $E[\hat{\beta}] = \beta$ *(não-viesado);*

(ii) $\hat{\beta}$ *é o melhor estimador linear não-viesado (Teorema de Gauss-Markov);*

(iii) $\hat{\beta} \sim N(\beta, \sigma^2(\mathbf{X}^T\mathbf{X})^{-1})$.

Demonstração. A propriedade (i) segue diretamente de $E[\hat{\beta}] = E[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}] = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T E[\mathbf{y}] = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}\beta = \beta$.

Para (ii), considere qualquer estimador linear não-viesado $\tilde{\beta} = \mathbf{A}\mathbf{y}$ de β, onde $\mathbf{A}$ é uma matriz $(p+1) \times n$. A não-viesade requer $\mathbf{A}\mathbf{X} = \mathbf{I}_{p+1}$. A matriz de covariância é

$\mathrm{Var}(\tilde{\beta}) = \sigma^2 \mathbf{A}\mathbf{A}^T$. Escrevendo $\mathbf{A} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T + \mathbf{C}$, temos que $\mathbf{C}\mathbf{X} = \mathbf{0}$ pela condição de não-viesade. Então:

$$\begin{aligned}
\mathrm{Var}(\tilde{\beta}) &= \sigma^2 \left[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T + \mathbf{C}\right] \left[\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1} + \mathbf{C}^T\right] \\
&= \sigma^2(\mathbf{X}^T\mathbf{X})^{-1} + \sigma^2\mathbf{C}\mathbf{C}^T \\
&= \mathrm{Var}(\hat{\beta}) + \sigma^2\mathbf{C}\mathbf{C}^T
\end{aligned}$$

Como $\mathbf{C}\mathbf{C}^T$ é semidefinida positiva, $\mathrm{Var}(\tilde{\beta}) \geq \mathrm{Var}(\hat{\beta})$ no sentido matricial, estabelecendo a optimalidade de $\hat{\beta}$.

A propriedade (iii) segue da normalidade de $\mathbf{y}$ e do fato de que $\hat{\beta}$ é uma transformação linear de $\mathbf{y}$. □

O estimador não-viesado de σ^2 é dado por:

$$\hat{\sigma}^2 = \frac{SSE}{n-p-1}, \quad \text{onde} \quad SSE = (\mathbf{y} - \mathbf{X}\hat{\beta})^T(\mathbf{y} - \mathbf{X}\hat{\beta}) \tag{3}$$

Proposição 1.1 (Distribuição de $\hat{\sigma}^2$ e Independência)**.** *Sob os pressupostos do modelo,* $\frac{(n-p-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p-1}$ *e é independente de* $\hat{\beta}$*.*

Demonstração. Seja $\mathbf{P} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ a matriz de projeção ortogonal no espaço coluna $\mathcal{C}(\mathbf{X})$ de $\mathbf{X}$. Então $\mathbf{P}\mathbf{y} = \mathbf{X}\hat{\beta} = \hat{\mathbf{y}}$ e $\mathbf{I}_n - \mathbf{P}$ é a projeção no complemento ortogonal. Temos:

$$SSE = \mathbf{y}^T(\mathbf{I}_n - \mathbf{P})\mathbf{y} = (\mathbf{y} - \mathbf{X}\beta)^T(\mathbf{I}_n - \mathbf{P})(\mathbf{y} - \mathbf{X}\beta)$$

pois $(\mathbf{I}_n - \mathbf{P})\mathbf{X}\beta = \mathbf{0}$.

Como $\mathbf{y} - \mathbf{X}\beta = \varepsilon \sim N(\mathbf{0}, \sigma^2\mathbf{I}_n)$ e $(\mathbf{I}_n - \mathbf{P})$ é uma projeção com traço igual a $n-p-1$, segue do Teorema de Cochran que $\frac{SSE}{\sigma^2} = \frac{(n-p-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p-1}$.

