

$$H_0: C\beta = m$$

$$C \text{ } p \times (k+1)$$

$$H_a: C\beta \neq m$$

Estatística de Teste:

$$W = \frac{(C\hat{\beta} - m)' [C(X'X)^{-1}C']^{-1} (C\hat{\beta} - m)}{p \text{ MSE}}$$

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

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

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

Rejeita-se H_0 se $W \geq w$, $w \mid$

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

Exemplos

$$1) H_0: \underset{\sim}{l}'\beta = b$$

$$\underset{\sim}{l}' = [l_0 \ l_1 \ l_2 \ \dots \ l_k]$$

$$H_a: \underset{\sim}{l}'\beta \neq b$$

$$p = 1$$

$$W = \frac{(\underline{l}'\hat{\beta} - b)' [\underline{l}'(X'X)^{-1}\underline{l}]^{-1} (\underline{l}'\hat{\beta} - b)}{MSE} =$$

$$= \frac{(\underline{l}'\hat{\beta} - b)^2}{[\underline{l}'(X'X)^{-1}\underline{l}] \hat{\sigma}^2} = \frac{(\underline{l}'\hat{\beta} - b)^2}{\widehat{\text{Var}}(\underline{l}'\hat{\beta})}$$

Rejeitamos H_0 se $W \geq F_{2, 1, n-k-1} = t_{\alpha/2, n-k-1}^2$

$$\Leftrightarrow \frac{\underline{l}'\hat{\beta} - b}{\sqrt{\widehat{\text{Var}}(\underline{l}'\hat{\beta})}} \geq t_{\alpha/2, n-k-1} \text{ ou } \frac{\underline{l}'\hat{\beta} - b}{\sqrt{\widehat{\text{Var}}(\underline{l}'\hat{\beta})}} \leq -t_{\alpha/2, n-k-1}$$

Esta é a RC em termos da estatística t .
Esta construção permite fazer o teste unicaudal.

Equivalentemente, não rejeitamos H_0 se e só se

$$b \in \left[\underline{l}'\hat{\beta} - t_{\alpha/2, n-k-1} \sqrt{\hat{\sigma}^2 \underline{l}'(X'X)^{-1}\underline{l}}, \underline{l}'\hat{\beta} + t_{\alpha/2, n-k-1} \sqrt{\hat{\sigma}^2 \underline{l}'(X'X)^{-1}\underline{l}} \right]$$

↓

IC com coeficiente $1 - \alpha$ para $\underline{l}'\beta$

Se $\underline{l} = [0 \ 0 \ \dots \ 1 \ \dots \ 0]$ e $b=0$, o teste fica

$$H_0: \beta_j = 0$$

$$H_a: \beta_j \neq 0$$

$$W = \frac{\hat{\beta}_j^2}{a_{jj} \hat{\sigma}^2}$$

a_{jj} - j -ésimo elemento da diagonal de

$$(X'X)^{-1} = \begin{bmatrix} a_{00} & a_{01} & \dots & a_{0k} \\ & a_{11} & & \\ & & & \\ & & & a_{kk} \end{bmatrix}$$

Rejeitamos H_0 se $W \geq F_{\alpha, 1, n-k-1}$

Além disso,

$$W = \frac{SSE_R - SSE_C}{SSE_C} \frac{n-k-1}{p} \stackrel{p=1}{=} \frac{SSE_R - SSE_C}{MSE_C}$$

Modelo Completo: $Y = \beta_0 + \beta_1 x_1 + \dots + \beta_j x_j + \beta_{j+1} x_{j+1} + \dots + \beta_k x_k + E$

Modelo Reduzido: $Y = \beta_0 + \beta_1 x_1 + \dots + \beta_{j-1} x_{j-1} + \beta_{j+1} x_{j+1} + \dots + \beta_k x_k + E$

Como aqui a redução sob H_0 não altera a variável dependente

$$W = \frac{SSR_C - SSR_R}{MSE_C} \rightarrow \text{Estatística do Teste F-parcial}$$

Intervalo de confiança para β_j :

$$\hat{\beta}_j \pm t_{\alpha/2, n-k-1} \sqrt{a_{jj} MSE}$$

$$2) \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \quad \begin{matrix} \beta_1 & (k+1-q) \times 1 \\ \beta_2 & q \times 1 \end{matrix} \quad 0 < q \leq k+1$$

$$H_0: \beta_2 = \underline{b}_2 \rightarrow \text{vector de constantes}$$

$$H_a: \beta_2 \neq \underline{b}_2$$

$$X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$$

$m \times (k+1-q) \quad m \times q$

$$X'X = \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}$$

$\begin{matrix} \xrightarrow{k+1-q} & \xrightarrow{q} \\ \xrightarrow{q} & \xrightarrow{k+1-q} \end{matrix}$

$$y = X\beta + \epsilon = \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \epsilon = X_1\beta_1 + X_2\beta_2 + \epsilon$$

$$C = \begin{bmatrix} 0 & I_q \end{bmatrix}$$

$q \times (k+1-q)$

$$\underline{m} = \underline{b}_2 \rightarrow C\beta = \beta_2$$

$$(X'X)^{-1} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$$

$\begin{matrix} (k+1-q) & (k+1-q) \times q \\ q \times (k+1-q) & q \times q \end{matrix}$

