

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon} \quad \underline{\varepsilon} \sim N(\underline{0}, \sigma^2 I)$$

$$\frac{SSE}{\sigma^2} \sim \chi^2_{n-k-1}$$

$$\frac{SSR}{\sigma^2} \sim \chi^2_{k, \lambda} \quad \lambda = \frac{1}{2\sigma^2} \underline{\beta}^{*'} X_1' \left[I - \frac{\underline{1}\underline{1}'}{n} \right] X_1 \underline{\beta}^*$$

$$X = \begin{bmatrix} \underline{1} & X_1 \end{bmatrix} \quad \underline{\beta}^* = [\beta_1 \ \beta_2 \dots \beta_k]'$$

SSR e SSE são variáveis aleatórias independentes.

Tabela de Análise de Variância

Fonte de Variação	GL	SA	MS	F
Regressão	k	$\hat{\underline{\beta}}' X' \underline{y} - n\bar{y}^2$	$MSR = \frac{SSR}{k}$	$\frac{MSR}{MSE}$
Resíduos	$n-k-1$	$\underline{y}' \underline{y} - \hat{\underline{\beta}}' X' \underline{y}$	$MSE = \frac{SSE}{n-k-1}$	
Total	$n-1$	$\underline{y}' \underline{y} - n\bar{y}^2$		

$$F = \frac{MSR}{MSE} = \frac{\frac{SSR}{k\sigma^2}}{\frac{SSE}{(n-k-1)\sigma^2}} \sim F_{k, n-k-1, \lambda}$$

Sob a hipótese de que $\beta^* = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} = \underline{0}$, $\lambda = 0$

$$e \quad F = \frac{MSR}{MSE} \sim F_{k, n-k-1}$$

Rejeitamos $H_0: \beta^* = \underline{0}$ ($\beta_1 = \beta_2 = \dots = \beta_k = 0$) se

$$F \geq F_c \quad P(F_{k, n-k-1} \geq F_c) = \alpha.$$

Este teste verifica a importância da inclusão de $X_1, X_2, \dots, X_k$ no modelo.

Para o cálculo do poder, $\beta^* \neq \underline{0}$ e $\lambda \neq 0$ e seria necessário utilizar a distribuição F não central.

Estudo do Modelo sem Intercepto

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \epsilon_i \quad i = 1, 2, \dots, n$$

$$\underline{y} = X \underline{\beta} + \underline{\epsilon} \quad X = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{k1} \\ x_{12} & x_{22} & & x_{k2} \\ \vdots & \vdots & & \vdots \\ x_{1n} & x_{2n} & & x_{kn} \end{bmatrix}_{n \times k} \quad \underline{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}_{k \times 1}$$

$$\hat{\underline{\beta}} = (X'X)^{-1} X' \underline{y} \quad SSE = (\underline{y} - X \hat{\underline{\beta}})' (\underline{y} - X \hat{\underline{\beta}}) = \underline{y}' \underline{y} - \hat{\underline{\beta}}' X' \underline{y}$$

Se $\beta_1 = \beta_2 = \dots = \beta_K = 0$ o modelo é $\hat{y}_i = 0$ e

$$SST_{mc} = \sum_{i=1}^n (y_i - 0)^2 = \sum_{i=1}^n y_i^2 \rightarrow \text{Soma de Quadrados Total não corrigida pela média}$$

$$SSR_{mc} = SST_{mc} - SSE = \underline{y}' \underline{y} - (\underline{y}' \underline{y} - \hat{\beta}' X' \underline{y}) = \hat{\beta}' X' \underline{y}$$

$$\frac{SSE}{\sigma^2} \sim \chi_{n-K}^2 \quad n[I - X(X'X)^{-1}X'] = n - K$$

$$\lambda = \frac{1}{2\sigma^2} \beta' X' [I - X(X'X)^{-1}X'] X \beta = 0$$

$$\frac{SSR_{mc}}{\sigma^2} = \frac{\underline{y}' X (X'X)^{-1} X' \underline{y}}{\sigma^2} \quad A = X(X'X)^{-1}X' \text{ é idempotente}$$

$$n[X(X'X)^{-1}X'] = \text{tr}[X(X'X)^{-1}X'] = \text{tr}[X'X(X'X)^{-1}] = \text{tr}(I_K) = K$$

$$\lambda = \frac{1}{2\sigma^2} \underbrace{\beta' X' X (X'X)^{-1} X' X \beta}_{\frac{1}{2} \underline{\mu}' A \underline{\mu}} \quad A = \frac{X(X'X)^{-1}X'}{\sigma^2}$$

$$\frac{1}{2} \underline{\mu}' A \underline{\mu}$$

$$= \frac{1}{2\sigma^2} \beta' X' X \beta$$

$$\therefore \frac{SSR_{mc}}{\sigma^2} \sim \chi_K^2, \frac{1}{2\sigma^2} \beta' X' X \beta$$

$$A \Sigma B = \frac{I - X(X'X)^{-1}X'}{\sigma^2} \sigma^2 I X (X'X)^{-1} X' = 0$$