Para a independência, observe que $\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$ e $\hat{\sigma}^2$ dependem de $\mathbf{P}\mathbf{y}$ e $(\mathbf{I}_n - \mathbf{P})\mathbf{y}$ respectivamente. Como $\mathbf{P}$ e $(\mathbf{I}_n - \mathbf{P})$ são projeções em subespaços ortogonais, $\mathbf{P}\mathbf{y}$ e $(\mathbf{I}_n - \mathbf{P})\mathbf{y}$ são independentes para vetores normais, implicando a independência entre $\hat{\beta}$ e $\hat{\sigma}^2$. □

1.4 Teorema de Cochran e Distribuições Qui-Quadrado

O resultado fundamental sobre distribuições qui-quadrado de formas quadráticas é dado pelo Teorema de Cochran, que fundamenta teoricamente a decomposição de somas de quadrados.

Teorema 1.2 (Teorema de Cochran)**.** *Seja* $\mathbf{y} \sim N(\boldsymbol{\mu}, \sigma^2\mathbf{I}_n)$ *e sejam* $\mathbf{Q}_1, \ldots, \mathbf{Q}_k$ *matrizes simétricas idempotentes tais que* $\sum_{i=1}^k \mathbf{Q}_i = \mathbf{I}_n$ *e* $\sum_{i=1}^k \mathrm{tr}(\mathbf{Q}_i) = n$*. Se* $\mathbf{Q}_i\boldsymbol{\mu} = \mathbf{0}$ *para* $i = 1, \ldots, k$*, então as formas quadráticas* $\mathbf{y}^T\mathbf{Q}_i\mathbf{y}$*,* $i = 1, \ldots, k$*, são independentes e* $\frac{\mathbf{y}^T\mathbf{Q}_i\mathbf{y}}{\sigma^2} \sim \chi^2_{\nu_i}$*, onde* $\nu_i = \mathrm{tr}(\mathbf{Q}_i)$*.*

O Teorema de Cochran estabelece que sob condições adequadas, somas de quadrados podem ser decompostas em componentes qui-quadrado independentes, sendo fundamental para a análise de variância e inferência em modelos lineares. No contexto do modelo de regressão, aplicamos este teorema com $\mathbf{Q}_1 = \mathbf{P} - \mathbf{P}_0$ (projeção no espaço gerado pelas variáveis explicativas, excluindo o intercepto) e $\mathbf{Q}_2 = \mathbf{I}_n - \mathbf{P}$ (projeção no complemento ortogonal), onde $\mathbf{P}_0 = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$ é a projeção no espaço gerado pelo vetor de uns.

1.5 Distribuições Assintóticas e Consistência

Embora os resultados finitos sejam exatos sob normalidade, é importante caracterizar o comportamento assintótico dos estimadores quando os pressupostos podem ser relaxados.

Teorema 1.3 (Consistência e Distribuição Assintótica). *Sob condições de regularidade (incluindo $n^{-1}\mathbf{X}^T\mathbf{X} \to \mathbf{\Sigma}$ positiva definida quando $n \to \infty$) e assumindo $E[\varepsilon_i] = 0$ e* $\text{Var}(\varepsilon_i) = \sigma^2$ *finita:*

(i) $\hat{\beta} \xrightarrow{p} \beta$ *(consistência);*

(ii) $\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(\mathbf{0}, \sigma^2\mathbf{\Sigma}^{-1})$ *(normalidade assintótica);*

(iii) $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$ *(consistência);*

(iv) $\frac{\hat{\beta}_j - \beta_j}{SE(\hat{\beta}_j)} \xrightarrow{d} N(0, 1)$ *assintoticamente, mesmo sem normalidade exata dos erros.*

Demonstração. A consistência de $\hat{\beta}$ segue de $\hat{\beta} = \beta + (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{\varepsilon}$ e da Lei dos Grandes Números aplicada a $n^{-1}\mathbf{X}^T\boldsymbol{\varepsilon}$, que converge em probabilidade para zero sob as condições assumidas.

Para (ii), observe que $\sqrt{n}(\hat{\beta} - \beta) = (n^{-1}\mathbf{X}^T\mathbf{X})^{-1}n^{-1/2}\mathbf{X}^T\boldsymbol{\varepsilon}$. Como $n^{-1}\mathbf{X}^T\mathbf{X} \to \mathbf{\Sigma}$ e, pelo Teorema Central do Limite, $n^{-1/2}\mathbf{X}^T\boldsymbol{\varepsilon} \xrightarrow{d} N(\mathbf{0}, \sigma^2\mathbf{\Sigma})$, segue o resultado por aplicação do Teorema de Slutsky.

