ACA228 - Modelos de Regressão e Previsão

Regressão Linear Multipla: Inferência

Prof. Carlos Trucíos
carlos.trucios@facc.ufrj.br
ctruciosm.github.io

Faculdade de Administração e Ciências Contábeis,
Universidade Federal do Rio de Janeiro

Aula 8

Introdução

Introdução

- Até agora conhecemos $\mathbb{E}(\hat{\beta})$ e $\mathbb{V}(\hat{\beta}|X)$

Introdução

- Até agora conhecemos $\mathbb{E}(\hat{\beta})$ e $\mathbb{V}(\hat{\beta}|X)$
- Mas para fazer inferência estatística precisamos mais do que apenas conhecer os dois primeiros momentos, precisamos conhecer a distribuição amostral toda

Introdução

- Até agora conhecemos $\mathbb{E}(\hat{\beta})$ e $\mathbb{V}(\hat{\beta}|X)$
- Mas para fazer inferência estatística precisamos mais do que apenas conhecer os dois primeiros momentos, precisamos conhecer a distribuição amostral toda
- Vamos assumir que u é *normalmente distribuido*

Introdução

- Até agora conhecemos $\mathbb{E}(\hat{\beta})$ e $\mathbb{V}(\hat{\beta}|X)$
- Mas para fazer inferência estatística precisamos mais do que apenas conhecer os dois primeiros momentos, precisamos conhecer a distribuição amostral toda
- Vamos assumir que u é *normalmente distribuido*

Introdução

- Até agora conhecemos $\mathbb{E}(\hat{\beta})$ e $\mathbb{V}(\hat{\beta}|X)$
- Mas para fazer inferência estatística precisamos mais do que apenas conhecer os dois primeiros momentos, precisamos conhecer a distribuição amostral toda
- Vamos assumir que u é *normalmente distribuido*

HRLM6: Normalidade

O erro populacional u é independente das variaveis explicativas (X) e $u \sim N(0, \sigma^2)$

Introdução

- Até agora conhecemos $\mathbb{E}(\hat{\beta})$ e $\mathbb{V}(\hat{\beta}|X)$
- Mas para fazer inferência estatística precisamos mais do que apenas conhecer os dois primeiros momentos, precisamos conhecer a distribuição amostral toda
- Vamos assumir que u é *normalmente distribuido*

HRLM6: Normalidade

O erro populacional u é independente das variaveis explicativas (X) e $u \sim N(0, \sigma^2)$

HRLM1 – HRLM6 são conhecidas como **hipóteses do modelo linear clássico**

Teste t

Teste t

```r
library(wooldridge)
modelo = lm(log(wage) ~ educ+exper+tenure, data = wage1)
coef(modelo)
```

```
## (Intercept)        educ       exper      tenure
## 0.284359555 0.092028987 0.004121109 0.022067217
```

Teste t

Distribuição amostral de $\hat{\beta}_j$

Sob HRLM1–HRLM6 e condicional aos valores amostrais de X,

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\mathbb{V}(\hat{\beta}_j|X)}} \sim N(0,1)$$

Teste t

Distribuição amostral de $\hat{\beta}_j$

Sob HRLM1–HRLM6 e condicional aos valores amostrais de X,

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\mathbb{V}(\hat{\beta}_j|X)}} \sim N(0,1)$$

- **O bom:** $\sim N(0,1)$
- **O problema:** $\mathbb{V}(\hat{\beta}_j|X)$ depende de σ^2 mas nós nunca conhecemos σ^2, então substituimos σ^2 por $\hat{\sigma}^2$, obtendo assim nosso $\widehat{\mathbb{V}}(\hat{\beta}_j|X)$

Teste t

Distribuição amostral de $\hat{\beta}_j$

Sob HRLM1–HRLM6 e condicional aos valores amostrais de X,

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\mathbb{V}(\hat{\beta}_j|X)}} \sim N(0,1)$$

- **O bom:** $\sim N(0,1)$
- **O problema:** $\mathbb{V}(\hat{\beta}_j|X)$ depende de σ^2 mas nós nunca conhecemos σ^2, então substituimos σ^2 por $\hat{\sigma}^2$, obtendo assim nosso $\widehat{\mathbb{V}}(\hat{\beta}_j|X)$

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\widehat{\mathbb{V}}(\hat{\beta}_j|X)}} \sim t_{n-(k+1)}$$

Teste t

No modelo

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + u$$

Geralmente, estamos interessados em testar

$$H_0 : \beta_j = b \quad vs \quad H_1 : \beta_j \neq b$$

$$H_0 : \beta_j \leq b \quad vs \quad H_1 : \beta_j > b$$

$$H_0 : \beta_j \geq b \quad vs \quad H_1 : \beta_j < b$$

Usualmente $b = 0$ (mas outros valores também são utilizados)

Teste t

Para testar hipóteses é preciso uma **estatística de teste**. A estatística utilizada no **Teste t** é chamada de **estatística t**

$$t_{\hat{\beta}_j} = \frac{\hat{\beta}_j - b}{\sqrt{\widehat{\mathbb{V}}(\hat{\beta}_j|X)}} \overset{H_0}{\sim} t_{n-(k+1)}$$

Quando:

- $H_0 : \beta_j = b \quad \text{vs} \quad H_1 : \beta_j \neq b$, **rejeitamos** H_0 se $|t_{\hat{\beta}_j}| > c_0$

Teste t

Para testar hipóteses é preciso uma **estatística de teste**. A estatística utilizada no **Teste t** é chamada de **estatística t**

$$t_{\hat{\beta}_j} = \frac{\hat{\beta}_j - b}{\sqrt{\widehat{\mathbb{V}}(\hat{\beta}_j|X)}} \overset{H_0}{\sim} t_{n-(k+1)}$$

Quando:

- $H_0 : \beta_j = b \quad \text{vs} \quad H_1 : \beta_j \neq b$, **rejeitamos** H_0 se $|t_{\hat{\beta}_j}| > c_0$
- $H_0 : \beta_j \geq b \quad \text{vs} \quad H_1 : \beta_j < b$, **rejeitamos** H_0 se $t_{\hat{\beta}_j} < c_1$