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$$

$$C\hat{\beta} - \underline{m} = \hat{\beta}_2 - \underline{b}_2$$

$$C(X'X)^{-1}C' = \begin{bmatrix} 0 & I_q \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} 0 \\ I_q \end{bmatrix}$$

$\stackrel{(X'X)^{-1}}{= T_{22}}$

$$W = \frac{(\hat{\beta}_2 - \underline{\beta}_2)' T_{22}^{-1} (\hat{\beta}_2 - \underline{\beta}_2)}{q \text{ MSE}}$$

$$X'X = \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1' X_1 & X_1' X_2 \\ X_2' X_1 & X_2' X_2 \end{bmatrix}$$

$$(X'X)^{-1} = \begin{bmatrix} - & - \\ - & \underbrace{\left[X_2' X_2 - X_2' X_1 (X_1' X_1)^{-1} X_1' X_2 \right]^{-1}}_{\parallel T_{22}} \end{bmatrix} \left[B_{22} - B_{21} B_{11}^{-1} B_{12} \right]^{-1}$$

$$T_{22}^{-1} = X_2' X_2 - X_2' X_1 (X_1' X_1)^{-1} X_1' X_2$$

$H_0: \beta_2 = 0$ (verifica a importância das q últimas variáveis

$H_a: \beta_2 \neq 0$ independentes)

$$W = \frac{\hat{\beta}_2' T_{22}^{-1} \hat{\beta}_2}{q \text{ MSE}}$$

Rejeitamos H_0 para $W > w$

$$P(F_{q, n-k-1} > w) = \alpha$$

$$3) H_0: \beta^* = 0$$

$$\beta^* = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}$$

Caso particular do anterior

$$q = k$$

$$\beta_1 = \beta_0$$

Verificar que, neste caso, $W = \frac{MSR}{MSE}$

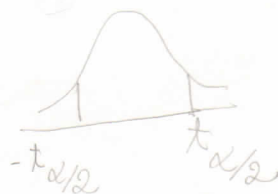
k) Intervalos de Confiança e de Previsão

Intervalo de Confiança para $E(Y_i | X_i) = X_i \beta$

$$X_i \beta = \tilde{l}' \beta \quad \text{para } \tilde{l}' = X_i$$

Portanto, o IC com coeficiente $1 - \alpha$ é

$$\tilde{l}' \hat{\beta} \pm t_{n-k-1, \alpha/2} \sqrt{\tilde{l}' (X'X)^{-1} \tilde{l} \hat{\sigma}^2}$$



Intervalo de Previsão para $Y|X_h$ quando a observação h não pertence à amostra

Definição

Seja T uma variável aleatória com valor desconhecido. A variável aleatória $\hat{T}$ é um previsor não viciado para T se

$$E(\hat{T} - T) = 0.$$

O erro quadrático médio de $\hat{T}$ é dado por

$$EQM(\hat{T}) = E(\hat{T} - T)^2 = \text{Var}(\hat{T} - T) =$$

↓
Se $\hat{T}$ é não viciado

$$\text{Var}(\hat{T} - T) = E[(\hat{T} - T)^2] - \underbrace{E^2(\hat{T} - T)}_{= 0} = E[(\hat{T} - T)^2]$$

$$\hat{Y}_h = \tilde{X}_h \hat{\beta} \rightarrow \text{previsores de } Y_h$$

$$E(\hat{Y}_h - Y_h) = \tilde{X}_h E(\hat{\beta}) - E(Y_h) = \tilde{X}_h \beta - \tilde{X}_h \beta = 0$$

$\therefore \hat{Y}_h$ é um previsor não viciado de Y_h .

