TÉCNICAS APLICADAS DE INVESTIGAÇÃO EM GESTÃO

1. Regressão linear múltipla – estimação e inferência

- Modelo de regressão linear múltipla
- Estimador dos mínimos quadrados
  - Interpretação dos parâmetros
  - Propriedades
- Inferência:
  - Intervalos de confiança
  - Testes de hipóteses para um só parâmetro
  - Testes de hipóteses para uma restrição linear dos parâmetros
  - Testes F

Motivação e interpretação

- **Modelo de Regressão Linear Simples**

1. Fertilidade: $Filhos = \beta_0 + \beta_1 educ + u$
2. Peso de um bebé: $peso = \beta_0 + \beta_1 cigs + u$
3. Salário: $sal = \beta_0 + \beta_1 educ + u$

- **Modelo de Regressão Linear Múltipla:** análise *ceteris paribus*

1. Fertilidade: $Filhos = \beta_0 + \beta_1 educ + \beta_2 idade + \beta_3 rend + u$
2. Peso de um bebé: $peso = \beta_0 + \beta_1 cigs + \beta_2 educ + \beta_3 nconsul + u$
3. Salário: $sal = \beta_0 + \beta_1 educ + \beta_2 exper + u$

O método dos mínimos quadrados

Formalização do Modelo de Regressão Linear Múltipla

O Modelo explica a variável y em função das variáveis $x_1, x_2, \ldots, x_k$

Termo independente

Coeficientes de declive

Parâmetros desconhecidos
A estimar

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + u$$

Variável dependente, Variável explicada, Variável de resposta
Observada

Variáveis independentes Variáveis explicativas, Regressores
Observadas

Variável residual
Erro
Não observado

O método dos mínimos quadrados (OLS)

- **Estimação OLS**
  - Amostra aleatória

$$\{(x_{i1}, x_{i2}, \ldots, x_{ik}, y_i) : \ i = 1, \ldots n\}$$

  - Valores ajustados: considere-se as estimativas $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k$

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \ldots + \hat{\beta}_k x_{ik}$$

  - Resíduos $\hat{u}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \ldots - \hat{\beta}_k x_{ik}$

  - Minimização da soma dos quadrados dos resíduos

$$\min \sum_{i=1}^{n} \hat{u}_i^2 \quad \rightarrow \quad \hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_k$$

O método dos mínimos quadrados (OLS)

- **Propriedades algébricas dos resíduos OLS**

$$\widehat{u}_i = y_i - \widehat{y}_i \quad \text{com} \quad \widehat{y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_{i1} + \widehat{\beta}_2 x_{i2} + \ldots + \widehat{\beta}_k x_{ik}$$

1. $\sum_{i=1}^{n} \widehat{u}_i = 0$

2. $\sum_{i=1}^{n} x_{ij} \widehat{u}_i = 0$

3. $\bar{y} = \widehat{\beta}_0 + \widehat{\beta}_1 \bar{x}_1 + \ldots + \widehat{\beta}_k \bar{x}_k$

As médias observadas na amostra da variável explicada e das variáveis explicativas estão na regressão

Estimação da variância da v. residual

- **Estimação da variância da variável residual (erro)**

- u é uma variável aleatória que verifica
  - $E(u) = 0$
  - $V(u) = \sigma^2$ ← Parâmetro a estimar

- Estimador da variância

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-k-1}$$

Graus de liberdade: $n-k-1$

Número de observações: n

Número de variáveis: k

Coeficiente de determinação

- **Coeficiente de determinação: medir a qualidade do ajustamento**

  - Decomposição da variação total (supondo modelo com termo independente)

$$SST = SSE + SSR$$

Variação Total: soma total dos quadrados

Variação explicada: soma explicada dos quadrados

Variação residual: soma dos quadrados dos resíduos

  - Coeficiente de determinação ou R^2

$$R^2 = SSE/SST = 1 - SSR/SST$$

Coeficiente de determinação

- **Observações**
  - $0 \leq R^2 \leq 1$
  - Sempre que se junta uma variável ao modelo o R^2 aumenta mesmo que o poder explicativo desta var. seja irrelevante
  - Não usar o R^2 para comparar a capacidade explicativa de modelos com
    1. diferente número de variáveis explicativas
    2. variáveis dependentes definidas em escalas muito diferentes