Teste t

Para testar hipóteses é preciso uma **estatística de teste**. A estatística utilizada no **Teste t** é chamada de **estatística t**

$$t_{\hat{\beta}_j} = \frac{\hat{\beta}_j - b}{\sqrt{\widehat{\mathbb{V}}(\hat{\beta}_j|X)}} \overset{H_0}{\sim} t_{n-(k+1)}$$

Quando:

- $H_0 : \beta_j = b$ vs $H_1 : \beta_j \neq b$, **rejeitamos** H_0 se $|t_{\hat{\beta}_j}| > c_0$
- $H_0 : \beta_j \geq b$ vs $H_1 : \beta_j < b$, **rejeitamos** H_0 se $t_{\hat{\beta}_j} < c_1$
- $H_0 : \beta_j \leq b$ vs $H_1 : \beta_j > b$, **rejeitamos** H_0 se $t_{\hat{\beta}_j} > c_2$

Teste t

Para testar hipóteses é preciso uma **estatística de teste**. A estatística utilizada no **Teste t** é chamada de **estatística t**

$$t_{\hat{\beta}_j} = \frac{\hat{\beta}_j - b}{\sqrt{\widehat{\mathbb{V}}(\hat{\beta}_j|X)}} \overset{H_0}{\sim} t_{n-(k+1)}$$

Quando:

- $H_0 : \beta_j = b$ vs $H_1 : \beta_j \neq b$, **rejeitamos** H_0 se $|t_{\hat{\beta}_j}| > c_0$
- $H_0 : \beta_j \geq b$ vs $H_1 : \beta_j < b$, **rejeitamos** H_0 se $t_{\hat{\beta}_j} < c_1$
- $H_0 : \beta_j \leq b$ vs $H_1 : \beta_j > b$, **rejeitamos** H_0 se $t_{\hat{\beta}_j} > c_2$

Teste t

Para testar hipóteses é preciso uma **estatística de teste**. A estatística utilizada no **Teste t** é chamada de **estatística t**

$$t_{\hat{\beta}_j} = \frac{\hat{\beta}_j - b}{\sqrt{\widehat{\mathbb{V}}(\hat{\beta}_j|X)}} \overset{H_0}{\sim} t_{n-(k+1)}$$

Quando:

- $H_0 : \beta_j = b \quad \text{vs} \quad H_1 : \beta_j \neq b$, **rejeitamos** H_0 se $|t_{\hat{\beta}_j}| > c_0$
- $H_0 : \beta_j \geq b \quad \text{vs} \quad H_1 : \beta_j < b$, **rejeitamos** H_0 se $t_{\hat{\beta}_j} < c_1$
- $H_0 : \beta_j \leq b \quad \text{vs} \quad H_1 : \beta_j > b$, **rejeitamos** H_0 se $t_{\hat{\beta}_j} > c_2$

onde c é um quantil da distribuição $t_{n-(k+1)}$ e depende do *nível de significância* α.

Teste t

Para um nível de significância α :

Teste Bilateral:

$$H_0 : \beta_j = b \quad \text{vs} \quad H_1 : \beta_j \neq b,$$

rejeitamos H_0 se $|t_{\hat{\beta}_j}| > c_0 = |t_{\alpha/2, n-(k+1)}| = t_{1-\alpha/2, n-(k+1)}$

Teste t

Para um nível de significância α :

Teste Bilateral:

$$H_0 : \beta_j = b \quad \text{vs} \quad H_1 : \beta_j \neq b,$$

rejeitamos H_0 se $|t_{\hat{\beta}_j}| > c_0 = |t_{\alpha/2, n-(k+1)}| = t_{1-\alpha/2, n-(k+1)}$

Teste Unilateral:

$$H_0 : \beta_j \geq b \quad \text{vs} \quad H_1 : \beta_j < b,$$

rejeitamos H_0 se $t_{\hat{\beta}_j} < c_1 = t_{\alpha, n-(k+1)}$.

Teste t

Para um nível de significância α :

Teste Bilateral:

$$H_0 : \beta_j = b \quad \text{vs} \quad H_1 : \beta_j \neq b,$$

rejeitamos H_0 se $|t_{\hat{\beta}_j}| > c_0 = |t_{\alpha/2, n-(k+1)}| = t_{1-\alpha/2, n-(k+1)}$

Teste Unilateral:

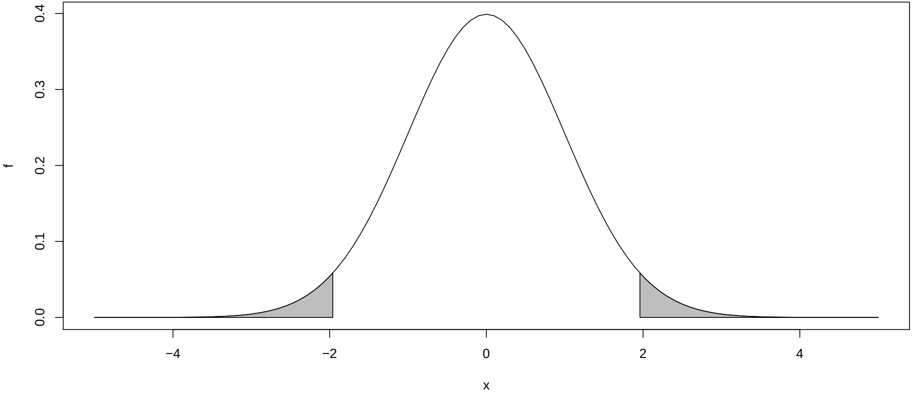
$$H_0 : \beta_j \geq b \quad \text{vs} \quad H_1 : \beta_j < b,$$

rejeitamos H_0 se $t_{\hat{\beta}_j} < c_1 = t_{\alpha, n-(k+1)}$.

