Econometria

Parte 2

Prof. Adalto Acir Althaus Junior oe

Sumário

- Regressão Multivariada
- Erros Heteroscedasticos versus Homoscedasticos
- Testes de hipóteses
- Significância econômica versus estatística

O modelo multivariado

- **Regressão múltipla:** permite a inclusão de diversos fatores para explicar a variável y, o que aproxima o efeito de cada variável da hipótese de ceteris paribus.
- MQO

- Forma geral com k variáveis explicativas:

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

O modelo multivariado

- β_0 continua sendo o intercepto da função e cada β_j é o parâmetro de inclinação associado a variável x_j.
- *u:* termo de erro que reúne os fatores relevantes para explicar *y*, exceto $x_1, x_2, ..., x_k$.
- Repare que, não importa quantos *x* incluímos no modelo, haverá sempre fatores não considerados, ou seja, ***u* continua existindo**.

- Salário/h e educação
- Nota escolar e n^0 de alunos em sala de aula
- Consumo e renda

Derivando os estimadores

- A principal hipótese aqui é uma extensão da média condicional zero do modelo de regressão simples:

$$E(u \mid x_1, x_2, \ldots, x_k) = E(u) = 0$$

- Isto significa que a distribuição de u é independente da distribuição conjunta de $(x_1, x_2, \ldots, x_k)$.
- Neste caso, esta hipótese tem uma implicação ainda mais forte: implica que acertamos a forma funcional que relaciona as *k* variáveis independentes a *y*.

Derivando os estimadores

- Obteremos os estimadores dos *β*'s novamente pelo método dos **mínimos quadrado ordinários**.
- Sabemos que este método é o que nos dá estimativas de *β*'s que minimizam a seguinte função:

$$\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \cdots - \hat{\beta}_k x_{ik})^2$$

- A diferença é que agora devemos escolher $k + 1$ estimadores que minimizem conjuntamente esta função.

Derivando os estimadores

- Neste caso, o problema de minimização gera as ***k* + 1 condições de primeira ordem** a seguir.
  - ✓ Lembrando que:

$$\boxed{\mathrm{E}(u) = 0,\ \mathrm{E}(x_1 u) = 0, \ldots, \mathrm{E}(x_k u) = 0}$$

  - ✓ Então

$$\boxed{\begin{array}{l}
\mathrm{E}(y - \beta_0 - \beta_1 x - \cdots - \beta_k x_k) = 0 \\
\mathrm{E}\left[x_1 (y - \beta_0 - \beta_1 x - \cdots - \beta_k x_k)\right] = 0 \\
\mathrm{E}\left[x_2 (y - \beta_0 - \beta_1 x - \cdots - \beta_k x_k)\right] = 0 \\
\qquad\qquad\vdots \\
\mathrm{E}\left[x_k (y - \beta_0 - \beta_1 x - \cdots - \beta_k x_k)\right] = 0
\end{array}}$$

Estimação dos Parâmetros

Método dos Momentos

– Portanto, os dados amostrais podem ser utilizados para solucionar o problema

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \cdots - \hat{\beta}_k x_{ki}) = 0$$

$$\frac{1}{n}\sum_{i=1}^{n}x_{1i}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \cdots - \hat{\beta}_k x_{ki}) = 0$$

$$\frac{1}{n}\sum_{i=1}^{n}x_{2i}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \cdots - \hat{\beta}_k x_{ki}) = 0$$

$$\frac{1}{n}\sum_{i=1}^{n}x_{3i}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \cdots - \hat{\beta}_k x_{ki}) = 0$$

$$\vdots$$

$$\frac{1}{n}\sum_{i=1}^{n}x_{ki}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \cdots - \hat{\beta}_k x_{ki}) = 0$$

Derivando os estimadores

- Simplificando, o problema de minimização das ***k* + 1 condições de primeira ordem** fica como seguir

$$\sum_{i=1}^{n}(y_i-\hat{\beta}_0-\hat{\beta}_1 x_{i1}-\cdots-\hat{\beta}_k x_{ik})^2=0$$

$$\sum_{i=1}^{n}x_{i1}(y_i-\hat{\beta}_0-\hat{\beta}_1 x_{i1}-\cdots-\hat{\beta}_k x_{ik})^2=0$$

$$\sum_{i=1}^{n}x_{i2}(y_i-\hat{\beta}_0-\hat{\beta}_1 x_{i1}-\cdots-\hat{\beta}_k x_{ik})^2=0$$

.

.

$$\sum_{i=1}^{n}x_{ik}(y_i-\hat{\beta}_0-\hat{\beta}_1 x_{i1}-\cdots-\hat{\beta}_k x_{ik})^2=0$$

Derivando os estimadores

- Repare que estas condições são análogas às condições que usamos para encontrar os estimadores no modelo simples, mas estendendo a cada x_k:

$$\boxed{\mathrm{E}(u) = 0}$$

$$\boxed{\mathrm{E}(u \mid x_1, x_2, \ldots, x_k) = \mathrm{E}(u) = 0}$$

$$\Downarrow$$

A correlação entre o erro e as variáveis explicativas é NULA!

Derivando os estimadores

- O estimador de MQO para o modelo multivariado, passa a ser um **vetor** de dimensão (k + 1) x 1 que resolve unicamente este sistema de equações.

- Explicitar a fórmula destes estimadores exige elementos de álgebra matricial.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki} + \varepsilon_i$$

Indivíduo	Y	X_1	X_2	...	X_k
1	Y_1	X_{11}	X_{21}	...	X_{k1}
2	Y_2	X_{12}	X_{22}	...	X_{k2}
3	Y_3	X_{13}	X_{23}	...	X_{k3}
.	.	.	.	.	.
.	.	.	.	.	.
n	Y_n	X_{1n}	X_{2n}	...	X_{kn}

Derivando os estimadores

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki} + \varepsilon_i$$

onde,

β_0 : o valor médio da variável resposta (Y) quando X_1, X_2, ..., X_k forem nulos;

β_1 : variação no valor médio da variável resposta (Y) no acréscimo de 1 unidade na variável X_1 fixadas as demais;

β_2 : variação no valor médio da variável resposta (Y) no acréscimo de 1 unidade na variável X_2 fixadas a demais;

...

β_k : variação no valor médio da variável resposta (Y) no acréscimo de 1 unidade na variável X_k fixadas a demais;

ε_i : erro aleatório em Y para a i-ésima observação, decorrente da influência de variáveis não consideradas no modelo.