Intervalo de Previsão para $Y|X_h$

$$\hat{Y}_h \pm t_{n-k-1, \alpha/2} \sqrt{\text{Var}(\hat{Y}_h - Y_h)} \\ \downarrow \text{EQM}(\hat{Y}_h)$$

$$\text{Var}(\hat{Y}_h) = \text{Var}(X_h \hat{\beta}) = X_h (X'X)^{-1} X_h' \sigma^2$$

$$\text{Var}(Y_h) = \sigma^2$$

$\hat{Y}_h$ e Y_h são independentes porque o elemento h não pertence à amostra $\Rightarrow$

$$\text{Var}(\hat{Y}_h - Y_h) = \text{Var}(\hat{Y}_h) + \text{Var}(Y_h) = X_h (X'X)^{-1} X_h' \sigma^2 + \sigma^2$$

Intervalo de Previsão

$$X_h \hat{\beta} \pm t_{n-k-1, \alpha/2} \sqrt{(1 + X_h (X'X)^{-1} X_h') \hat{\sigma}^2}$$

Ex: No modelo de regressão linear simples

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad \hat{Y}_l = \hat{\beta}_0 + \hat{\beta}_1 x_l$$

IC:

$$\hat{Y}_l \pm t_{n-k-1, \alpha/2} \sqrt{\left(\frac{1}{n} + \frac{(x_l - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right) \text{MSE}}$$

I Previsão

$$\hat{Y}_l \pm t_{n-k-1, \alpha/2} \sqrt{\left(1 + \frac{1}{n} + \frac{(x_l - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right) \text{MSE}}$$

Modelo Linear Geral $\underline{\epsilon} \sim N_n(\underline{0}, \sigma^2 V)$

$V \neq I$ conhecida

$$\underline{y} = X\underline{\beta} + \underline{\epsilon}$$

$$\underline{\epsilon} \sim N_n(\underline{0}, \sigma^2 V)$$

V matriz positiva definida ($V \neq I$) conhecida

V diagonal, $V \neq I \Rightarrow$ heteroscedasticidade

V não diagonal $\Rightarrow \epsilon_i, \epsilon_j$ correlacionados, $i \neq j$
 $i, j = 1, 2, \dots, n$

V positiva definida $\Rightarrow \exists P$ não singular | $V = PP'$

$$\underline{z} = P^{-1}\underline{y} \sim N_n(P^{-1}X\underline{\beta}, \sigma^2 I) \text{ pois}$$

$$\begin{aligned} \text{Var}(\underline{z}) &= \text{Var}(P^{-1}\underline{y}) = P^{-1} \text{Var}(\underline{y}) P^{-1'} \sigma^2 = \\ &= P^{-1} P P' P^{-1'} \sigma^2 = \sigma^2 I \end{aligned}$$

$$P \text{ e } A^{-1} = A^{-1}$$

Transformamos o modelo

$$Y = X\beta + \varepsilon \quad \varepsilon \sim N(0, \sigma^2 V) \quad \text{em}$$

$$P^{-1}Y = P^{-1}X\beta + P^{-1}\varepsilon$$

$$Z = Q\beta + \eta \quad Z = P^{-1}Y \quad Q = P^{-1}X \quad \eta = P^{-1}\varepsilon$$

$$\eta \sim N_n(0, \sigma^2 I)$$

O modelo transformado satisfaz às condições do modelo linear geral, caso 1 ($\varepsilon \sim N_n(0, \sigma^2 I)$)

Os parâmetros de interesse β e σ^2 do modelo original são os parâmetros do modelo transformado.

O estimador ~~de~~ ótimo de β é $\hat{\beta} = (Q'Q)^{-1}Q'Z$, estimador de m.q. (e de m.v.) de β no modelo transformado.

$$\hat{\beta} = (Q'Q)^{-1}Q'Z = \left(X' \underbrace{P^{-1'}}_{V^{-1}} P^{-1} X \right)^{-1} X' \underbrace{P^{-1}}_{V^{-1}} Y$$

$$= (X'V^{-1}X)^{-1}X'V^{-1}Y$$

$$V = PP'$$

$$V^{-1} = P'^{-1}P^{-1} = P'^{-1}P^{-1}$$

$$\begin{aligned}\text{Var}(\hat{\beta}) &= \text{Var}[(X'V^{-1}X)^{-1}X'V^{-1}\underline{y}] = \\ &= (X'V^{-1}X)^{-1}X'V^{-1}V V^{-1}X(X'V^{-1}X)^{-1}\sigma^2 = (X'V^{-1}X)^{-1}\sigma^2.\end{aligned}$$

$$\hat{\beta}_G = (X'V^{-1}X)^{-1}X'V^{-1}\underline{y} \rightarrow \text{estimador de mínimos quadrados generalizados de } \beta \text{ no modelo}$$

$$\underline{y} = X\beta + \underline{\varepsilon}$$

coincide com o estimador de min. quadrados usual $[(Q'Q)^{-1}Q'\underline{z}]$ do modelo transformado

$$\hat{\sigma}_G^2 = \frac{1}{n-k-1} (\underline{z} - Q\hat{\beta})'(\underline{z} - Q\hat{\beta}) = \frac{1}{n-k-1} (\underline{z}'\underline{z} - \hat{\beta}'Q'\underline{z})$$

↳ em função das variáveis transformadas

$$= \frac{1}{n-k-1} \underline{y}' [V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}] \underline{y}$$

↳ em função das variáveis originais

$\hat{\beta}_G$, $\text{Var}(\hat{\beta}_G)$ e $\hat{\sigma}_G^2$ são as quantidades necessárias para estimação pontual e por intervalo e testes de hipóteses.

$$H_0: C\beta = m$$

$$H_a: C\beta \neq m$$

$$W = \frac{(C\hat{\beta}_G - m)' [C(X'V^{-1}X)^{-1}C']^{-1} (C\hat{\beta}_G - m)}{p \hat{\sigma}_G^2}$$

Ols: Muitas vezes, a transformação $P^{-1}y = P^{-1}X\beta + P^{-1}\varepsilon$ pode resultar em um modelo sem intercepto, mesmo que o modelo de partida possua intercepto.