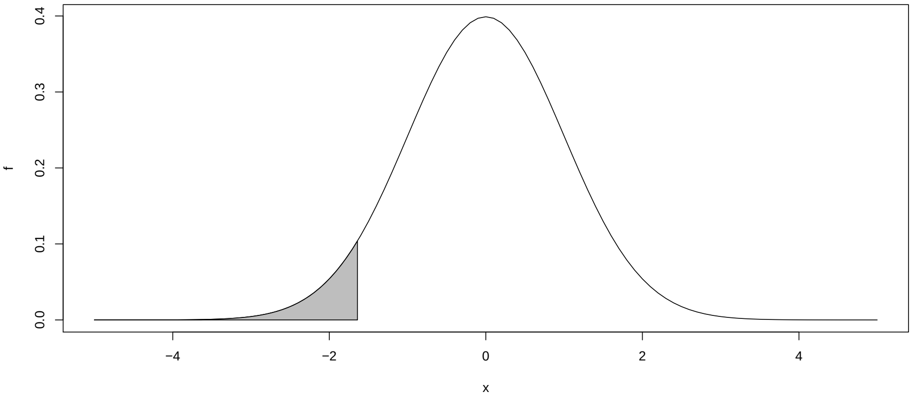
$$H_0 : \beta_j \leq b \quad \text{vs} \quad H_1 : \beta_j > b,$$

rejeitamos H_0 se $t_{\hat{\beta}_j} > c_2 = t_{1-\alpha, n-(k+1)}$.

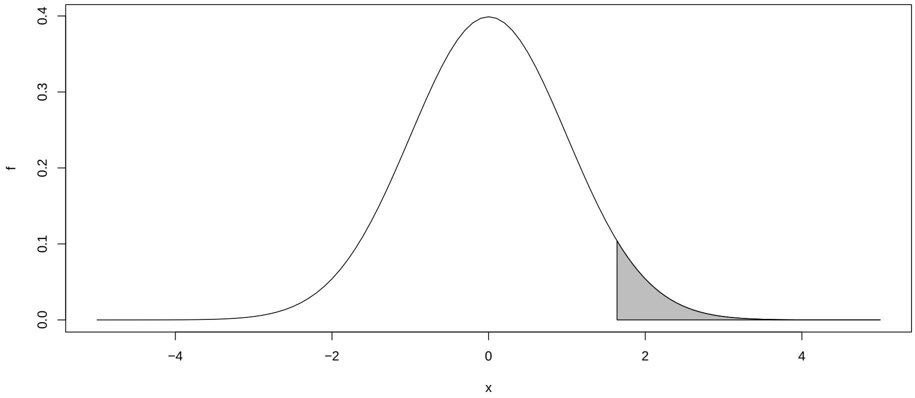
Teste t: Bilateral $H_1 : \beta_j \neq 0$



Teste t: Unilateral $H_1 : \beta_j < 0$



Teste t: Unilateral $H_1 : \beta_j > 0$



Teste t

Resumindo, para testar hipóteses precisamos:

0. Definir o nível de significância α.
1. Calcular a estatística de teste $t_{\hat{\beta}_j}$.
2. Comparar:
   - Para $H_0 : \beta_j = b \quad vs. \quad H_1 : \beta_j \neq b$, rejeitar H_0 se
   $$|t_{\hat{\beta}_j}| > |t_{\alpha/2, n-(k+1)}|$$
   - Para $H_0 : \beta_j \geq b$ vs $H_1 : \beta_j < b$, rejeitar H_0 se
   $$t_{\hat{\beta}_j} < t_{\alpha, n-(k+1)}$$
   - Para $H_0 : \beta_j \leq b$ vs $H_1 : \beta_j > b$, rejeitar H_0 se
   $$t_{\hat{\beta}_j} > t_{1-\alpha, n-(k+1)}.$$
3. Se não temos evidência para rejeitar H_0, então **não rejeitamos** H_0 (nunca falamos que aceitamos H_0, apenas não rejeitamos H_0)

Teste t

- O valor c é obtido do quantil da distribuição $t_{n-(k+1)}$, por exemplo:

```
n = 100; k = 5; alpha = 0.05
# Hipótese Bilateral
qt(1-alpha/2, df = n-(k+1))
```

```
## [1] 1.985523
```

```
# Hipótese Unilateral
qt(alpha, df = n-(k+1)) # ou
```

```
## [1] -1.661226
```

```
qt(1-alpha, df = n-(k+1))
```

```
## [1] 1.661226
```

Teste t

- A maioria de softwares testam a hipótese
  $H_0 : \beta_j = 0 \quad \text{vs} \quad H_1 : \beta_j \neq 0$, e fornecem os *p-valores*.

Teste t

- A maioria de softwares testam a hipótese
  $H_0 : \beta_j = 0 \quad \text{vs} \quad H_1 : \beta_j \neq 0$, e fornecem os *p-valores*.
- Olhando para os *p-valores* podemos rejeitar ou não H_0 sem precisar calcular c

Teste t

- A maioria de softwares testam a hipótese $H_0 : \beta_j = 0 \quad \text{vs} \quad H_1 : \beta_j \neq 0$, e fornecem os *p-valores*.
- Olhando para os *p-valores* podemos rejeitar ou não H_0 sem precisar calcular c
- **Cuidado**, se nosso interesse é testar $H_0 : \beta_j = b$ (com $b \neq 0$) precisaremos fazer as contas **manualmente**

Teste t

- O *p-valor* é a probabilidade (sob H_0) de obter um valor tão extremo quanto aquele observado (Sob H_0, qual a probabilidade de, por acaso, obter um valor tão ou mais extremo quanto $t_{\hat{\beta}_j}$?)

Teste t

- O *p-valor* é a probabilidade (sob H_0) de obter um valor tão extremo quanto aquele observado (Sob H_0, qual a probabilidade de, por acaso, obter um valor tão ou mais extremo quanto $t_{\hat{\beta}_j}$?)
- Valores pequenos do *p-valor* são evidências contra H_0, pois indicam que o resultado ocorre com pequena probabilidade se H_0 for verdadeira.(ou seja a probabilidade de que esse acontecimento seja devido ao acaso é pequena)

Teste t

- O *p-valor* é a probabilidade (sob H_0) de obter um valor tão extremo quanto aquele observado (Sob H_0, qual a probabilidade de, por acaso, obter um valor tão ou mais extremo quanto $t_{\hat{\beta}_j}$?)
- Valores pequenos do *p-valor* são evidências contra H_0, pois indicam que o resultado ocorre com pequena probabilidade se H_0 for verdadeira.(ou seja a probabilidade de que esse acontecimento seja devido ao acaso é pequena)
- O *p-valor* fornece o menor nível de significância no qual H_0 deve ser rejeitada.