Estimação dos Parâmetros

Mínimos Quadrados Ordinários

- Da mesma forma que na regressão linear simples os estimadores

$$\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_k$$

são chamados de estimadores de mínimos quadrados e podem ser estimados por meio da minimização da soma do quadrado dos resíduos:

$$\boxed{\sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \cdots - \hat{\beta}_k x_{ki})^2}$$

- As condições de primeira ordem são

$$\boxed{\begin{gathered}
\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \cdots - \hat{\beta}_k x_{ki}) = 0 \\
\sum_{i=1}^{n} x_{1i}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \cdots - \hat{\beta}_k x_{ki}) = 0 \\
\vdots \\
\sum_{i=1}^{n} x_{ki}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \cdots - \hat{\beta}_k x_{ki}) = 0
\end{gathered}}$$

Estimação dos Parâmetros

- Voltando ao problema...
- Suponha que os parâmetros do modelo de regressão linear simples são estimados de forma a tornar a soma dos quadrados dos resíduos $\boxed{\sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{1i} - \cdots - \hat{\beta}_k x_{ki})^2}$

  o menor possível.
- $S(b) = S(\hat{\beta}_0, , \ldots, \hat{\beta}_n)$
- A solução é obtida ao derivar$S(b)$ em relação aos b

$$\boxed{\frac{\partial S}{\partial \hat{\beta}_{is}} = -2X'y + 2X'Xb = 0}$$

$$\boxed{b = (X'X)^{-1}X'y}$$

Estimação dos Parâmetros

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \dots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & x_{22} & \dots & x_{k2} \\ 1 & x_{13} & x_{23} & \dots & x_{k3} \\ 1 & x_{14} & x_{24} & \dots & x_{k4} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} & \dots & x_{kn} \end{pmatrix} \times \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ \dots \\ b_k \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ \dots \\ e_n \end{pmatrix}$$

- A solução é

$$\boxed{b = (X'X)^{-1}X'y}$$

Interpretação dos estimadores

- Cada estimador de β_k tem aqui a mesma interpretação que no modelo simples, com apenas uma diferença:

  ✓ mede o efeito de x_k sobre y supondo que fixos os **fatores não-observáveis** (u) e também as **demais variáveis explicativas** que não x_k.

- Assim, nos aproximamos mais do efeito *ceteris paribus* de x_k, pois agora estamos, de fato, fixando parte dos demais fatores associados a y.

Regressão Múltipla

Interpretação Revisitada

- Considere um modelo com apenas duas variáveis explicativas.

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

- Portanto,

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

- O estimador do parâmetro β_1 pode ser escrito como

$$\boxed{\hat{\beta}_1 = \sum_{i=1}^{n} \hat{r}_{1i} y_i \Big/ \sum_{i=1}^{n} \hat{r}_{1i}^2}$$

- onde r_{1i} são os distúrbios (erros) da regressão de x_1 contra x_2.
- Qual a interpretação para o coeficiente β_1?

Regressão Múltipla

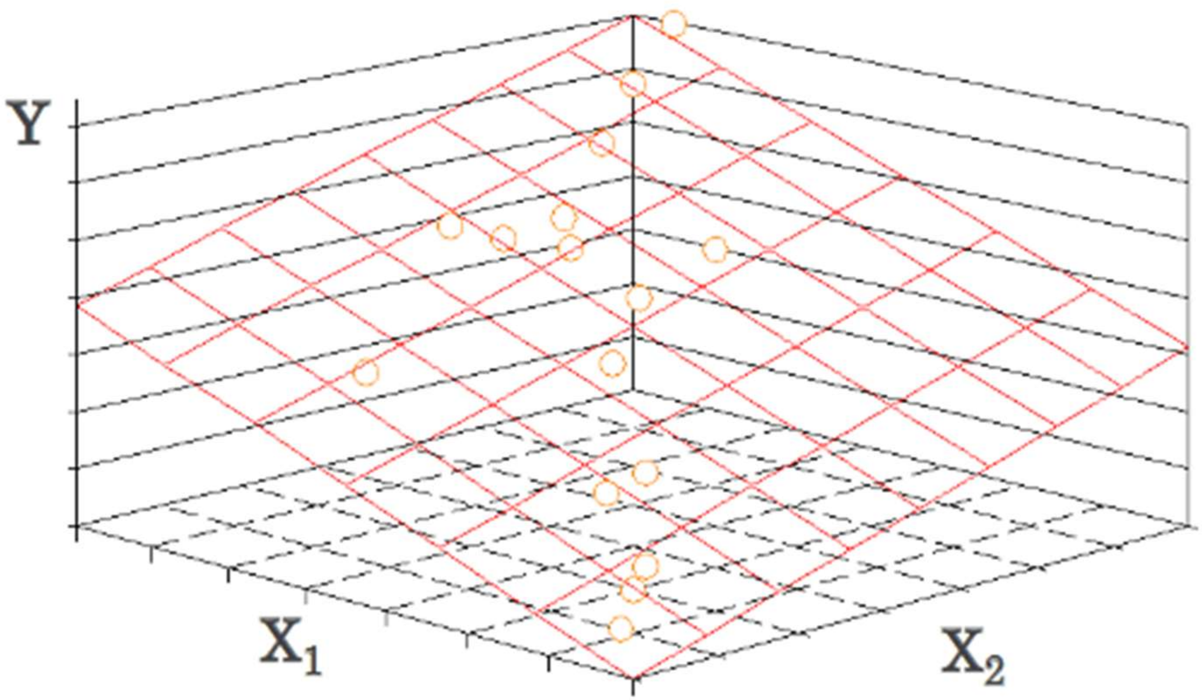
Interpretação Revisitada

- Ou seja, podemos obter a estimativa β_1 por MQO pelo seguinte procedimento:
  - ✓ estimamos uma regressão simples de x_1 contra x_2 e obtemos seu resíduo ($x_1 = \alpha_0 + \alpha_1 x_2 + w$), isto é, a parte de x_1 não associada a x_2 .
  - ✓ então regredimos um modelo simples de y contra este resíduo (r) e obtemos o mesmo β_1 da regressão múltipla de y contra x_1 e x_2 .
- Esta fórmula está nos mostrando, agora matematicamente, que β_1 do modelo múltiplo nos dá a **associação entre *y* e x_1 , descontado (*partialled out*) os efeitos de x_2** .

Interpretação dos estimadores

- Com apenas duas variáveis

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



Interpretação dos estimadores

- Outra forma de interpretar...
- Para o caso de um modelo com 2 variáveis explicativas, temos, em termos de variações:

$$\Delta\hat{y} = \hat{\beta}_1 \Delta x_1 + \hat{\beta}_2 \Delta x_2 + \Delta u$$

- Já mantendo u fixo (Δu=0):

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

- Se mantivermos x_1 fixo (Δx_1=0), temos: $\hat{\beta}_2 = \dfrac{\Delta\hat{y}}{\Delta x_2}$

Interpretação dos estimadores

- Perceba que podemos interpretar β_k como sendo o efeito de x_k mantendo os demais fatores fixos mesmo sem, de fato, ter coletados informações fixas para estes fatores.

- É por esta razão que a regressão múltipla é uma poderosa ferramenta para estabelecer relações de *ceteris paribus* entre fenômenos econômicos mesmo com dados não experimentais.