    Alternativa para R^2 na situação **1**: $\Rightarrow$ $\bar{R}^2$ ajustado
  - o R^2 é igual ao quadrado do coeficiente de correlação empírica entre os valores observados e os valores ajustados da variável dependente

Coeficiente de determinação

- **Observações** (continuação)
  - Um R^2 elevado não significa que exista realmente uma relação causal entre as variáveis
  - Um R^2 baixo não é necessariamente um sintoma de que os efeitos parciais das variáveis não sejam estimados com precisão
- Cálculo do $\bar{R}^2$

$$R^2 = 1 - \frac{SSR}{SST} = 1 - \frac{(SSR/n)}{(SST/n)}$$ é uma estimativa de $$1 - \frac{\sigma_u^2}{\sigma_y^2}$$

R^2 da população

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

Correção dos graus de liberdade do numerador e do denominador

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

Pode ser negativo

Forma funcional e interpretação dos parâmetros

- **Interpretação de $\hat{\beta}_j$ em termos de efeito parcial depois de removido o efeito das outras variáveis**

- Mostra-se que no MRLM $\hat{\beta}_j$ pode ser obtido em 2 passos:

1) Regressão de x_j em termos de todas as outras variáveis explicativas

$\Downarrow$

$\boxed{\text{expurga de } x_j \text{ a influência de todas as outras variáveis explicativas}}$

2) Regressão de y em função dos resíduos da regressão anterior

- Os resíduos da 1º regressão representam a parte de x_j que não depende das outras variáveis explicativas
- O coeficiente de declive da segunda regressão é igual a $\hat{\beta}_j$ e representa assim o impacto de x_j depois de expurgados (removidos) os efeitos das outras variáveis

Forma funcional e interpretação dos parâmetros

- Note-se que: $\beta_j = \dfrac{\partial y}{\partial x_j}$

Mede a variação de y quando x_j varia de uma unidade mantendo tudo o resto constante

- O MRLM permite manter na análise o valor das outras variáveis explicativas fixo ou constante mesmo que sejam correlacionadas com a variável em análise

Análise ceteris paribus

- Tem que se admitir ainda que o erro não varia quando x_j varia

Forma funcional e interpretação dos parâmetros

- **Interpretação de** $\hat{\beta}_j$

$$\hat{y}_I = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_k x_k \quad \text{e} \quad \hat{y}_F = \hat{\beta}_0 + \hat{\beta}_1 (x_1 + 1) + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_k x_k$$

$$\hat{y}_F - \hat{y}_I = \Delta\hat{y} =$$

$$\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_k x - \hat{\beta}_0 - \hat{\beta}_1 x_1 - \hat{\beta}_2 x_2 - \ldots - \hat{\beta}_k x$$

Então

$$\boxed{\hat{\beta}_j = \Delta\hat{y} \quad \text{if} \quad \Delta x_j = 1 \;\; ceteris\ paribus}$$

- **Variação total de** y

$$\Delta\hat{y} = \hat{\beta}_1 \Delta x_1 + \hat{\beta}_2 \Delta x_2 + \ldots + \hat{\beta}_k \Delta x_k$$

Exemplo

Dependent Variable: PRICE
Method: Least Squares
Included observations: 88

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-21.77031	29.47504	-0.738601	0.4622
ALOTE	0.022257	0.006912	3.220096	0.0018
ACASA	1.321574	0.142486	9.275093	0.0000
BDRMS	13.85252	9.010145	1.537436	0.1279

R-squared	0.672362	Mean dependent var	293.5460
Adjusted R-squared	0.660661	S.D. dependent var	102.7134
S.E. of regression	59.83348	Akaike info criterion	11.06540
Sum squared resid	300723.8	Schwarz criterion	11.17800
Log likelihood	-482.8775	Hannan-Quinn criter.	11.11076
F-statistic	57.46023	Durbin-Watson stat	2.109796
Prob(F-statistic)	0.000000		

Exemplo

Dependent Variable: LOG(PRICE)
Method: Least Squares
Sample: 1 88
Included observations: 88

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.765972	0.440114	1.740396	0.0855
LOG(ALOTE)	0.167967	0.038281	4.387714	0.0000
LOG(ACASA)	0.700232	0.092865	7.540306	0.0000
BDRMS	0.036958	0.027531	1.342415	0.1831