Teste t

- O *p-valor* é a probabilidade (sob H_0) de obter um valor tão extremo quanto aquele observado (Sob H_0, qual a probabilidade de, por acaso, obter um valor tão ou mais extremo quanto $t_{\hat{\beta}_j}$?)
- Valores pequenos do *p-valor* são evidências contra H_0, pois indicam que o resultado ocorre com pequena probabilidade se H_0 for verdadeira.(ou seja a probabilidade de que esse acontecimento seja devido ao acaso é pequena)
- O *p-valor* fornece o menor nível de significância no qual H_0 deve ser rejeitada.
- Rejeitamos H_0 se *p-valor* $< \alpha$

Teste t

p-valor

$$p\text{-}valor = P(|T| > |t_0| \big| H_0) = P_{H_0}(|T| > |t_0|)$$

$$p\text{-}valor = P(T < t_0 \big| H_0) = P_{H_0}(T < t_0)$$

$$p\text{-}valor = P(T > t_0 \big| H_0) = P_{H_0}(T > t_0)$$

- Rejetiamos H_0 se $p\text{-}valor < \alpha$, escolhas comuns para α são 0.05, 0.01, 0.10.

Teste t

```r
library(wooldridge)
modelo = lm(log(wage) ~ educ+exper+tenure, data = wage1)
round(summary(modelo)$coefficients,4)
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.2844     0.1042  2.7292   0.0066
## educ          0.0920     0.0073 12.5552   0.0000
## exper         0.0041     0.0017  2.3914   0.0171
## tenure        0.0221     0.0031  7.1331   0.0000
```

Teste t

- E se quisermos testar $H_0 : \beta_{educ} \geq 0 \quad \text{vs} \quad \beta_{educ} < 0$?

```
round(summary(modelo)$coefficients,4)
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.2844     0.1042  2.7292   0.0066
## educ          0.0920     0.0073 12.5552   0.0000
## exper         0.0041     0.0017  2.3914   0.0171
## tenure        0.0221     0.0031  7.1331   0.0000
```

Teste t

- E se quisermos testar $H_0 : \beta_{educ} \geq 0 \quad \text{vs} \quad \beta_{educ} < 0$?

```
round(summary(modelo)$coefficients,4)
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.2844     0.1042  2.7292   0.0066
## educ          0.0920     0.0073 12.5552   0.0000
## exper         0.0041     0.0017  2.3914   0.0171
## tenure        0.0221     0.0031  7.1331   0.0000
```

- **Não podemos utilizar** $P(> |t|)$ (Por quê?)

Teste t

- E se quisermos testar $H_0 : \beta_{educ} \geq 0 \quad \text{vs} \quad \beta_{educ} < 0$?

```
round(summary(modelo)$coefficients,4)
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.2844     0.1042  2.7292   0.0066
## educ          0.0920     0.0073 12.5552   0.0000
## exper         0.0041     0.0017  2.3914   0.0171
## tenure        0.0221     0.0031  7.1331   0.0000
```

- **Não podemos utilizar** $P(> |t|)$ (Por quê?)
- Note que $P_{H_o}(|T| > |t|) = 2P_{H_0}(T > |t|) = 2P_{H_0}(T < -|t|)$

Teste t

- E se quisermos testar $H_0 : \beta_{educ} \geq 0 \quad \text{vs} \quad \beta_{educ} < 0$?

```r
round(summary(modelo)$coefficients,4)
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.2844     0.1042  2.7292   0.0066
## educ          0.0920     0.0073 12.5552   0.0000
## exper         0.0041     0.0017  2.3914   0.0171
## tenure        0.0221     0.0031  7.1331   0.0000
```

- **Não podemos utilizar** $P(> |t|)$ (Por quê?)
- Note que $P_{H_o}(|T| > |t|) = 2P_{H_0}(T > |t|) = 2P_{H_0}(T < -|t|)$
- Então: *p-valor* unilateral $= P_{H_o}(|T| > |t|)/2$

Teste t

- E se quisermos testar $H_0 : \beta_{educ} = 1 \quad \text{vs} \quad \beta_{educ} \neq 1$?

```
round(summary(modelo)$coefficients,4)
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.2844     0.1042  2.7292   0.0066
## educ          0.0920     0.0073 12.5552   0.0000
## exper         0.0041     0.0017  2.3914   0.0171
## tenure        0.0221     0.0031  7.1331   0.0000
```

Teste t

- E se quisermos testar $H_0 : \beta_{educ} = 1 \quad \text{vs} \quad \beta_{educ} \neq 1$?

```
round(summary(modelo)$coefficients,4)
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.2844     0.1042  2.7292   0.0066
## educ          0.0920     0.0073 12.5552   0.0000
## exper         0.0041     0.0017  2.3914   0.0171
## tenure        0.0221     0.0031  7.1331   0.0000
```

- **Não podemos utilizar nem "t value" nem** $P(> |t|)$ (Por quê?)

Teste t

- E se quisermos testar $H_0 : \beta_{educ} = 1 \quad \text{vs} \quad \beta_{educ} \neq 1$?

```
round(summary(modelo)$coefficients,4)
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.2844     0.1042  2.7292   0.0066
## educ          0.0920     0.0073 12.5552   0.0000
## exper         0.0041     0.0017  2.3914   0.0171
## tenure        0.0221     0.0031  7.1331   0.0000
```

- **Não podemos utilizar nem "t value" nem** $P(> |t|)$ (Por quê?)
- $t_{\hat{\beta}_{educ}} = \dfrac{\hat{\beta}_{educ} - 1}{\sqrt{\widehat{V}(\hat{\beta}_{educ})}} = \dfrac{0.0920 - 1}{0.0073} = -124.3836$

Teste t

```r
round(summary(modelo)$coefficients,4)
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.2844     0.1042  2.7292   0.0066
## educ          0.0920     0.0073 12.5552   0.0000
## exper         0.0041     0.0017  2.3914   0.0171
## tenure        0.0221     0.0031  7.1331   0.0000
```

- Para um nível de significância $\alpha = 0.05$,
  $c = t_{1-\alpha/2, n-(k+1)} = 1.964519$