Mínimos Quadrados Ordinários

Qualidade do Ajuste

- Define-se
  - Soma total dos quadrados (SST – *Total Sum of Squares*)

$$SST \equiv \sum_{i=1}^{n} (y_i - \bar{y})^2$$

  - Soma dos quadrados ajustados/explicados (SSE – *Explained Sum of Squares*)

$$SSE \equiv \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$

  - Soma dos quadrados dos resíduos (SSR – *Residual Sum of Squares*)

$$SSR \equiv \sum_{i=1}^{n} \hat{u}_i^2$$

Mínimos Quadrados Ordinários

Qualidade do Ajuste

- Pela definição de *SST*, *SSE* e *SSR*, chega-se a seguinte relação

$$\boxed{SST = SSE + SSR}$$

- Como na regressão simples pode-se definir o coeficiente de determinação ou R^2

$$\boxed{R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}}$$

$$\boxed{R^2 = \frac{\left(\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})\right)^2}{\left(\sum_{i=1}^{n}(y_i - \bar{y})^2\right)\left(\sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2\right)}}$$

- **ATENÇÃO**: a medida que novas variáveis são incluídas no modelo de regressão linear múltipla o valor do R^2 **nunca decresce**!

Qualidade do ajustamento

- A análise da qualidade de ajustamento usando R^2 em modelo de regressão múltipla é igual a de modelos simples.
- Com notação mais simples, podemos definir R^2 da mesma forma:

$$R^2 \equiv \frac{SQE}{SQT}$$

ou

$$\boxed{R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}}$$

Qualidade do ajustamento

- No caso da regressão múltipla é preciso ressaltar a seguinte propriedade algébrica do R^2 :
  - ✓ ele sempre aumenta quando se inclui uma nova variável explicativa, mesmo que tal variável não tenha poder explicativo sobre y.
- Assim, quando adicionamos uma variável explicativa ao modelo e o R^2 se eleva, devemos tomar cuidado ao interpretar este aumento como uma melhora ajustamento de um modelo devido a inclusão desta nova variável.

Mínimos Quadrados Ordinários

Propriedades Estatísticas dos Estimadores

- Algumas hipóteses importantes:
  - (H1) Modelo populacional é linear

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

  - (H2) Uma amostra aleatória de tamanho *n*

$$\boxed{\{(x_{1i}, x_{2i}, \ldots, x_{ki} y_i) : i = 1, \ldots, n\}}$$

pode ser construída a partir do modelo populacional.

$$\boxed{\mathrm{E}(u \mid x_1, x_2, \ldots, x_k) = \mathrm{E}(u) = 0}$$

Mínimos Quadrados Ordinários

Propriedades Estatísticas dos Estimadores

- Algumas hipóteses importantes continuação:
  - (H3) Média condicional nula

$$\boxed{\mathrm{E}(u \mid x_1, x_2, \ldots, x_k) = \mathrm{E}(u) = 0}$$

  - (H4) Não existe colinearidade perfeita entre as variáveis explicativas
    - Nenhuma variável é constante e não há relação linear entre as variáveis
  - (H5) Homocedasticidade $\boxed{\mathrm{Var}(u \mid x) = \sigma^2}$

Mínimos Quadrados Ordinários

Propriedades Estatísticas dos Estimadores

- Teorema 1: sob as hipóteses (H1) - (H4) os estimadores de mínimos quadrados ordinários são não-tendenciosos, isto é

$$\boxed{\begin{array}{l} E\left(\hat{\beta}_0\right) = \beta_0 \\ E\left(\hat{\beta}_1\right) = \beta_1 \\ E\left(\hat{\beta}_2\right) = \beta_2 \\ \quad\vdots \\ E\left(\hat{\beta}_k\right) = \beta_k \end{array}}$$

Mínimos Quadrados Ordinários

Propriedades Estatísticas dos Estimadores

- O que acontece quando variáveis **irrelevantes** são incluídas no modelo?
  - Considere que o modelo abaixo tenha sido especificado

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

  - Considere ainda que o efeito de x_3 em y, após a inclusão de x_1 e x_2 no modelo, seja nulo. Isto é,

$$\begin{aligned} &\beta_3 = 0 \Rightarrow \mathrm{E}(y \mid x_1, x_2, x_3) = \mathrm{E}(y \mid x_1, x_2) \\ &\mathrm{E}(y \mid x_1, x_2) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \end{aligned}$$

  - Mas na prática não se sabe *a priori* que β_3=0. O que acontecerá com os estimadores?

Mínimos Quadrados Ordinários

Propriedades Estatísticas dos Estimadores

- O que acontece quando variáveis **relevantes** não são incluídas no modelo?
  - Os estimadores serão viesados (tendenciosos).
  - O viés é geralmente chamado de viés de variáveis omitidas.
  - Considere o seguinte modelo populacional

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

  - Agora, suponha que no modelo estimado a variável x_2 não foi incluída

$$\tilde{y} = \tilde{\beta}_0 + \tilde{\beta}_1 x_1$$

$$\Downarrow$$

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n} (x_{1i} - \bar{x}_1) y_i}{\sum_{i=1}^{n} (x_{1i} - \bar{x}_1)^2}$$

Mínimos Quadrados Ordinários

Propriedades Estatísticas dos Estimadores

- Viés de variáveis omitidas (continuação)
  - Pode-se mostrar que

$$\boxed{E(\tilde{\beta}_1) = \beta_1 + \beta_2 \tilde{\delta}_1}$$

onde

$$\boxed{\tilde{\delta}_1 = \frac{\sum_{i=1}^{n} (x_{1i} - \bar{x}_1) x_{2i}}{\sum_{i=1}^{n} (x_{1i} - \bar{x}_1)^2}}$$

Comparação com o modelo simples

- Na grande maioria dos casos, o β_j estimado por regressão simples gera estimativas diferentes daquele estimado por regressão múltipla, após a inclusão de outras variáveis explicativas

- Mas há dois casos em que as estimativa são as iguais.

Comparação com o modelo simples

- Para enxergar estes casos, consideremos novamente um modelo com duas variáveis explicativas:

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

- Considere agora um modelo simples que inclua apenas x_1:

$$\tilde{y} = \tilde{\beta}_0 + \tilde{\beta}_1 x_1$$

- É possível estabelecer a seguinte relação entre as duas estimativas:

$$\tilde{\beta}_1 = \hat{\beta}_1 + \hat{\beta}_2 \tilde{\delta}_1$$

Comparação com o modelo simples

- Em que δ_1 é o coeficiente da regressão entre x_2 e x_1.
- A partir desta relação podemos concluir as duas situações em que as estimativas dos dois modelos são iguais:
  - ✓ Quando não há relação parcial entre x_2 e y, isto é:

$$\hat{\beta}_2 = 0$$

  - ✓ Quando não há relação parcial entre x_2 e x_1, ou seja:

$$\tilde{\delta}_1 = 0$$

- Nestes casos: $\hat{\beta}_1 = \tilde{\beta}_1$

Comparação com o modelo simples

- Estendendo esta análise para um modelo de k variáveis temos as estimativas de β_1 serão iguais entre modelos simples e múltiplo:

  - ✓ Se x_1 for não correlacionado com as demais variáveis explicativas $x_2, \ldots, x_k$.
  - ✓ Se cada variável $x_2, \ldots, x_k$ for não correlacionada com y.

- Nestas situações é possível dizer que não há qualquer ganho em se usar um modelo de regressão múltipla para estimar β_1.