R-squared	0.642965	Mean dependent var	5.633180
Adjusted R-squared	0.630214	S.D. dependent var	0.303573
S.E. of regression	0.184603	Akaike info criterion	-0.496833
Sum squared resid	2.862563	Schwarz criterion	-0.384227
Log likelihood	25.86066	Hannan-Quinn criter.	-0.451467
F-statistic	50.42374	Durbin-Watson stat	2.088996
Prob(F-statistic)	0.000000		

Exemplo

Dependent Variable: PRICE
Method: Least Squares
Sample: 1 88
Included observations: 88

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1345.801	141.4617	-9.513543	0.0000
LOG(ALOTE)	61.45707	12.30436	4.994739	0.0000
LOG(ACASA)	224.9734	29.84882	7.537093	0.0000
BDRMS	19.35056	8.849135	2.186718	0.0315

R-squared	0.677797	Mean dependent var	293.5460
Adjusted R-squared	0.666290	S.D. dependent var	102.7134
S.E. of regression	59.33513	Akaike info criterion	11.04867
Sum squared resid	295735.3	Schwarz criterion	11.16128
Log likelihood	-482.1415	Hannan-Quinn criter.	11.09404
F-statistic	58.90179	Durbin-Watson stat	2.241032
Prob(F-statistic)	0.000000		

Hipóteses e Propriedades estatísticas

- Hipótese DS.1: Modelo linear nos parâmetros

Na população, a relação entre y e os Coeficientes é linear

$$\boxed{y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + u}$$

O modelo inclui uma variável residual (não observada)

- Hipótese DS.2: Amostragem aleatória

$$\{(x_{i1}, x_{i2}, \ldots, x_{ik}, y_i) : \ i = 1, \ldots n\}$$

Os dados são constituem uma amostra puramente aleatória da população

Cada observação verifica o modelo da população

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + u_i$$

Hipóteses e Propriedades estatísticas

- Hipótese DS.3: não existe multicolinearidade perfeita entre as variáveis explicativas
  - Nenhuma variável pode ser constante
  - Nenhuma variável pode ser uma combinação linear da outra

- Exemplo de multicolienaridade perfeita

$$voteA = \beta_0 + \beta_1 shareA + \beta_2 shareB + u$$

$voteA$: Votação em A num sistema bipartidário

$shareA$: Proporção de votos em A nas eleições anteriores

$shareB$: Proporção de votos em B nas eleições anteriores

$$\boxed{shareA+shareB = 1}$$

Hipóteses e Propriedades estatísticas

- Hipótese DS.4

O valor das variáveis explicativas não deve conter informação sobre a média do erro

$$E(u_i|x_{i1},x_{i2},\ldots,x_{ik})=0$$

$$\Downarrow$$

$$Corr(u,x_j)=0 \ \forall j=1,\ldots,k$$

  - Variáveis explicativas que são correlacionadas com o erro são consideradas endógenas
  - Variáveis explicativas que não são correlacionadas com o erro são consideradas exógenas
  - A hipótese DS.4 verifica-se quando todas as variáveis são exógenas

Hipóteses e Propriedades estatísticas

- **Teorema 1**

Sob as hipóteses DS.1 a DS.4 demonstra-se que o estimador OLS é **centrado**:

$$E\left(\hat{\beta}_j\right) = \beta_j \quad j = 1,\ldots,k$$

- Hipótese DS.5: Homocedasticidade

$$Var(u_i|x_{i1}, x_{i2}, \ldots, x_{ik}) = \sigma^2$$

O valor das variáveis explicativas não pode conter qualquer informação sobre a variância do erro

A variância do erro não pode depender do valor das variáveis explicativas

Hipóteses e Propriedades estatísticas

- **Variância do estimador OLS**

Sob as hipóteses DS.1 a DS.5 então,

Variância do erro

$$Var(\hat{\beta}_j) = \frac{\sigma^2}{SST_j(1 - R_j^2)}, \quad j = 1, \ldots, k$$

Variação total na amostra da v. explicativa *j*

R^2 da regressão da variável explicativa *j* sobre todas as outras variáveis explicativas (incluindo a constante)

$$\sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2$$

Hipóteses e Propriedades estatísticas

- **Desvios padrão e erros padrão de $\hat{\beta}_j$**
  - Desvio padrão de $\hat{\beta}_j$ (desconhecido)

$$sd(\hat{\beta}_j) = \sqrt{Var(\hat{\beta}_j)} = \sqrt{\sigma^2 / \left[SST_j(1 - R_j^2)\right]}$$

  - Erro padrão de $\hat{\beta}_j$

$$se(\hat{\beta}_j) = \sqrt{\widehat{Var}(\hat{\beta}_j)} = \sqrt{\hat{\sigma}^2 / \left[SST_j(1 - R_j^2)\right]}$$

Valor dado pelos softwares.