A consistência de $\hat{\sigma}^2$ decorre da consistência de $\hat{\beta}$ e da aplicação da Lei dos Grandes Números à soma de quadrados dos resíduos.

O resultado (iv) segue da normalidade assintótica de $\hat{\beta}_j$ e da consistência de $\hat{\sigma}^2$. □

Estes resultados garantem que a inferência baseada em testes t e F permanece válida assintoticamente mesmo quando a normalidade exata não se verifica, desde que os erros tenham variância finita e sejam não correlacionados.

2 Teste de Hipótese $H_0 : \beta_1 = \mathbf{0}_p$

2.1 Formulação do Teste

Particionando $\beta = (\beta_0, \beta_1^T)^T$ onde $\beta_1 = (\beta_1, \ldots, \beta_p)^T$, o teste avalia se as variáveis explicativas têm efeito significativo sobre Y:

$$H_0 : \beta_1 = \mathbf{0}_p \quad \text{versus} \quad H_1 : \beta_1 \neq \mathbf{0}_p \tag{4}$$

onde $\mathbf{0}_p$ é o vetor nulo de dimensão p. Sob H_0, o modelo reduz-se a $Y_i = \beta_0 + \varepsilon_i$.

2.2 Estatística F e Distribuição

A estatística de teste F é definida como:

$$F = \frac{MSR}{MSE} = \frac{SSR/p}{SSE/(n-p-1)} \tag{5}$$

onde $SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$ é a Soma dos Quadrados da Regressão e $MSR = SSR/p$, $MSE = SSE/(n-p-1)$ são os quadrados médios.

Teorema 2.1 (Distribuição da Estatística F). *Sob $H_0 : \beta_1 = \mathbf{0}_p$ e os pressupostos do modelo:*

$$F \sim F_{p,n-p-1} \tag{6}$$

2.3 Fundamentação Teórica da Estatística F

A escolha da estatística F para o teste $H_0 : \beta_1 = \mathbf{0}_p$ decorre de propriedades teóricas fundamentais que conectam a teoria de modelos lineares com a teoria geral de testes de hipótese. Nesta seção, desenvolvemos a fundamentação completa, mostrando como a estrutura distribucional do modelo normal linear conduz naturalmente à estatística F como uma estatística pivotal apropriada.

2.3.1 Construção como Estatística Pivotal

Sob os pressupostos do modelo normal linear, temos que $\frac{SSR}{\sigma^2} \sim \chi^2_p$ e $\frac{SSE}{\sigma^2} \sim \chi^2_{n-p-1}$ são independentes, conforme estabelecido anteriormente e fundamentado pelo Teorema de Cochran. A estatística F emerge como a razão adequadamente normalizada dessas quantidades, eliminando o parâmetro de nuisance σ^2.

Proposição 2.1 (Eliminação do Parâmetro de Nuisance)**.** *Sejam $Q_1 \sim \sigma^2\chi^2_{\nu_1}$ e $Q_2 \sim \sigma^2\chi^2_{\nu_2}$ variáveis aleatórias independentes com parâmetro de escala comum $\sigma^2 > 0$. Então a estatística*

$$R = \frac{Q_1/\nu_1}{Q_2/\nu_2}$$

segue uma distribuição F_{ν_1,ν_2} que não depende de σ^2.

Demonstração. Como $Q_1 = \sigma^2 U$ e $Q_2 = \sigma^2 V$ onde $U \sim \chi^2_{\nu_1}$ e $V \sim \chi^2_{\nu_2}$ são independentes, temos:

$$R = \frac{(\sigma^2 U)/\nu_1}{(\sigma^2 V)/\nu_2} = \frac{\sigma^2}{\sigma^2} \cdot \frac{U/\nu_1}{V/\nu_2} = \frac{U/\nu_1}{V/\nu_2}$$

A última quantidade segue uma distribuição F_{ν_1,ν_2} por definição da distribuição F, e não depende de σ^2. □

Aplicando este resultado ao contexto do teste, obtemos que $F = \frac{MSR}{MSE} = \frac{(SSR/\sigma^2)/p}{(SSE/\sigma^2)/(n-p-1)}$ é uma estatística pivotal, cuja distribuição sob H_0 é completamente especificada e não depende de parâmetros desconhecidos.