Teste t

```
round(summary(modelo)$coefficients,4)
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.2844     0.1042  2.7292   0.0066
## educ          0.0920     0.0073 12.5552   0.0000
## exper         0.0041     0.0017  2.3914   0.0171
## tenure        0.0221     0.0031  7.1331   0.0000
```

- Para um nível de significância $\alpha = 0.05$, $c = t_{1-\alpha/2,n-(k+1)} = 1.964519$
- Como é um teste de Hipóteses Bilateral rejeitamos H_0 se

$$|t_{\hat{\beta}_{educ}}| > |c|,$$

Teste t

```
round(summary(modelo)$coefficients,4)
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.2844     0.1042  2.7292   0.0066
## educ          0.0920     0.0073 12.5552   0.0000
## exper         0.0041     0.0017  2.3914   0.0171
## tenure        0.0221     0.0031  7.1331   0.0000
```

- Para um nível de significância $\alpha = 0.05$, $c = t_{1-\alpha/2, n-(k+1)} = 1.964519$
- Como é um teste de Hipóteses Bilateral rejeitamos H_0 se

$$|t_{\hat{\beta}_{educ}}| > |c|,$$

- $t_{\hat{\beta}_{educ}} = -124.3836$ e $c = 1.964519$. Como $|-124.3836| > 1.964519$, então rejeitamos H_0 a um nível de significancia de 5%.

Teste t

Estatísticamente significativo: Quando testamos $H_0 : \beta_j = 0 \quad vs. \quad H_1 : \beta_j \neq 0$ e rejeitamos H_0 é comum dizer que β_j é estatísticamente significante ou que a j-ésima variável é estatisticamente significativa.

Teste F

Teste F

- A **estatística t** é utilizada para testar **individualmente** os parâmetros do modelo.

Teste F

- A **estatística t** é utilizada para testar **individualmente** os parâmetros do modelo.
- Seja o modelo

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_8 X_8$$

e suponha que queremos testar

$$H_0 : \beta_2 = 0, \beta_3 = 0, \beta_5 = 0 \quad \text{vs} \quad H_1 : H_0 \text{ não é verdadeiro}$$

Teste F

- A **estatística t** é utilizada para testar **individualmente** os parâmetros do modelo.
- Seja o modelo

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_8 X_8$$

e suponha que queremos testar

$$H_0 : \beta_2 = 0, \beta_3 = 0, \beta_5 = 0 \quad \text{vs} \quad H_1 : H_0 \text{ não é verdadeiro}$$

- Fazer um **teste t** para cada β?

Teste F

- A **estatística t** é utilizada para testar **individualmente** os parâmetros do modelo.
- Seja o modelo

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_8 X_8$$

e suponha que queremos testar

$$H_0 : \beta_2 = 0, \beta_3 = 0, \beta_5 = 0 \quad \text{vs} \quad H_1 : H_0 \text{ não é verdadeiro}$$

- Fazer um **teste t** para cada β?
- Não!. Precisamos fazer o teste de forma conjunta!

Teste F

- A **estatística t** é utilizada para testar **individualmente** os parâmetros do modelo.
- Seja o modelo

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_8 X_8$$

e suponha que queremos testar

$$H_0 : \beta_2 = 0, \beta_3 = 0, \beta_5 = 0 \quad \text{vs} \quad H_1 : H_0 \text{ não é verdadeiro}$$

- Fazer um **teste t** para cada β?
- Não!. Precisamos fazer o teste de forma conjunta!
- O teste F nos permite testar H_0 de forma conjunta!

Teste F

Seja o modelo **irrestrito**

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + u$$

E seja $H_0 : \beta_1 = 0, \beta_2 = 0, \ldots, \beta_q = 0$. Então, o modelo **restrito** (sob H_0) é dado por

$$y = \beta_0 + \beta_{q+1} x_{q+1} + \beta_{q+2} x_{q+2} + \ldots + \beta_k x_k + u$$

Teste F

Sob HRLM1–HRLM6, o **teste F** é dado por

$$F = \frac{(SQR_r - SQR_i)/q}{SQR_i/(n-(k+1))} \overset{H_0}{\sim} F_{q,n-(k+1)}$$

Teste F

$$F = \frac{(SQR_r - SQR_i)/q}{SQR_i/(n-(k+1))} \overset{H_0}{\sim} F_{q,n-(k+1)}$$

- Intuitivamente, estamos analisando quanto aumenta a SQR quando retiramos algumas das variáveis.

Teste F

$$F = \frac{(SQR_r - SQR_i)/q}{SQR_i/(n-(k+1))} \overset{H_0}{\sim} F_{q,n-(k+1)}$$

- Intuitivamente, estamos analisando quanto aumenta a SQR quando retiramos algumas das variáveis.
- Se o aumento no SQR for grande (em relação ao SQR do modelo irrestrito) rejeitamos H_0 (pois a eliminação dessas variáveis aumeta muito a SQR)

Teste F

$$F = \frac{(SQR_r - SQR_i)/q}{SQR_i/(n-(k+1))} \overset{H_0}{\sim} F_{q,n-(k+1)}$$

- Intuitivamente, estamos analisando quanto aumenta a SQR quando retiramos algumas das variáveis.
- Se o aumento no SQR for grande (em relação ao SQR do modelo irrestrito) rejeitamos H_0 (pois a eliminação dessas variáveis aumeta muito a SQR)
- Isto implica que se F for suficientemente grande, rejeitamos H_0

Teste F

$$F = \frac{(SQR_r - SQR_i)/q}{SQR_i/(n-(k+1))} \overset{H_0}{\sim} F_{q,n-(k+1)}$$