Hipóteses e Propriedades estatísticas

- **Componentes da variância OLS**

1. Variância do Erro

   - Quando maior a variância do erro, menor a precisão com que se estimam os coeficientes (maior a variância do OLS) porque existe mais "distúrbio" no modelo

   - A variância do erro não diminui aumentando o nº de observações na amostra

2. Variação total do regressor *j*

   - Quanto maior a variação total de x_j maior a precisão com que se estima β_j

   - a variação total de x_j pode aumentar quando aumenta a dimensão da amostra

Hipóteses e Propriedades estatísticas

3. Correlação linear entre os regressores do modelo

- O R_j^2 (coeficiente de determinação da regressão de x_j sobre todas as outras variáveis explicativas) é tanto maior quanto maior for a correlação entre x_j e as outras variáveis do modelo.

- Quanto maior for R_j^2 menor a precisão com que se estima β_j .

Evitar de introduzir no modelo variáveis explicativas muito correlacionadas

- Quando a correlação entre as variáveis é muito elevada tem-se um problema de multicolinearidade, i.e. existe x_j : $R_j \rightarrow 1$

Hipóteses e Propriedades estatísticas

- **O problema da Multicolienaridade**
  - Quando as variáveis estão muito correlacionadas torna-se difícil discriminar o efeito de cada uma delas.
  - O problema da falta de precisão na estimação provocado pela multicolienaridade afeta a estimação dos efeitos parciais apenas das variáveis que têm multicolinearidade e não de todas as variáveis do modelo.
  - A multicolinearidade não representa uma violação da hipótese DS.3
  - Pode ser identificada através dos fatores VIF (*variance inflation factor*)

$$VIF_j = 1/(1 - R_j^2)$$

VIF > 10 $\longrightarrow$ **sintomas de multicolinearidade**

Hipóteses e Propriedades estatísticas

- **Teorema 2: Teorema de Gauss Markov**
  - Sob as hipóteses DS.1 a DS.5 o estimador OLS para os coeficientes do modelo $\beta_j \;\; j = 0,1,\ldots,k$ é o **mais eficiente** na classe dos estimadores centrados e lineares em y.

$$Var(\hat{\beta}_j) \leq Var(\tilde{\beta}_j) \quad j = 0, 1, \ldots, k$$

Para qualquer $\tilde{\beta}_j = \sum_{i=1}^{n} w_{ij} y_i$ que verifique $E(\tilde{\beta}_j) = \beta_j, j = 0, \ldots, k$

- **Teorema 3:**
  - Sob as hipóteses DS.1 a DS.5 $E\left(\hat{\sigma}^2\right) = \sigma^2$

Hipóteses e Propriedades estatísticas

Inclusão de variáveis irrelevantes

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$$

x_3 irrelevante $\Rightarrow \beta_3 = 0$

- como OLS centrado então

$$E\left(\hat{\beta}_3\right) = 0$$

- No entanto a inclusão de variáveis irrelevantes pode levar a que as **variâncias sejam maiores**

Omissão de variáveis relevantes

Verdadeiro modelo

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$

x_2 relevante $\Rightarrow \beta_2 \neq 0$

$$y = \alpha_0 + \alpha_1 x_1 + w$$

Modelo especificado: omissão de X_2

Hipóteses e Propriedades estatísticas

- Enviesamento por omissão de variáveis

$$x_2 = \delta_0 + \delta_1 x_1 + v$$

Se x_2 estiver correlacionado com x_1 pode assumir-se que verificam uma relação linear

$$\Rightarrow \quad y = \beta_0 + \beta_1 x_1 + \beta_2(\delta_0 + \delta_1 x_1 + v) + u$$

Substituindo a relação no verdadeiro modelo obtém-se o modelo especificado

$$= \underbrace{(\beta_0 + \beta_2\delta_0)}_{\alpha_0} + \underbrace{(\beta_1 + \beta_2\delta_1)}_{\alpha_1} x_1 + \underbrace{(\beta_2 v + u)}_{w}$$