Exemplo

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad \varepsilon \sim N(0, \sigma^2 V)$$

$$V = \text{diag}(w_1, w_2, \dots, w_n) \Rightarrow \text{Var}(\varepsilon_i) = w_i \sigma^2$$

conhecida, $w_i > 0$ $\varepsilon_i, \varepsilon_j$ var. a. indep $i \neq j$

$$V = P P' \Rightarrow P = \text{diag}(\sqrt{w_1}, \sqrt{w_2}, \dots, \sqrt{w_n})$$

$$P^{-1} = \text{diag}\left(\frac{1}{\sqrt{w_1}}, \frac{1}{\sqrt{w_2}}, \dots, \frac{1}{\sqrt{w_n}}\right)$$

$$P^{-1}y = \begin{bmatrix} y_1/\sqrt{w_1} \\ y_2/\sqrt{w_2} \\ \vdots \\ y_n/\sqrt{w_n} \end{bmatrix} \quad P^{-1}X = \begin{bmatrix} 1/\sqrt{w_1} & x_1/\sqrt{w_1} \\ 1/\sqrt{w_2} & x_2/\sqrt{w_2} \\ \vdots & \vdots \\ 1/\sqrt{w_n} & x_n/\sqrt{w_n} \end{bmatrix}$$

Modelo Transformado:

$$\frac{Y_i}{\sqrt{w_i}} = \frac{\beta_0}{\sqrt{w_i}} + \beta_1 \frac{x_i}{\sqrt{w_i}} + \frac{\epsilon_i}{\sqrt{w_i}} \quad \text{sem intercepto}$$

$$SST_{nc} = \sum_{i=1}^n \frac{y_i^2}{w_i}$$

Usar a estatística W normalmente.

Se usar ANOVA, trabalhar com SST_{nc} e SSR_{nc}
(ver ANOVA para modelo sem intercepto)

Inferência $\underline{\epsilon} \sim N_n(\underline{0}, \underline{\Sigma})$ $\underline{\Sigma}$ desconhecida

$$\underline{y} = \underline{X}\underline{\beta} + \underline{\epsilon} \quad \underline{\epsilon} \sim N_n(\underline{0}, \underline{\Sigma})$$

$\underline{\Sigma}$ positiva definida ($\underline{\Sigma} \neq \underline{I}$) desconhecida

$$\hat{\underline{\beta}}_G = (\underline{X}' \underline{\Sigma}^{-1} \underline{X})^{-1} \underline{X}' \underline{\Sigma}^{-1} \underline{y} \quad \text{não é um estimador}$$



MAE 5725- Modelos Lineares

1) Um psicólogo está investigando a relação entre o tempo que o indivíduo leva para reagir a um certo estímulo (Y) e algumas de suas características tais como: sexo, idade e acuidade visual (medida em porcentagem).

	Y (seg)	W	X_1	X_2
indivíduo	tempo de reação	sexo	idade	acuidade visual
1	96	H	20	90
2	92	M	20	100
3	106	H	20	80
4	100	M	20	90
5	98	M	25	100
6	104	H	25	90
7	110	M	25	80
8	101	M	25	90
9	116	H	30	70
10	106	H	30	90
11	109	H	30	90
12	100	M	30	80
13	112	M	35	90
14	105	M	35	80
15	118	H	35	70
16	108	H	35	90
17	113	M	40	90
18	112	M	40	90
19	127	H	40	60
20	117	H	40	80

a) Obtenha o modelo de regressão ajustado

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$$

onde $\hat{\beta}_0$, $\hat{\beta}_1$ e $\hat{\beta}_2$ são as estimativas de mínimos quadrados de β_0 , β_1 e β_2 .

b) Construa a Tabela de Análise de Variância e calcule o coeficiente de explicação do modelo.

c) Admitindo que $Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i$, $\epsilon_i \sim N(0, \sigma^2)$, $i = 1, 2, \dots, 20$ e ϵ_i, ϵ_j v.a. independentes para $i \neq j$, teste a hipótese

$H_0: \beta_1 = \beta_2 = 0$ contra

H_a : Pelo menos um dos $\beta_l \neq 0$, $l = 1, 2$.