$\Rightarrow$ SSE e SSR_{nc} são independentes

$$\beta = 0 \iff \frac{SSR_{nc}}{\sigma^2} \sim \chi^2_k$$

ANOVA

FV	GL	SS	MS	F
regressão	K	$\hat{\beta}'X'Y$	$\hat{\beta}'X'Y/K$	$\frac{MSR}{MSE}$
resíduos	n-K	$Y'Y - \hat{\beta}'X'Y$	$(Y'Y - \hat{\beta}'X'Y)/(n-K)$	
total	n	$Y'Y$		

Sob $H_0: \beta = 0$ $\frac{MSR}{MSE} \sim F_{K, n-K}$

$\beta_1 = \beta_2 = \dots = \beta_K = 0$

Rejeita-se H_0 se $\frac{MSR}{MSE} \geq F_c$ $F_c \mid P(F \geq F_c) = \alpha$

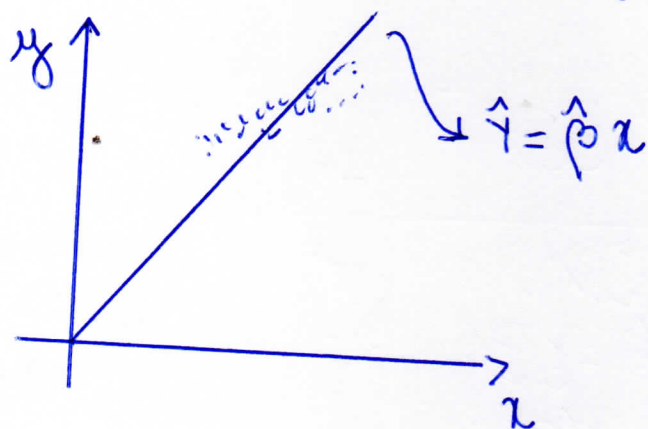
$F \sim F_{\text{snedecor}}$ K gl no numerador e n-K no denom.

Considerações sobre o modelo sem intercepto

Coefficiente de explicação $R^2 = \frac{\hat{\beta}' X' Y}{Y' Y}$

$$\hat{Y}_i = \hat{\beta} x_i \quad x=0 \Rightarrow \hat{Y}=0$$

Reta passa pela origem



Só utilizar modelo sem intercepto em situações em que $E(Y|x=0) = 0$ e o conjunto de dados tiver $x=0$.

Significado prático de β_0 :

$$Y = \beta_0 + \beta_1 x + \epsilon$$

↓
custo

↓
quantidade

β_0 : custo fixo

Considerações sobre R^2 (modelo com intercepto)

$$R^2 = \frac{SSR}{SST} \quad 0 \leq R^2 \leq 1$$

Tende a ser alto quando n é pequeno.

Mede uma tendência de crescimento que pode não ser linear:

Para verificar a hipótese de linearidade:

- teste de falta de ajustamento
- análise de resíduos

R^2 sempre aumenta com o acréscimo de uma nova variável preditora, mesmo que não seja tão importante.

Para levar em conta o tamanho da amostra e o número de variáveis explicativas pode-se utilizar

$$R^2_{aj} = 1 - \frac{n-1}{n-k-1} (1 - R^2)$$

Pode diminuir com o acréscimo de variáveis pouco importantes.

Exemplo

$$Y_i = \beta x_i + \epsilon_i \quad \epsilon_i \sim N(0, \sigma^2) \quad \text{v.a. independentes}$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \beta = [\beta]$$

$$\hat{\beta} = (X'X)^{-1}X'Y = \left(\sum x_i^2\right)^{-1} \sum x_i Y_i = \frac{\sum x_i Y_i}{\sum x_i^2} \quad \text{Var}(\hat{\beta}) = (X'X)^{-1}\sigma^2 = \frac{\sigma^2}{\sum x_i^2}$$

$$SST_{mc} = \sum Y_i^2$$

$$SSR_{mc} = \hat{\beta}'X'Y = \frac{\sum x_i Y_i}{\sum x_i^2} \sum x_i Y_i = \frac{(\sum x_i Y_i)^2}{\sum x_i^2}$$

$$SSE = Y'Y - \hat{\beta}'X'Y = \sum Y_i^2 - \frac{(\sum x_i Y_i)^2}{\sum x_i^2}$$

Tabela de Análise de Variância

FV	GL	SS	F
regressão	1	SSR_{mc}	$\frac{SSR_{mc}}{\frac{SSE}{n-1}} = F$
resíduos	$n-1$	SSE	
total	n	SST_{mc}	

$$H_0: \beta = 0$$

Rejeita-se H_0 se $F > F_c$ F_c quantil de ordem $1-\alpha$ da distribuição F 1 gl no numerador e $n-1$ no denominador.