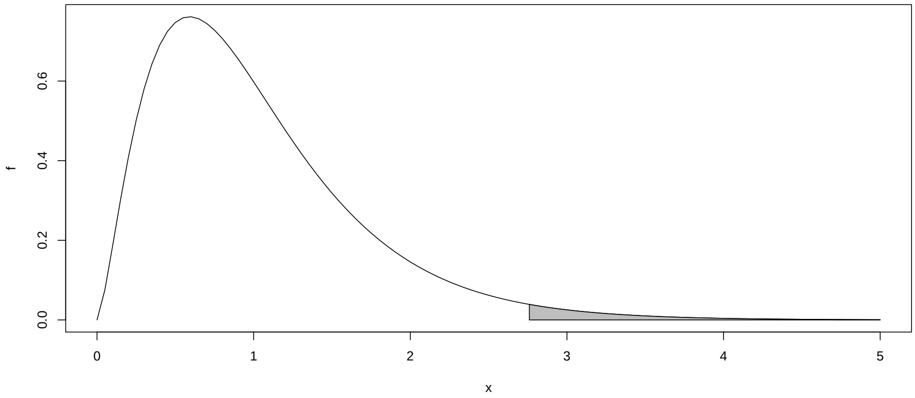
- Intuitivamente, estamos analisando quanto aumenta a SQR quando retiramos algumas das variáveis.
- Se o aumento no SQR for grande (em relação ao SQR do modelo irrestrito) rejeitamos H_0 (pois a eliminação dessas variáveis aumeta muito a SQR)
- Isto implica que se F for suficientemente grande, rejeitamos H_0
- Quão grande? depende do nível de significância α.

Teste F

$$F = \frac{(SQR_r - SQR_i)/q}{SQR_i/(n-(k+1))} \overset{H_0}{\sim} F_{q,n-(k+1)}$$

- Intuitivamente, estamos analisando quanto aumenta a SQR quando retiramos algumas das variáveis.
- Se o aumento no SQR for grande (em relação ao SQR do modelo irrestrito) rejeitamos H_0 (pois a eliminação dessas variáveis aumeta muito a SQR)
- Isto implica que se F for suficientemente grande, rejeitamos H_0
- Quão grande? depende do nível de significância α.
- Como $SQR_r \geq SQR_i$ (sempre) o numerador será sempre ≥ 0, e valores grandes de F nos levarão a rejeitar H_0 (teste unilateral).

Teste F



Teste F

No modelo $\log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + u$

Queremos testar: $H_0 : \beta_1 = 0, \beta_3 = 0 \quad \text{vs} \quad H_1 : H_0$ não é verdadeira

$$F = \frac{(SQR_r - SQR_i)/q}{SQR_i/(n-(k+1))} \overset{H_0}{\sim} F_{q,n-(k+1)}$$

```
modeloi = lm(log(wage) ~ educ + exper + tenure, data = wage1)
modelor = lm(log(wage) ~ exper, data = wage1)
SQRi = sum(residuals(modeloi)^2)
SQRr = sum(residuals(modelor)^2)
n = nrow(wage1); k = 3; q = 2
```

Teste F

```r
F = ((SQRr-SQRi)/SQRi)*((n-k-1)/q)
F
```

```
## [1] 115.8532
```

```r
# Para alpha = 0.05
qf(p = 0.95, df1 = q, df2 = n-k-1)
```

```
## [1] 3.012991
```

Como $\underbrace{F}_{115.8532} > \underbrace{c}_{3.012991}$, então rejeitamos H_0 com um nível de significância $\alpha = 0.05$.

Teste F

```
anova(modelor, modeloi)
```

```
## Analysis of Variance Table
##
## Model 1: log(wage) ~ exper
## Model 2: log(wage) ~ educ + exper + tenure
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1    524 146.49
## 2    522 101.46  2    45.034 115.85 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Teste F (significância geral do modelo)

Dado um modelo da forma

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_k X_k + u \tag{1}$$

um teste bastante rotineiro nos modelos de regressão é:

$$H_0 : \beta_1 = 0, \beta_2 = 0, \ldots, \beta_k = 0 \quad \text{vs} \quad H_0 : H_1 \text{ não é verdadeiro}$$

Neste teste, (1) é o modelo irrestrito e (2) é o modelo restrito.

$$y = \beta_0 + u \tag{2}$$

Teste F (significância geral do modelo)

No modelo

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + u$$

Queremos testar: $H_0 : \beta_1 = 0, \beta_2 = 0, \beta_3 = 0$

```
modeloi = lm(log(wage) ~ educ+exper+tenure, data = wage1)
modelor = lm(log(wage) ~ 1, data = wage1)
SQRi = sum(residuals(modeloi)^2)
SQRr = sum(residuals(modelor)^2)
n = nrow(wage1); k = 3; q = 3
((SQRr-SQRi)/SQRi)*((n-k-1)/q)
```

```
## [1] 80.39092
```

Teste F (significância geral do modelo)

Para saber o valor da estatística F, bem como o *p*-valor associado ao teste, basta fazer

```
summary(modeloi)$fstatistic
```

```
##      value     numdf     dendf
##   80.39092   3.00000 522.00000
```

Se utizarmos o valor da estatística F, então para um nível de significância $\alpha = 0.01$

```
qf(p = 1-0.01, df1 = 3, df2 = 522)
```

```
## [1] 3.819327
```

e rejeitamos H_0 pois $\underbrace{F}_{80.39092} > \underbrace{c}_{3.819327}$.

Teste F

- O **teste F** a diferença do **teste t**, permite que testemos conjuntamente hipóteses com mais de uma restrição

Teste F

- O **teste F** a diferença do **teste t**, permite que testemos conjuntamente hipóteses com mais de uma restrição
- Às vezes os resultados obtidos pelos testes **t** e **F** podem levar a conclusões diferentes, nestes casos é preciso analisar cuidadosamente cada caso.

Teste F

- O **teste F** a diferença do **teste t**, permite que testemos conjuntamente hipóteses com mais de uma restrição
- Às vezes os resultados obtidos pelos testes **t** e **F** podem levar a conclusões diferentes, nestes casos é preciso analisar cuidadosamente cada caso.
- No **teste F** quando $q = 1$, o **teste F** e o **teste t** são equivalentes.

Testes mais gerais

Testes mais gerais

Seja o modelo

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + u$$

e sejam as hipóteses

- $H_0 : \beta_1 = \beta_2 \quad \text{vs} \quad H_1 : \beta_1 \neq \beta_2$

Testes mais gerais

Seja o modelo

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + u$$

e sejam as hipóteses

- $H_0 : \beta_1 = \beta_2 \quad \text{vs} \quad H_1 : \beta_1 \neq \beta_2$
- $H_0 : \beta_1 = 1, \beta_2 = 0, \ldots, \beta_k = 0 \quad \text{vs} \quad H_1 : H_0$ não é verdadeira

Testes mais gerais

Seja o modelo

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + u$$

e sejam as hipóteses

- $H_0 : \beta_1 = \beta_2 \quad \text{vs} \quad H_1 : \beta_1 \neq \beta_2$
- $H_0 : \beta_1 = 1, \beta_2 = 0, \ldots, \beta_k = 0 \quad \text{vs} \quad H_1 : H_0$ não é verdadeira

Testes mais gerais

Seja o modelo

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + u$$

e sejam as hipóteses

- $H_0 : \beta_1 = \beta_2 \quad \text{vs} \quad H_1 : \beta_1 \neq \beta_2$
- $H_0 : \beta_1 = 1, \beta_2 = 0, \ldots, \beta_k = 0 \quad \text{vs} \quad H_1 : H_0$ não é verdadeira

Não podemos usar diretamente as estatísticas de teste nem p-valores reportados por padrão, precisaremos faze-lo *manualmente*.

O primeiro caso será visto aqui, o segundo caso será visto só no laboratório de R.

Testes mais gerais