2.3.2 Derivação via Teste de Razão de Verossimilhança

O teste F pode ser derivado como uma transformação monotônica da estatística de razão de verossimilhança, conectando-o com a teoria geral de testes de hipótese. Sob os pressupostos do modelo normal linear, a função de verossimilhança é:

$$L(\beta, \sigma^2) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{X}\beta)^T(\mathbf{y} - \mathbf{X}\beta)\right\} \tag{7}$$

Sob $H_0 : \beta_1 = \mathbf{0}_p$, o modelo reduzido é $Y_i = \beta_0 + \varepsilon_i$. Particionando a matriz $\mathbf{X} = [\mathbf{1}_n \,|\, \mathbf{X}_1]$, onde $\mathbf{1}_n$ é o vetor de uns e $\mathbf{X}_1$ é a matriz $n \times p$ das variáveis explicativas, temos $\beta = (\beta_0, \beta_1^T)^T$. Sob H_0, o estimador de máxima verossimilhança do intercepto é $\hat{\beta}_0^{(0)} = \bar{y}$ e do parâmetro de variância é $\hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^n (y_i - \bar{y})^2 = \frac{SSE_0}{n}$.

Sob o modelo completo (alternativa), os estimadores de máxima verossimilhança são $\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$ e $\hat{\sigma}^2 = \frac{SSE}{n}$.

A estatística de razão de verossimilhança é:

$$\Lambda = \frac{\sup_{H_0} L(\beta, \sigma^2)}{\sup L(\beta, \sigma^2)} = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{n/2} = \left(\frac{SSE}{SSE_0}\right)^{n/2}$$

Como $SSE_0 = SSR + SSE$, onde $SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$ é a soma de quadrados da regressão, temos:

$$\Lambda = \left(\frac{SSE}{SSE + SSR}\right)^{n/2} = \left(1 + \frac{SSR}{SSE}\right)^{-n/2} = \left(1 + \frac{pF}{n-p-1}\right)^{-n/2}$$

onde utilizamos a relação $SSR = p \cdot MSR = p \cdot F \cdot MSE = \frac{p(n-p-1)}{n-p-1} F \cdot MSE = \frac{pF}{n-p-1} \cdot SSE$.

A estatística $-2 \log \Lambda = n \log(1 + \frac{pF}{n-p-1})$ é monotonicamente crescente em F, estabelecendo que o teste F é equivalente ao teste de razão de verossimilhança. Sob H_0, temos $-2 \log \Lambda \sim \chi^2_p$ assintoticamente, e para amostras finitas sob normalidade, a estatística F possui distribuição exata.

2.3.3 Interpretação e Propriedades

A estatística F compara a variância explicada pelo modelo (MSR) com a variância residual (MSE). Valores grandes de F indicam que a variância explicada é significativamente maior que a residual, fornecendo evidência contra H_0. Além disso, a natureza multivariada da hipótese $H_0 : \beta_1 = \mathbf{0}_p$, que envolve p parâmetros simultaneamente, exige um teste conjunto adequado, sendo o teste F apropriado para esta tarefa.

2.4 Derivação via Decomposição de Soma de Quadrados e Teorema de Cochran

Sob $H_0 : \beta_1 = \mathbf{0}_p$, o modelo reduzido é $Y_i = \beta_0 + \varepsilon_i$, cujo E.M.Q. é $\hat{\beta}_0 = \bar{y}$. A soma de quadrados do modelo reduzido é:

$$SSE_0 = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \mathbf{y}^T\mathbf{y} - n\bar{y}^2 = \mathbf{y}^T(\mathbf{I}_n - \mathbf{P}_0)\mathbf{y} \tag{8}$$

onde $\mathbf{P}_0 = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$ é a matriz de projeção no espaço gerado pelo vetor de uns.

A diferença $SSR = SSE_0 - SSE$ representa a redução na soma de quadrados devido à inclusão das p variáveis explicativas. Em termos de matrizes de projeção, temos:

$$SSR = \mathbf{y}^T(\mathbf{P} - \mathbf{P}_0)\mathbf{y}, \quad SSE = \mathbf{y}^T(\mathbf{I}_n - \mathbf{P})\mathbf{y}$$

onde $\mathbf{P} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ é a matriz de projeção no espaço coluna de $\mathbf{X}$.