Testes de Hipóteses

$$Ex: y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \epsilon$$

Hipóteses

$$1) H_0: \beta_2 = 0 \quad \beta_4 = 0 \quad \Leftrightarrow C\beta = m$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} \quad C = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad m = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Seja H_0 , o modelo é $y = \beta_0 + \beta_1 x_1 + \beta_3 x_3 + \epsilon$.

$$2) H_0: \beta_1 + \beta_3 = 0 \quad \Leftrightarrow C\beta = m$$

$$C = [0 \quad 1 \quad 0 \quad 1 \quad 0] \quad m = [0]$$

Modelo para H_0 : $y = \beta_0 + \beta_2 x_2 + \beta_4 x_4 + \beta_1 (x_1 - x_3) + \epsilon$

$$3) H_0: \beta_2 = 4 \quad \beta_3 = 5 \quad \Leftrightarrow C\beta = m$$

$$C = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad m = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

$$4) H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 \Leftrightarrow C \underline{\beta} = \underline{m}$$

$$C = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 \end{bmatrix} \quad \underline{m} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\beta_1 - \beta_2 = 0$$

$$\beta_1 - \beta_3 = 0$$

$$\beta_1 - \beta_4 = 0$$

Obs: As matrizes C e $\underline{m}$ não são necessariamente únicas.
No exemplo 4, $H_0: C_1 \underline{\beta} = \underline{m}_1$

$$C_1 = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \quad \underline{m}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\beta_1 - \beta_2 = 0$$

$$\beta_2 - \beta_3 = 0$$

$$\beta_3 - \beta_4 = 0$$

Teste de Hipóteses

Testar $H_0: C\beta = \underline{m}$ contra C -matriz $p \times (k+1)$
 $H_a: C\beta \neq \underline{m}$ de constantes, $r(C) = p$
 $\leq k+1$

no modelo $\underline{Y} = X\beta + \underline{\epsilon}$ $\underline{m}$ vetor $p \times 1$
 $\underline{\epsilon} \sim N(0, \sigma^2 I)$

Teste da Razão de Verossimilhança Generalizada

$L(\theta, x_1, \dots, x_n)$ função de verossimilhança de uma amostra $X_1, X_2, \dots, X_n$, $\theta \in \mathcal{H}$. Para testar

$H_0: \theta \in \mathcal{H}_0 \subset \mathcal{H}$ contra $H_a: \theta \in \mathcal{H} - \mathcal{H}_0$, a estatística

$$\Lambda = \lambda(x_1, x_2, \dots, x_n) = \frac{\sup_{\theta \in \mathcal{H}_0} L(\theta, x_1, x_2, \dots, x_n)}{\sup_{\theta \in \mathcal{H}} L(\theta, x_1, \dots, x_n)}$$

é chamada razão de verossimilhança generalizada

O teste da razão de verossimilhança generalizada rejeita H_0 se $\Lambda \leq \lambda_0$, sendo que $0 \leq \lambda_0 \leq 1$ e tal que

$$P(\Lambda(x_1, x_2, \dots, x_n) \leq \lambda_0 \mid H_0 \text{ é verdadeira}) = \alpha.$$

$\Lambda \leq 1$ Pequenos valores de Λ indicam que $\theta \in \mathcal{H}_0$ é improvável.

O teste da razão de verossimilhanças generalizada rejeita H_0 se $\Lambda \leq \lambda_0$, onde $0 \leq \lambda_0 \leq 1$ e tal que $P(\Lambda(X_1, X_2, \dots, X_n) | H_0 \text{ é verdadeira}) = \alpha$.

$$L(\beta, \sigma^2 | \underline{y}, X) = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left[-\frac{1}{2\sigma^2} (\underline{y} - X\beta)' (\underline{y} - X\beta) \right]$$

$$\textcircled{H} = \{(\beta, \sigma^2), -\infty < \beta_j < \infty, j=0, 1, 2, \dots, k, \sigma^2 > 0\}$$

$$\textcircled{H}_0 = \{(\beta, \sigma^2) \in \textcircled{H} \mid C\beta = \underline{m}, \sigma^2 > 0\}$$

$$\Lambda = \frac{\sup_{(\beta, \sigma^2) \in \textcircled{H}_0} L(\beta, \sigma^2 | \underline{y}, X)}{\sup_{(\beta, \sigma^2) \in \textcircled{H}} L(\beta, \sigma^2 | \underline{y}, X)}$$