- Em conclusão: não se consegue estimar β_0, β_1 e β_2
- **Os coeficientes estimados são enviesados** para estimar os coeficientes do verdadeiro modelo quando a variável omitida está correlacionada com as outras variáveis do modelo

Inferência - Introdução

- **Exemplos**

Peso de um bebé à nascença

$$peso = \beta_0 + \beta_1 cigs + \beta_2 idade + \beta_3 educ + \beta_4 nconsul + u$$

- Será o rendimento uma variável relevante para explicar o peso?

$$H_0: \beta_4 = 0 \qquad H_1: \beta_4 \neq 0$$

Teste de significância individual

- Será que se a mãe tiver mais uma consulta durante a gravidez compensa o efeito negativo no peso decorrente de fumar mais 2 cigarros?

$$H_0: 2\beta_1 + \beta_4 = 0 \qquad H_1: 2\beta_1 + \beta_4 \neq 0$$

Teste de uma restrição linear sobre os parâmetros

- Será que o peso depende apenas do número de cigarros?

$$H_0: \beta_2 = 0,\ \beta_3 = 0,\ \beta_4 = 0 \qquad H_1: \beta_2 \neq 0 \vee \beta_3 \neq 0 \vee \beta_4 \neq 0$$

Teste de significância conjunta
Várias restrições em simultâneo

$$\boxed{peso = \beta_0 + \beta_1 cigs + u} \qquad \boxed{peso = \beta_0 + \beta_1 cigs + \beta_2 educ + \beta_3 idade + \beta_4 nconsul + u}$$

Modelo sob H_0

Modelo sob H_1

Inferência – Distribuição do estimador OLS

- Hipótese DS.6: Distribuição Normal dos erros

  $u_i \sim N(0, \sigma^2)$ Independente de $x_{i1}, x_{i2}, \ldots, x_{ik}$

- Teorema 1: Sob as hipóteses DS.1 a DS.6 prova-se que,

  $\hat{\beta}_j \sim N(\beta_j, Var(\hat{\beta}_j))$ equivalentemente $\dfrac{\hat{\beta}_j - \beta_j}{sd(\hat{\beta}_j)} \sim N(0,1)$

- Teorema 2: Sob as hipóteses DS.1 a DS.6 prova-se que,

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \sim t_{n-k-1}$$

Estatística *t*

Note-se que $sd(\hat{\beta}_j)$ depende de σ^2 que é desconhecido enquanto que $se(\hat{\beta}_j)$ depende de $\hat{\sigma}^2$. A substituição de σ^2 por $\hat{\sigma}^2$ altera a distribuição do rácio da N(0,1) para a distribuição t-student.

Inferência - Intervalos de Confiança

Intervalo de confiança para β_j

$$\beta_j \in \left(\hat{\beta}_j - se(\hat{\beta}_j)t_{\alpha/2}; \hat{\beta}_j + se(\hat{\beta}_j)t_{\alpha/2}\right) \text{ com } (1-\alpha)\times 100\% \text{ de confiança}$$

$$\text{com } t_{\alpha/2} : P\left(t(n-k-1) > t_{\alpha/2}\right) = \frac{\alpha}{2}$$

Esta probabilidade é igual a 0.025 para um nível de confiança de 95%

Exemplo

Dependent Variable: PESO (gramas)
Method: Least Squares
Included observations: 1644 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3170.187	115.8661	27.36079	0.0000
CIGS	-10.49749	3.338919	-3.143977	0.0017
IDADE	3.558060	3.130502	1.136578	0.2559
EDUC	0.975131	7.147705	0.136426	0.8915
NCONSUL	11.46307	3.765319	3.044381	0.0024

R-squared	0.013998	Mean dependent var	3409.99
Adj. R-squared	0.011592	S.D. dependent var	570.58
F-statistic	5.817211	Prob(F-statistic)	0.00012

INTERVALO DE CONFIANÇA A 95% para β_1

$\beta_1 \in$ (-10.50 – 1.96x3.34 ; -10.50 + 1.96x3.34)