Mínimos Quadrados Ordinários

Variância dos Estimadores

- Teorema 2: sob as hipóteses (H1) - (H5) a variância dos estimadores é dada por

$$\mathrm{Var}(\beta_j) = \frac{\sigma^2}{SST_j(1 - R_j^2)}$$

σ^2 é a variância do erro da regressão, u

onde

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

e

$$R_j^2 = \frac{SSE_j}{SST_j} = \frac{\sum_{i=1}^{n} (\hat{x}_{ji} - \bar{x}_j)^2}{\sum_{i=1}^{n} (x_{ji} - \bar{x}_j)^2}$$

Mínimos Quadrados Ordinários

Variância dos Estimadores

- Três fatores influenciam a variância dos estimadores
  - Variância do erro
  - Variação de x_j
  - Grau de relação linear entre as variáveis explicativas

Mínimos Quadrados Ordinários

Variância do Erro

- Como estimar σ^2?

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

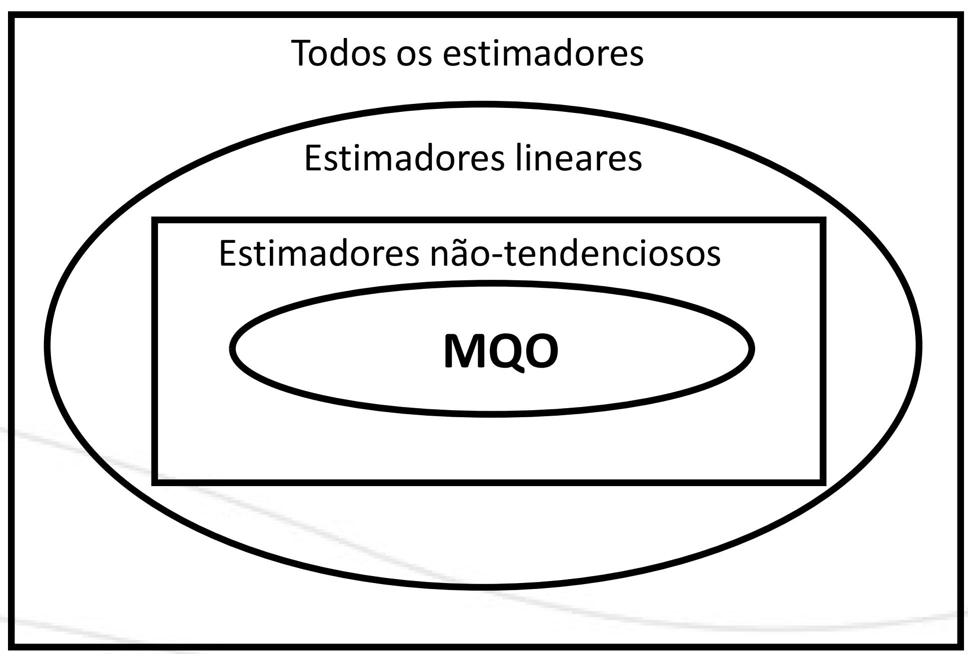
- Teorema 3: sob as hipótese (H1) - (H5)

$$\boxed{E\left(\hat{\sigma}^2\right) = \sigma^2}$$

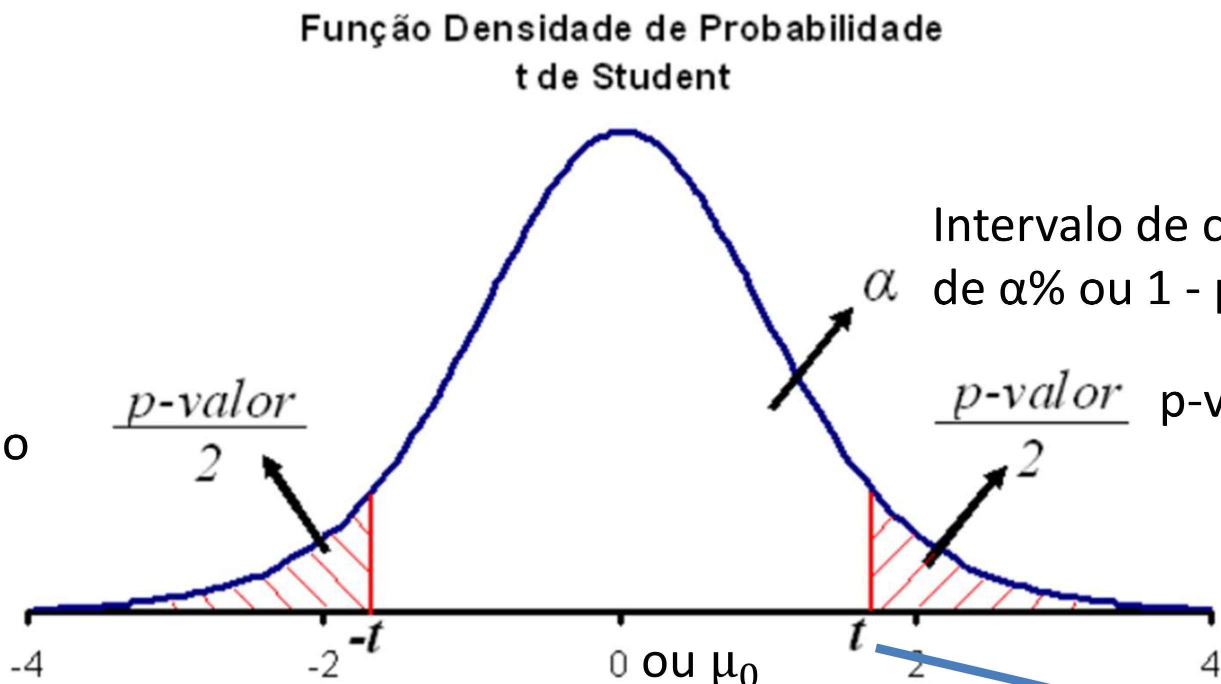
Mínimos Quadrados Ordinários

Teorema de Gauss-Markov

- Teorema 4: sob as hipóteses (H1) - (H5) os estimadores de **MQO** são **BLUE** (*best linear unbiased estimators*), isto é, são os melhores estimadores, no sentido de possuírem menor variância (maior eficiência), dentro da classe dos estimadores lineares e não-viesados**.**



Teste de Hipóteses



Intervalo de confiança de α% ou 1 - p-valor

p-valor de “corte”

Valor “*t*” calculado a ser comparado com valor “*t*” de “corte”

0 ou μ_0 valor a ser comparado

Valor “*t*” crítico para o p-valor de “corte”

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Distribuição *t* de *student* com n-1 graus de liberdade

Desvio padrão da amostra ou σ se for da população

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Teste de Hipóteses

- Antes de chegar ao teste de hipóteses, o que nos permite dizer algo como "nossa estimativa é estatisticamente significativa", é útil primeiro analisar a variância do OLS