```r
library(wooldridge)
modelo <- lm(lwage ~ exper + jc + univ, data = twoyear)
coef(modelo)
```

```
## (Intercept)       exper          jc        univ
## 1.472325579 0.004944224 0.066696719 0.076876249
```

Testes mais gerais

```r
library(wooldridge)
modelo <- lm(lwage ~ exper + jc + univ, data = twoyear)
coef(modelo)
```

```
## (Intercept)       exper          jc        univ
## 1.472325579 0.004944224 0.066696719 0.076876249
```

$$H_0 : \beta_{jc} = \beta_{univ} \quad \text{vs} \quad H_1 : \beta_{jc} \neq \beta_{univ}$$

Testes mais gerais

```
library(wooldridge)
modelo <- lm(lwage ~ exper + jc + univ, data = twoyear)
coef(modelo)
```

```
## (Intercept)       exper          jc        univ
## 1.472325579 0.004944224 0.066696719 0.076876249
```

$$H_0 : \beta_{jc} = \beta_{univ} \quad \text{vs} \quad H_1 : \beta_{jc} \neq \beta_{univ}$$

$$H_0 : \beta_{jc} - \beta_{univ} = 0 \quad \text{vs} \quad H_1 : \beta_{jc} - \beta_{univ} \neq 0$$

Testes mais gerais

Construindo uma estatística t:

Testes mais gerais

Construindo uma estatística t:

$$\frac{\hat{\beta}_{jc} - \hat{\beta}_{univ}}{\sqrt{\widehat{\mathbb{V}}(\hat{\beta}_{jc} - \hat{\beta}_{univ})}} = \frac{\hat{\beta}_{jc} - \hat{\beta}_{univ}}{\sqrt{\widehat{\mathbb{V}}(\hat{\beta}_{jc}) + \widehat{\mathbb{V}}(\hat{\beta}_{univ}) - 2\widehat{Cov}(\hat{\beta}_{jc}, \hat{\beta}_{univ})}} \overset{H_0}{\sim} t_{n-(k+1)}$$

```
summary(modelo)$coefficients
```

```
##                Estimate   Std. Error   t value      Pr(>|t|)
## (Intercept) 1.472325579 0.0210602393 69.910202  0.000000e+00
## exper       0.004944224 0.0001574735 31.397175 4.122772e-202
## jc          0.066696719 0.0068287940  9.766984  2.193055e-22
## univ        0.076876249 0.0023087290 33.298083 2.955236e-225
```

Testes mais gerais

```r
coef(modelo)
```

```
## (Intercept)       exper          jc        univ
## 1.472325579 0.004944224 0.066696719 0.076876249
```

```r
betajc = coef(modelo)[3]; betauniv = coef(modelo)[4]
vcov(modelo)
```

```
##               (Intercept)         exper            jc
## (Intercept)  4.435337e-04 -3.104756e-06 -1.741432e-05 -1.57347
## exper       -3.104756e-06  2.479792e-08 -1.718296e-08  3.93349
## jc          -1.741432e-05 -1.718296e-08  4.663243e-05  1.92792
## univ        -1.573472e-05  3.933491e-08  1.927929e-06  5.33023
```

Testes mais gerais

```
COV = vcov(modelo)
s2jc = COV[3,3]; s2univ = COV[4,4]; covjc_univ = COV[3,4]
EstatisticaT = (betajc-betauniv)/sqrt(s2jc+s2univ-2*covjc_univ)
EstatisticaT
```

```
##        jc
## -1.467657
```

Testes mais gerais

```
COV = vcov(modelo)
s2jc = COV[3,3]; s2univ = COV[4,4]; covjc_univ = COV[3,4]
EstatisticaT = (betajc-betauniv)/sqrt(s2jc+s2univ-2*covjc_univ)
EstatisticaT
```

```
##         jc
## -1.467657
```

Para realizar o teste de hipóteses podemos comparar o valor da estatística de teste com o quantil da distribuição sob H_0 ou calcular o p-valor.

Testes mais gerais

```
EstatisticaT
```

```
##        jc
## -1.467657
```

Testes mais gerais

```
EstatisticaT
```

```
##         jc
## -1.467657
```

Para um nível de significância de $\alpha = 0.05$,

```
n = nrow(twoyear); k = 3 #(3 variáveis explicativas)
alpha = 0.05
qt(1-alpha/2, df =  n-k-1)
```