$\beta_1 \in$ (-17.05 ; -3.95) com 95% de confiança

Inferência - Testes de hipóteses para um parâmetro

- **Teste de significância individual**

A variável x_j não tem efeito na determinação de y depois de controlar pelo efeito de todas as outras variáveis explicativas (o efeito adicional é nulo)

$$H_0 : \beta_j = 0 \qquad H_1 : \beta_j \neq 0$$

- Estatística de teste sob H_0

$$t_j = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \sim t(n-k-1)$$

- Valor-p (p-value)
  - é o menor nível de significância estatística para o qual ainda se rejeita H_0
  - Mede a evidência favorável a H_0
  - Depende da hipótese alternativa
  - Alternativa bilateral: $H_1 : \beta_j \neq 0$ então $p = 2P\left(t(n-k-1) > |t_{obsj}|\right)$
  - Alternativa $H_1 : \beta_j > 0$ então $p = P\left(t(n-k-1) > t_{obsj}\right)$
  - Alternativa $H_1 : \beta_j < b$ então $p = P\left(t(n-k-1) < t_{obsj}\right)$

Rejeitar H_0 sempre que $p < \alpha$ com $\alpha = 0.01$ ou 0.05 ou 0.1

Inferência - Testes de hipóteses para um parâmetro

- **Teste sobre um valor para β_j**

$$H_0 : \beta_j = b$$

- Estatística de teste sob H_0

$$t = \frac{\hat{\beta}_j - b}{se(\hat{\beta}_j)} \sim t(n-k-1)$$

- Valor-p (p-value) e região de rejeição a $\alpha \times 100\%, W_\alpha$

$H_1 : \beta_j \neq b$ então $p = 2P\left(t(n-k-1) > |t_{obs}|\right)$; $W_\alpha = \left\{t : |t| > t_{\alpha/2}\right\}$; $t_{\alpha/2} : P(t(n-k-1) > t_{\alpha/2}) = \alpha/2$

$H_1 : \beta_j > b$ então $p = P\left(t(n-k-1) > t_{obs}\right)$; $W_\alpha = \left\{t : t > t_\alpha\right\}$; $t_\alpha : P(t(n-k-1) > t_\alpha) = \alpha$

$H_1 : \beta_j < b$ então $p = P\left(t(n-k-1) < t_{obs}\right)$; $W_\alpha = \left\{t : t < -t_\alpha\right\}$; $t_\alpha : P(t(n-k-1) > t_\alpha) = \alpha$

- **Exemplo:** $H_0 : \beta_1 \geq -10 \quad H_1 : \beta_1 < -10$

valor observado da estatística de teste: $t = \dfrac{-10.50 - (-10)}{3.34} = -0.150 > -1.645$

Não rejeitar H_0

Inferência - Testes para uma combinação linear dos parâmetros

- $H_0 : c_1\beta_1 + c_2\beta_2 = b$
- Estatística de teste sob H_0

$$t = \frac{c_1\hat{\beta}_1 + c_2\hat{\beta}_2 - b}{se(c_1\hat{\beta}_1 + c_2\hat{\beta}_2)} \sim t(n-k-1)$$

- Cálculo do erro padrão

$$se\left(c_1\hat{\beta}_1 + c_2\hat{\beta}_2\right) = \sqrt{c_1^2\hat{V}ar(\hat{\beta}_1) + c_2^2\hat{V}ar(\hat{\beta}_2) + 2c_1c_2\hat{C}ov(\hat{\beta}_1, \hat{\beta}_2)}$$

- **Exemplo:** Wald Test:

Test Statistic	Value	df	Probability
t-statistic	-1.231046	1639	0.2185

Null Hypothesis: 2*C(2)+C(5)=0
Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
2*C(2) + C(5)	-9.531905	7.742929

Inferência – Testes F

- **Teste de significância global**
  - Estatística de teste sob H_0

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

Com valor-p $p = P\left(F(k, n-k-1) > F_{obs}\right)$

- **Teste de *q* restrições lineares sobre os parâmetros**
  - Estatística de teste sob H_0

$$F \sim F(q, n-k-1)$$

Com valor-p $p = P\left(F(q, n-k-1) > F_{obs}\right)$

- **Exemplo**

Wald Test:

Test Statistic	Value	df	Probability
F-statistic	0.784908	(2, 1639)	0.4563

Null Hypothesis: C(3)=C(4)=0

Normalized Restriction (= 0)	Value	Std. Err.
C(3)	3.558060	3.130502
C(4)	0.975131	7.147705