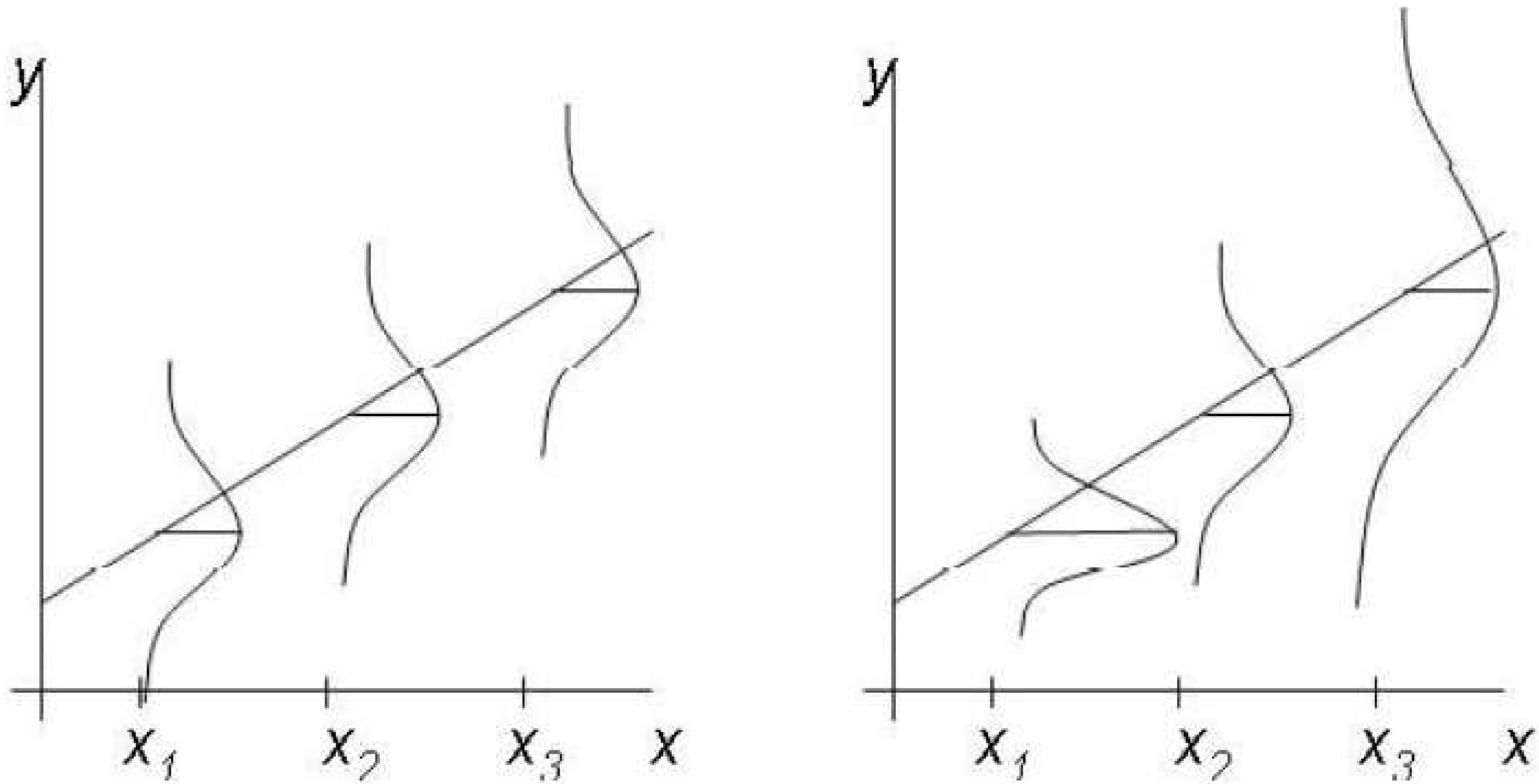
- Entender isso e as suposições feitas para obtê-lo podem nos ajudar a obter os erros padrão corretos para nossos testes de hipóteses posteriores

Variância dos Estimadores OLS

- Homocedasticidade implica $Var\ (u \mid x) = \sigma^2$
  - ✓ A variância dos erros, u, não depende do nível do x observado

- Heteroscedasticidade implica $Var\ (u \mid x) = f\ (x)$
  - ✓ A variância dos erros, u, depende do nível de x de alguma forma

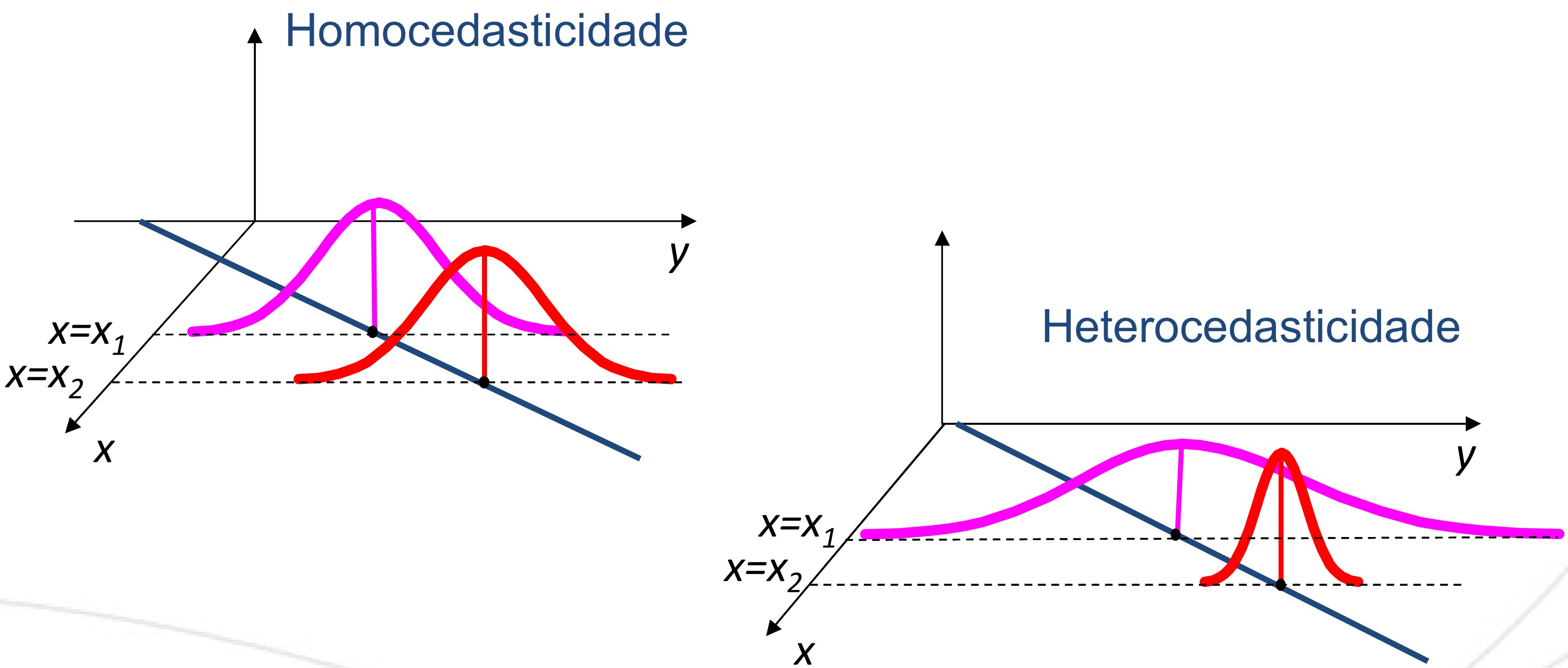
Variância visualmente....