```
## [1] 1.960315
```

Como $\underbrace{|T|}_{1.467657} < \underbrace{c}_{1.960315}$, não rejeitamos H_0 (com um nível de significância de 5%)

Intervalos de Confiança

Intervalos de Confiança

Sabemos que $\dfrac{\hat{\beta}_j - \beta_j}{\sqrt{\widehat{V}(\hat{\beta}_j)}} \sim t_{n-k-1}$

Então, sob HMRLM1 – HMRLM6, calcular IC para os β_j é simples.

IC para β_j

Um intervalo de confiança (1-α)% para β_j, é dado por

$$\Big(\underbrace{\hat{\beta}_j - t_{1-\alpha/2}\sqrt{\widehat{V}(\hat{\beta}_j)}}_{\underline{\beta}_j} \quad ; \quad \underbrace{\hat{\beta}_j + t_{1-\alpha/2}\sqrt{\widehat{V}(\hat{\beta}_j)}}_{\overline{\beta}_j}\Big) \tag{3}$$

onde $t_{1-\alpha/2} = F^{-1}_{t_{n-k-1}}(1-\alpha/2)$

Intervalos de Confiança

- Se as a.a. fossem obtidas repetidas vezes, e em todas elas calcularmos o IC 95%, o valor de β_j estará dentro de $(\underline{\beta}_j; \overline{\beta}_j)$ em (1-α)100% das vezes.

Intervalos de Confiança

- Se as a.a. fossem obtidas repetidas vezes, e em todas elas calcularmos o IC 95%, o valor de β_j estará dentro de ($\underline{\beta}_j$; $\overline{\beta}_j$) em (1-α)100% das vezes.
- Na prática, esperamos ter uma a.a. que seja umas das (1-α)100% em que β_j estará dentro de ($\underline{\beta}_j$; $\overline{\beta}_j$), mas não temos essa certeza.

Intervalos de Confiança

- Se as a.a. fossem obtidas repetidas vezes, e em todas elas calcularmos o IC 95%, o valor de β_j estará dentro de $(\underline{\beta}_j; \overline{\beta}_j)$ em (1-α)100% das vezes.
- Na prática, esperamos ter uma a.a. que seja umas das (1-α)100% em que β_j estará dentro de $(\underline{\beta}_j; \overline{\beta}_j)$, mas não temos essa certeza.
- Lembre-se, quando $n - k - 1$ for grande, podemos aproximar a distribuição t com uma distribuição Normal

Intervalos de Confiança

- Se as a.a. fossem obtidas repetidas vezes, e em todas elas calcularmos o IC 95%, o valor de β_j estará dentro de $(\underline{\beta}_j; \overline{\beta}_j)$ em (1-α)100% das vezes.
- Na prática, esperamos ter uma a.a. que seja umas das (1-α)100% em que β_j estará dentro de $(\underline{\beta}_j; \overline{\beta}_j)$, mas não temos essa certeza.
- Lembre-se, quando $n - k - 1$ for grande, podemos aproximar a distribuição t com uma distribuição Normal
- Para calcular IC, utilizamos a função `confint(modelo, level)`

Intervalos de Confiança

```r
modelo = lm(log(wage) ~ educ + exper + tenure, data = wage1)
confint(modelo, level = 0.95)
```

```
##                     2.5 %     97.5 %
## (Intercept) 0.0796755842 0.48904353
## educ        0.0776292137 0.10642876
## exper       0.0007356983 0.00750652
## tenure      0.0159896850 0.02814475
```

MQO assimptotico

MQO assimptotico

- Até aqui, temos utilizado a HMRLM6 (Normalidade) para construir os testes de hipóteses e os intervalos de confiança.

MQO assimptotico

- Até aqui, temos utilizado a HMRLM6 (Normalidade) para construir os testes de hipóteses e os intervalos de confiança.
- Quando u não é normalmente distribuido, a **estatística t** não tem mais uma distribuição t, e a **estatística F** não tem mais uma distribuição F.

MQO assimptotico

- Até aqui, temos utilizado a HMRLM6 (Normalidade) para construir os testes de hipóteses e os intervalos de confiança.
- Quando u não é normalmente distribuido, a **estatística t** não tem mais uma distribuição t, e a **estatística F** não tem mais uma distribuição F.
- Felizmente, em amostras grandes ($n \to \infty$), mesmo quando HRLM6 não acontece, temos que **estatística t/F** tem aproximadamente distribuição t/F.

MQO assimptotico

- Até aqui, temos utilizado a HMRLM6 (Normalidade) para construir os testes de hipóteses e os intervalos de confiança.
- Quando u não é normalmente distribuido, a **estatística t** não tem mais uma distribuição t, e a **estatística F** não tem mais uma distribuição F.
- Felizmente, em amostras grandes ($n \to \infty$), mesmo quando HRLM6 não acontece, temos que **estatística t/F** tem aproximadamente distribuição t/F.
- Além disso, veremos algumas propriedades interessantes quando $n \to \infty$

MQO assimptotico

Consistencia

Sob HRLM1–HRLM4,

$$\lim_{n\to\infty} \hat{\beta}_j = \beta_j \quad j = 1, \ldots, k \quad \text{(em probabilidade)}$$

Isto significa que se $n \to \infty$, então $\forall \epsilon > 0$,

$$\mathbb{P}(|\hat{\beta}_{\gimel} - \beta_{\gimel}| > \epsilon) = 0$$

- HRLM4 ($\mathbb{E}(u|X) = 0$) pode ser susbtituida por

$$HRLM4' : \quad \mathbb{E}(u) = 0 \quad e \quad \mathbb{C}ov(x_i, u) = 0 \quad \forall i$$

MQO assimptotico

Consistencia

Sob HRLM1–HRLM4,

$$\lim_{n\to\infty} \hat{\beta}_j = \beta_j \quad j = 1, \dots, k \quad \text{(em probabilidade)}$$

Isto significa que se $n \to \infty$, então $\forall \epsilon > 0$,

$$\mathbb{P}(|\hat{\beta}_{\gimel} - \beta_{\gimel}| > \epsilon) = 0$$

- HRLM4 ($\mathbb{E}(u|X) = 0$) pode ser susbtituida por

$$HRLM4' : \quad \mathbb{E}(u) = 0 \quad e \quad \mathbb{C}ov(x_i, u) = 0 \quad \forall i$$

- Se $\mathbb{C}ov(x_i, u) \neq 0$ para algum i, todos os estimadores MQO serão geralmente incosistentes.

MQO assimptotico

Normalidade assintótica

Sob HRLM1–HRML5,

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\widehat{V}(\hat{\beta}_j)}} \overset{a}{\sim} N(0,1)$$

Leituras recomendadas

Leituras recomendadas

- Wooldridge, Jeffrey M. *Introdução à Econometria: Uma abordagem moderna*. (2016). Cengage Learning. – **Cap 4** e **Cap 5**
- Johnston, Jack e Dinardo, John. *Econometric Methods*. (1997), Mc Graw Hill, 4ed. – **Section 3.4.5**