O denominador é $L(\hat{\beta}, \hat{\sigma}_{mv}^2 | \underline{y}, X)$

$$\hat{\beta} = (X'X)^{-1} X' \underline{y}$$

$$\hat{\sigma}_{mv}^2 = \frac{(\underline{y} - X\hat{\beta})' (\underline{y} - X\hat{\beta})}{n}$$

Cálculo do numerador

Maximizar $L(\beta, \sigma^2 | \underline{y}, X)$ sujeito à condição

$$C\beta = \underline{m}.$$

$$\hat{\sigma}_0^2 = \frac{(\underline{y} - X\hat{\beta}_0)'(\underline{y} - X\hat{\beta}_0)}{n}$$

$$\hat{\beta}_0 = \hat{\beta} - (X'X)^{-1}C'[C(X'X)^{-1}C']^{-1}(C\hat{\beta} - \underline{m})$$

Estes estimadores coincidem com os estimadores de máxima verossimilhança no modelo reduzido sob H_0 .

$$\Lambda = \left(\frac{\hat{\sigma}_{mv}^2}{\hat{\sigma}_0^2} \right)^{n/2} \rightarrow \text{estatística de teste}$$

Região Crítica: $\Lambda \leq \lambda_0$

Usar a estatística

$$W = \left(\Lambda^{-2/n} - 1 \right) \frac{n-k-1}{p}$$

devido à facilidade de obtenção da distribuição.

W é uma função monótona de Λ .

$$\Lambda \leq \lambda_0 \Leftrightarrow \left(\Lambda^{-2/n} - 1 \right) \frac{n-k-1}{p} \geq \left(\lambda_0^{-2/n} - 1 \right) \frac{n-k-1}{p}$$

$$\Leftrightarrow W \geq w \quad \text{forma da região crítica}$$

Após cálculos

$$W = \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}_{mv}^2} - 1 \right) \frac{n-k-1}{p} = \frac{\hat{\sigma}_0^2 - \hat{\sigma}_{mv}^2}{\hat{\sigma}_{mv}^2} \frac{n-k-1}{p}$$

$$W = \frac{(C\hat{\beta} - \underline{m})' [C(X'X)^{-1}C']^{-1} (C\hat{\beta} - \underline{m})}{p \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{n-k-1}}$$

↪ MSE do modelo inicial

Distribuição de W

$$\frac{(C\hat{\beta} - \underline{m})' [C(X'X)^{-1}C']^{-1} (C\hat{\beta} - \underline{m})}{\sigma^2} \sim \chi^2_{p,2}$$

$$\hat{\beta} \sim N_{k+1}(\underline{\beta}, (X'X)^{-1}\sigma^2)$$

$$(C\hat{\beta} - \underline{m}) \sim N_p(C\underline{\beta} - \underline{m}, C(X'X)^{-1}C'\sigma^2)$$

$$A \Sigma = \frac{[C(X'X)^{-1}C']^{-1} C(X'X)^{-1}C'\sigma^2}{\sigma^2} = I$$

$\therefore A \Sigma$ é idempotente

$$\lambda = \frac{1}{2} (C\hat{\beta} - \underline{m})' \frac{[C(X'X)^{-1}C']^{-1}}{2} (C\hat{\beta} - \underline{m})$$

$$\frac{(Y - X\hat{\beta})' (Y - X\hat{\beta})}{\sigma^2} \sim \chi^2_{n-k-1}$$

$\hat{\beta}$ e SSE são independentes

$$\therefore W \sim F_{p, n-k-1, \lambda}$$

Sob H_0 , $C\hat{\beta} = \underline{m}$, $\lambda = 0$, $W \sim F_{p, n-k-1}$ central

TRVG RC: $W \geq w$

$$w \mid P(F_{p, n-k-1} \geq w) = \alpha$$

Observações:

$$1) W = \frac{\hat{\sigma}_0^2 - \hat{\sigma}_{mv}^2}{\hat{\sigma}_{mv}^2} \frac{n-k-1}{p}$$

$\hat{\sigma}_{mv}^2$ - estimador de max. veross. de σ^2 no modelo inicial

$\hat{\sigma}_0^2$ - estimador de max. veross de σ^2 no modelo reduzido sob H_0 .