Homoskedasticity **Heteroskedasticity**

Variância visualmente....

$$\boxed{\mathrm{Var}(u \mid x) = \sigma^2}$$



O que é mais realista?

- Numa regressão sobre a variável investimentos, o que é mais realista, Homocedasticidade ou Heteroscedasticidade?

$$Investment = \alpha + \beta Q + u$$

Resposta: A heterocedasticidade parece ser uma suposição muito mais segura de se fazer; Não é difícil chegar a histórias sobre como a homoscedasticidade pode ser violada

Heteroscedastidade e viés

- A existência de heteroscedasticidade causa viés?

$$Investment = \alpha + \beta Q + u$$

Não! $E(u|x)$=0 (que é o que precisamos para estimativas consistentes)
Isso é muito diferente. Heteroscedasticidade afeta somente os Erros Padrões (*Standard Errors*, SEs)

Heteroscedasticidade significa apenas que a estimação OLS pode não ser o mais eficiente (preciso) estimador linear

Heteroscedastidade e viés

- Por que se preocupar com heteroscedasticidade?

- Erros padrão estimados por programas como o Stata assumem o HOMOSCEDASTICIDADE

- Se os erros padrão forem heterocedásticos, as inferências estatísticas feitas a partir desses erros padrão podem estar incorretas…
- **Como corrigimos isso?**

Erros Padrão Robustos

- Use a opção "robust" para obter erros padrão (para testes de hipóteses) que são robustos à heterocedasticidade

- Tipicamente aumenta o SE, mas geralmente não traz maiores complicações na prática
- Se os erros padrão diminuírem, poderá ter problemas; use os erros padrão maiores!

- Erros padrão (ou SE - standard erros) : é uma estimativa do desvio padrão

$$t = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \qquad t_j = \frac{\beta_j}{\frac{u'u}{n-(k+1)} \cdot \sqrt{a_{jj}}}$$

, onde a_{jj} é o j-ésimo termo da diagonal $(X'X)^{-1}$

Usando WLS para lidar com Hek

- Mínimos Quadrados Ponderados (WLS) às vezes são usados quando estamos preocupados com a heterocedasticidade

- WLS basicamente pondera a observação de x usando uma estimativa da variância para esse valor de x

- Feito corretamente, pode melhorar a precisão das estimativas

WLS uma recomendação de uso...

- Recomendação de Angrist-Pischke [Veja a Seção 3.4.1]:
  - ✓ **não se incomode com o WLS**

- OLS é consistente, então por que se incomodar? Pode apenas usar erros padrão robustos
- Propriedades de amostra finitas podem ser ruins [e pode não ser mais eficiente]
- Mais difícil para interpretar do que apenas usando OLS [que ainda é melhor aprox. linear do CEF]
- Se a forma funcional do CEF estiver errada, erros padrão e t-stats estarão errados

Teste de Hipóteses

- Você já deve ter ouvido um tipo de frase comum: “A estimativa, β, é estatisticamente significativa”
- O que isso significa?

Resposta = “Significância estatística” geralmente significa que uma estimativa é estatisticamente diferente de zero

Mas, de onde vem isso?

Teste de Hipóteses

- Quando pensamos em significância, é útil lembrar algumas coisas…
- Estimativas de $\beta 1$, $\beta 2$, etc. são funções de variáveis aleatórias; assim, eles são variáveis aleatórias com variâncias e covariâncias entre si
- Estas variações e covariâncias podem ser estimadas [Ver livros didáticos para várias derivações]
- O **erro padrão** é apenas a raiz quadrada da variância estimada de uma estimativa ($\sqrt{\sigma^2} = \sigma$)

Teste de Hipóteses

- A estatística t (t-stat) informada está nos dizendo quantos desvios padrão nossa estimativa de amostra, $\hat{\beta}$, está longe de zero
- Isto é, está testando a hipótese nula: $\beta = 0$
- p-valor é apenas a probabilidade de obtermos uma estimativa $\hat{\beta}$ que esteja, por sorte, tantos desvios-padrão longe de zero no caso de que seja verdade que $\beta = 0$

Teste de Hipóteses

- Veja livros didáticos para mais detalhes sobre como fazer outros testes de hipóteses; Por exemplo.
- $\beta 1 = \beta 2$
- $\beta 1 = \beta 2 = \beta 3 = 0$
- Dado que estes são geralmente feitos facilmente em programas como o Stata, eu não quero gastar muito tempo revisando a matemática

- TIP: Dê uma olhada nas opções de postestimation do Stata

Teste de Hipóteses

- Considere o modelo

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

- Como testar hipóteses sobre o parâmetro β_j?
  - Por exemplo:
    - H_0: $\beta_j = 0$?
- Sob as hipóteses do modelo

erro padrão

$$\boxed{\frac{\hat{\beta}_j - \beta_j}{\mathrm{se}\left(\hat{\beta}_j\right)} \sim t_{n-k-1}} \qquad t_j = \frac{\beta_j}{\frac{u'u}{n-(k+1)} \cdot \sqrt{a_{jj}}}$$

Teste de Hipóteses

Testes de Hipóteses Sobre um Único Parâmetro

- Um exemplo (saída do E-Views)

Dependent Variable: LOG(WAGE)
Method: Least Squares
Date: 04/15/02 Time: 07:23
Sample: 1 526
Included observations: 526

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.284360	0.104190	2.729230	0.0066
EDUC	0.092029	0.007330	12.55525	0.0000
EXPER	0.004121	0.001723	2.391437	0.0171
TENURE	0.022067	0.003094	7.133070	0.0000

R-squared	0.316013	Mean dependent var	1.623268
Adjusted R-squared	0.312082	S.D. dependent var	0.531538
S.E. of regression	0.440862	Akaike info criterion	1.207406
Sum squared resid	101.4556	Schwarz criterion	1.239842
Log likelihood	-313.5478	F-statistic	80.39092
Durbin-Watson stat	1.768805	Prob(F-statistic)	0.000000