Aplicando o Teorema de Cochran com as matrizes $\mathbf{Q}_1 = \mathbf{P} - \mathbf{P}_0$ e $\mathbf{Q}_2 = \mathbf{I}_n - \mathbf{P}$, verificamos que são simétricas, idempotentes, $\mathbf{Q}_1 + \mathbf{Q}_2 + \mathbf{P}_0 = \mathbf{I}_n$, e sob H_0 temos $\mathbf{Q}_1\mathbf{X}\beta = (\mathbf{P} - \mathbf{P}_0)\mathbf{X}\beta = \mathbf{0}$ quando $\beta_1 = \mathbf{0}_p$. Portanto, sob H_0:

$$\frac{SSR}{\sigma^2} \sim \chi^2_p, \quad \frac{SSE}{\sigma^2} \sim \chi^2_{n-p-1} \tag{9}$$

e essas quantidades são independentes, conforme o Teorema de Cochran. Portanto, a razão:

$$F = \frac{SSR/p}{SSE/(n-p-1)} = \frac{(SSR/\sigma^2)/p}{(SSE/\sigma^2)/(n-p-1)} \sim F_{p,n-p-1} \tag{10}$$

segue uma distribuição F com p e $n-p-1$ graus de liberdade. Valores grandes de F_{obs} (por exemplo, $F_{\text{obs}} > F_{p,n-p-1;\alpha}$) indicam rejeição de H_0, pois a variância explicada (MSR) é significativamente maior que a variância residual (MSE).

2.5 Região de Rejeição e Testes Complementares

Para nível de significância α, rejeitamos H_0 se $F > F_{p,n-p-1;\alpha}$ ou se p-valor $= P(F_{p,n-p-1} > F_{\text{obs}}) < \alpha$.

Para testes individuais $H_0 : \beta_j = 0$, utiliza-se $t_j = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \sim t_{n-p-1}$, que complementa o teste F global para identificar quais variáveis específicas são significativas.

3 Análise de Variância e Conclusão

3.1 Decomposição ANOVA

A decomposição fundamental da variabilidade total é $SST = SSR + SSE$, onde $SST = \sum_{i=1}^{n}(y_i - \bar{y})^2$ é a Soma Total dos Quadrados, representando a variabilidade total da variável resposta. Sob H_0, temos $\frac{SSR}{\sigma^2} \sim \chi^2_p$ e $\frac{SSE}{\sigma^2} \sim \chi^2_{n-p-1}$ independentes, fundamentando teoricamente o teste F.

A tabela ANOVA resume essa decomposição:

Fonte	Soma de Quadrados	Graus de Liberdade	Quadrado Médio	F
Regressão	SSR	p	$MSR = SSR/p$	$F = MSR/MS$
Erro	SSE	$n-p-1$	$MSE = SSE/(n-p-1)$	
Total	SST	$n-1$		

Tabela 1: Tabela ANOVA para Regressão Linear Múltipla

O coeficiente de determinação $R^2 = SSR/SST$ mede a proporção da variabilidade total explicada pelo modelo. Valores próximos de 1 indicam bom ajuste do modelo aos dados.

3.2 Interpretação e Considerações Finais

O teste $H_0 : \beta_1 = \mathbf{0}_p$ permite avaliar a significância global das variáveis explicativas no modelo. Rejeitar H_0 significa que pelo menos uma das variáveis explicativas contribui de forma estatisticamente significativa para explicar a variabilidade de Y.

O teste F global deve ser complementado por testes t individuais para identificar quais variáveis específicas são responsáveis pela significância. A validade da inferência requer verificação cuidadosa dos pressupostos clássicos; a violação deles compromete a validade dos testes F e t. Além disso, a significância estatística não implica necessariamente significância prática, sendo importante considerar também o tamanho do efeito e o contexto do problema.

Referências Bibliográficas

- Casella, G. & Berger, R. L. (2002). *Statistical Inference.* 2nd ed. Duxbury Press.
- Montgomery, D. C., Peck, E. A. & Vining, G. G. (2012). *Introduction to Linear Regression Analysis.* 5th ed. Wiley.
- Seber, G. A. F. & Lee, A. J. (2012). *Linear Regression Analysis.* 2nd ed. Wiley.