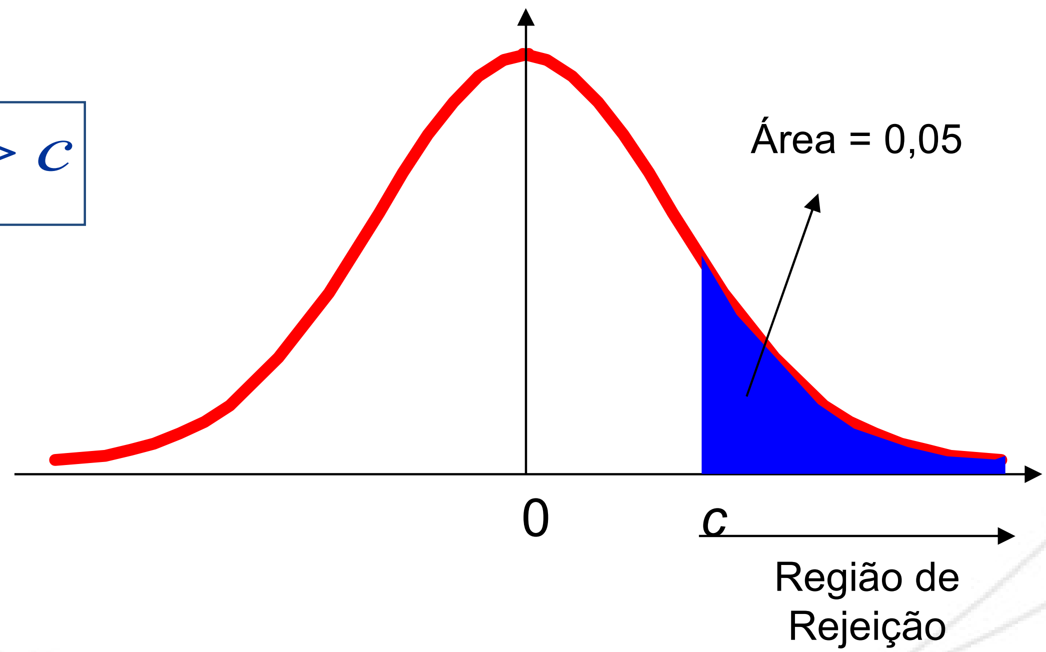
t-Statistic
2.729230
12.55525
2.391437
7.133070

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

Teste de Hipóteses

Testes de Hipóteses Sobre um Único Parâmetro

- Qual é a regra para rejeição de H_0?
  - Depende da hipótese alternativa e do nível de significância do teste!
  - Considere H_1: $\beta_j > 0$ e o nível de significância α = 5% (mais utilizado)
  - Regra de rejeição: $\boxed{t_{\hat{\beta}_j} > c}$



Teste de Hipóteses

Testes de Hipóteses Sobre um Único Parâmetro

- Quando o número de graus de liberdade é grande, a distribuição *t* pode ser aproximada pela distribuição normal padrão.
- Exemplo no E-Views

Dependent Variable: LOG(WAGE)
Method: Least Squares
Date: 04/15/02 Time: 07:23
Sample: 1 526
Included observations: 526

Graus de lib. = 526-4=522

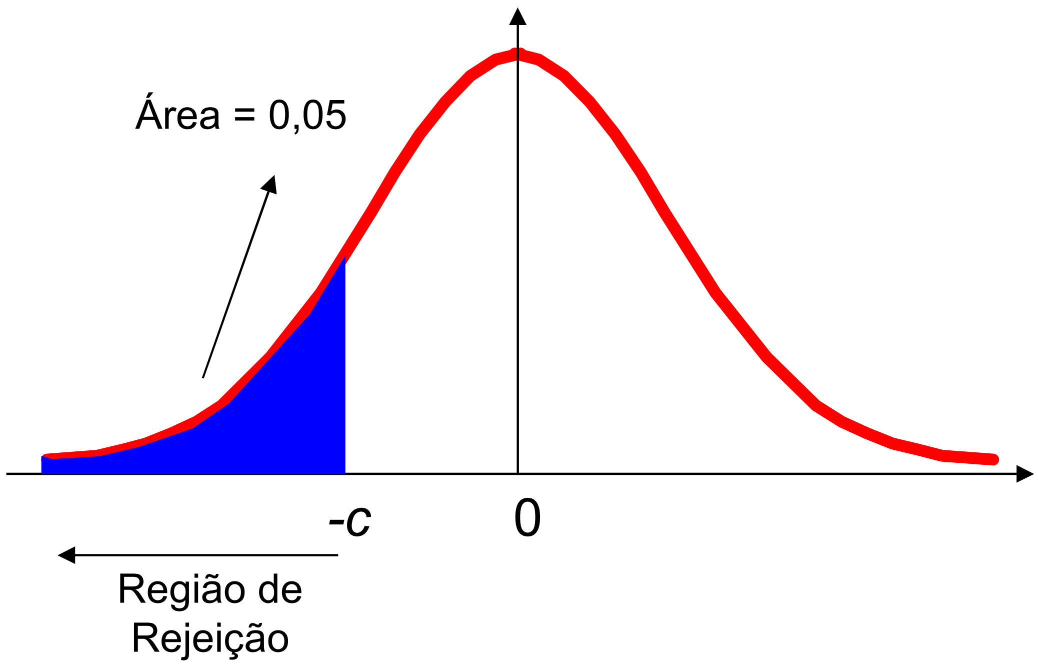
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.284360	0.104190	2.729230	0.0066
EDUC	0.092029	0.007330	12.55525	0.0000
EXPER	0.004121	0.001723	2.391437	0.0171
TENURE	0.022067	0.003094	7.133070	0.0000

R-squared	0.316013	Mean dependent var	1.623268
Adjusted R-squared	0.312082	S.D. dependent var	0.531538
S.E. of regression	0.440862	Akaike info criterion	1.207406
Sum squared resid	101.4556	Schwarz criterion	1.239842
Log likelihood	-313.5478	F-statistic	80.39092
Durbin-Watson stat	1.768805	Prob(F-statistic)	0.000000

Teste de Hipóteses

Testes de Hipóteses Sobre um Único Parâmetro

- Considere H_1: $\beta_j < 0$ e o nível de significância α = 5%

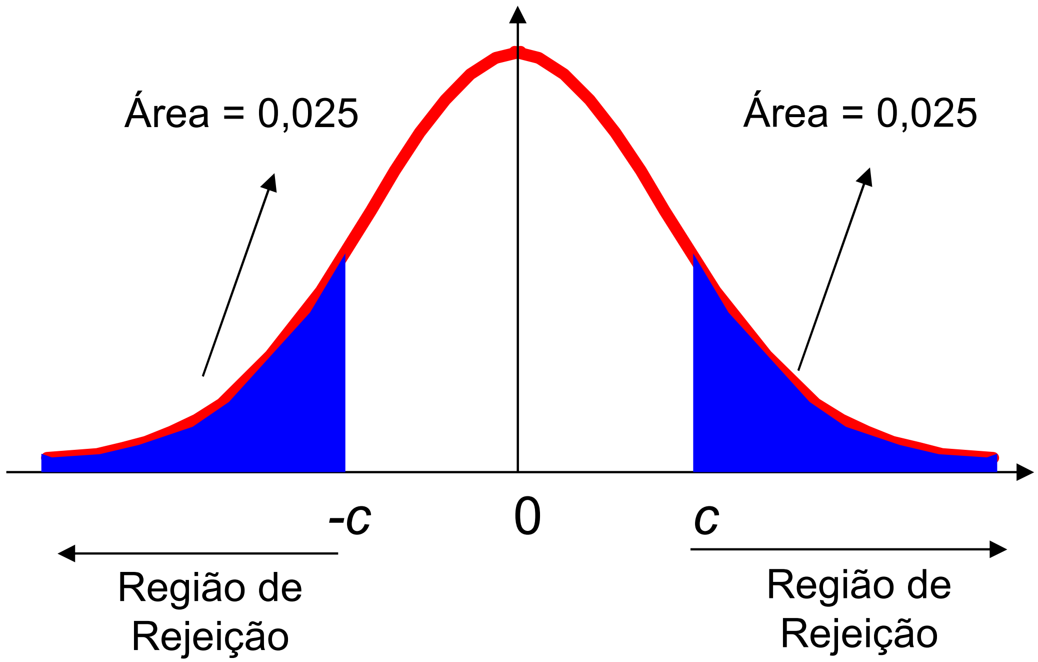


- Regra de rejeição: $\boxed{t_{\hat{\beta}_j} < -c}$

Teste de Hipóteses

Testes de Hipóteses Sobre um Único Parâmetro

- Considere H_1: $\beta_j \neq 0$ e o nível de significância α = 5%



- Regra de rejeição: $\boxed{\left| t_{\hat{\beta}_j} \right| > c}$

Teste de Hipóteses

Testes de Hipóteses Sobre um Único Parâmetro

- Dado o valor da estatística *t*, qual o menor nível de significância para o qual a hipótese nula é rejeitada?
  - Este nível de significância é conhecido como o *p*-valor do teste e é definido por

$$\boxed{\Pr\left(|T| > |t|\right)}$$

onde *T* é uma variável aleatória com distribuição *t* e *n*-*k*-1 graus de liberdade e *t* é o valor numérico observado da estatística de teste.

Teste de Hipóteses

- Exemplo no E-Views

Dependent Variable: LOG(WAGE)
Method: Least Squares
Date: 04/15/02 Time: 07:23
Sample: 1 526
Included observations: 526

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.284360	0.104190	2.729230	0.0066
EDUC	0.092029	0.007330	12.55525	0.0000
EXPER	0.004121	0.001723	2.391437	0.0171
TENURE	0.022067	0.003094	7.133070	0.0000

R-squared	0.316013	Mean dependent var	1.623268
Adjusted R-squared	0.312082	S.D. dependent var	0.531538
S.E. of regression	0.440862	Akaike info criterion	1.207406
Sum squared resid	101.4556	Schwarz criterion	1.239842
Log likelihood	-313.5478	F-statistic	80.39092
Durbin-Watson stat	1.768805	Prob(F-statistic)	0.000000

Prob.
0.0066
0.0000
0.0171
0.0000

$$\Pr\left(|T| > |t|\right)$$

Teste de Hipóteses

- Vimos que sob as hipóteses clássicas do modelo de regressão linear

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

- Desta forma podem ser criados intervalos de confiança para os parâmetros estimados, definidos por $\boxed{\hat{\beta}_j \pm c \cdot \mathrm{se}(\hat{\beta}_j)}$

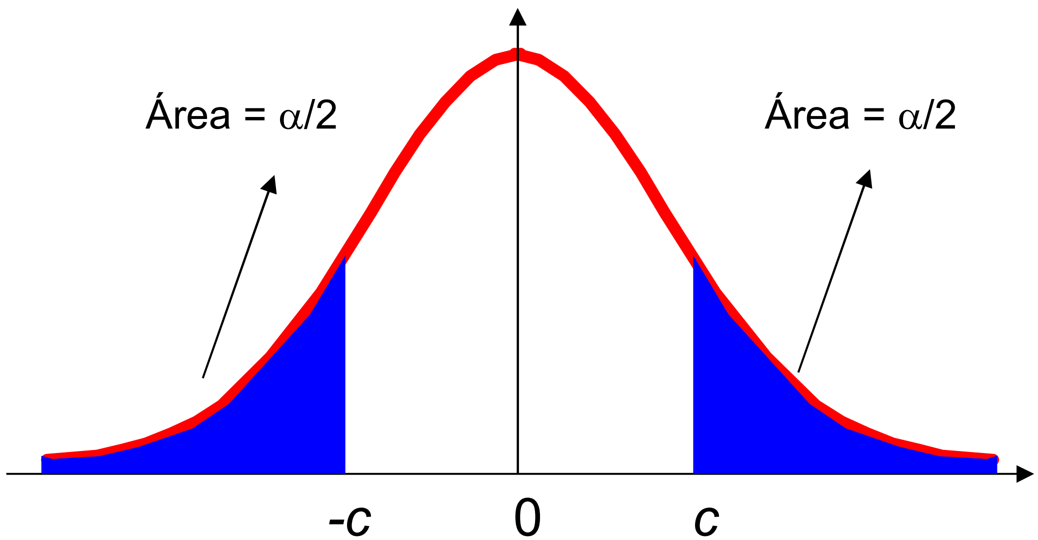
  onde *c* é o percentil adequado da distribuição t com *n* -*k* -1 graus de liberdade.

  – Por exemplo: se o intervalo de confiança for de 95% e o número de graus de liberdade for igual à 25, então *c* vale 2,06.

Teste de Hipóteses

Intervalos de Confiança

- Se o nível de confiança do teste for 1-α, então *c* é o percentil da distribuição t referente à $\alpha/2$.



Mínimos Quadrados Ordinários

- Um exemplo no E-Views

Dependent Variable: LOG(WAGE)
Method: Least Squares
Date: 04/15/02 Time: 07:23
Sample: 1 526
Included observations: 526

Graus de lib. = 526-4=522
A distribuição t é aproximadamente igual à distribuição normal padronizada

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.284360	0.104190	2.729230	0.0066
EDUC	0.092029	0.007330	12.55525	0.0000
EXPER	0.004121	0.001723	2.391437	0.0171
TENURE	0.022067	0.003094	7.133070	0.0000

R-squared	0.316013	Mean dependent var	1.623268
Adjusted R-squared	0.312082	S.D. dependent var	0.531538
S.E. of regression	0.440862	Akaike info criterion	1.207406
Sum squared resid	101.4556	Schwarz criterion	1.239842
Log likelihood	-313.5478	F-statistic	80.39092
Durbin-Watson stat	1.768805	Prob(F-statistic)	0.000000

Intervalos de 95%

$$\left[\hat{\beta}_0\right] = 0{,}284360 \pm 1{,}96 \cdot 0{,}104190$$
$$\left[\hat{\beta}_1\right] = 0{,}092029 \pm 1{,}96 \cdot 0{,}007330$$
$$\left[\hat{\beta}_2\right] = 0{,}004121 \pm 1{,}96 \cdot 0{,}001723$$
$$\left[\hat{\beta}_3\right] = 0{,}022067 \pm 1{,}96 \cdot 0{,}003094$$

Significância Estatística vs Econômica

- Isto não é a mesma coisa!
- O coeficiente pode ser estatisticamente significativo, mas economicamente pequeno
- Você pode obter isso em grandes amostras, ou quando você tem muita variação em x (ou outliers)
- O coeficiente pode ser economicamente grande, mas estatisticamente insignificante
- Pode ser apenas um pequeno tamanho de amostra ou pouca variação em x para obter uma estimativa precisa

Significância Estatística vs Econômica

- Você deve sempre verificar a significância econômica dos coeficientes

- Por exemplo: Quão grande é a mudança implícita em y para uma mudança do tamanho de um desvio padrão em x?

- **E importante**, isso é plausível? Se não, você pode ter um problema de especificação

Outras Questões