ANÁLISE DE REGRESSÃO LINEAR SIMPLES E MÚLTIPLA

BIBLIOGRAFIA RECOMENDADA

- FÁVERO. L. P. **Análise de dados: modelos de regressão com Excel®, STATA® e SPSS®**. Rio de Janeiro: Elsevier, 2015.
- GUJARATI, D. **Econometria Básica,** Ed. Campus, São Paulo, 2006.

MODELOS E ANÁLISE DE DADOS

- Uma das preocupações de um estatístico ao analisar um conjunto de dados é criar modelos que explicitem estruturas do fenômeno sob observação, as quais frequentemente estão misturadas com variações aleatórias.

- Portanto, uma estratégia conveniente de análise é supor que cada observação seja formada por duas partes:

observação = previsível + aleatório

observação = previsível + aleatório

- Conhecimento que o pesquisador tem sobre o fenômeno
- Usualmente expressa por uma função matemática, com **parâmetros desconhecidos**.

- Componente aleatório (ou não previsível)
- Representa aquilo que o pesquisador não pode controlar
- São impostas algumas suposições a esse componente
    - Obedeça a algum modelo probabilístico específico, caracterizado por **parâmetros desconhecidos**.

MODELOS E ANÁLISE DE DADOS

- Dentro desse cenário, o trabalho do estatístico passa a ser o de estimar os parâmetros desconhecidos das duas partes do modelo, baseado em amostras observadas.

EXEMPLO: A RETA CARACTERÍSTICA DE UMA AÇÃO

$$r_{it} = \alpha_i + \beta_i r_{mt} + u_{it}$$

- r_{it}: retorno do ativo i no momento t
- r_{mt}: retorno da carteira de mercado no momento t
- α_i: intercepto, representa o retorno esperado em excesso de um ativo. Tanto no caso de ações quanto de portfolios, espera-se $\alpha = 0$.
- β_i: coeficiente angular
  - Representa o risco do ativo em relação ao mercado

LINGUAGEM

DADOS DE SÉRIE TEMPORAL

- “Conjunto de observações dos valores que uma variável assume em diferentes momentos.”

DADOS

Dados de Série temporal

Taxas de Inflação em cinco países industrializados, 1960-1970 (% anual)					
Ano	**EUA**	**Reino Unido**	**Japão**	**Alemanha**	**França**
1960	1,50	1,00	3,60	1,50	3,60
1961	1,10	3,40	5,40	2,30	3,40
1962	1,10	4,50	6,70	4,50	4,70
1963	1,20	2,50	7,70	3,00	4,80
1964	1,40	3,90	3,90	2,30	3,40
1965	1,60	4,60	6,50	3,40	2,60
1966	2,80	3,70	6,00	3,50	2,70
1967	2,80	2,40	4,00	1,50	2,70
1968	4,20	4,80	5,50	1,80	4,50
1969	5,00	5,20	5,10	2,60	6,40
1970	5,90	6,50	7,60	3,70	5,50

Fonte: GUJARATI (2005:11-12)

LINGUAGEM

DADOS DE SÉRIE TEMPORAL

- “Conjunto de observações dos valores que uma variável assume em diferentes momentos.”

DADOS DE CORTE (Cross-Section)

- “São dados de uma ou mais variáveis coletados no mesmo ponto do tempo.”

DADOS

Dados "Cross Section"

Taxas de Inflação em cinco países industrializados, 1960-1970 (% anual)

Ano	EUA	Reino Unido	Japão	Alemanha	França
1960	1,50	1,00	3,60	1,50	3,60
1961	1,10	3,40	5,40	2,30	3,40
1962	1,10	4,50	6,70	4,50	4,70
1963	1,20	2,50	7,70	3,00	4,80
1964	1,40	3,90	3,90	2,30	3,40
1965	1,60	4,60	6,50	3,40	2,60
1966	2,80	3,70	6,00	3,50	2,70
1967	2,80	2,40	4,00	1,50	2,70
1968	4,20	4,80	5,50	1,80	4,50
1969	5,00	5,20	5,10	2,60	6,40
1970	5,90	6,50	7,60	3,70	5,50

Fonte: GUJARATI (2005:11-12)

LINGUAGEM

DADOS DE SÉRIE TEMPORAL

- "Conjunto de observações dos valores que uma variável assume em diferentes momentos."

DADOS DE CORTE (Cross-Section)

- "São dados de uma ou mais variáveis coletados no mesmo ponto do tempo."

DADOS COMBINADOS (Painel)

- "Há elementos tanto de séries temporais como de dados de corte."

DADOS

Painel de dados

Taxas de Inflação em cinco países industrializados, 1960-1970 (% anual)

Ano	EUA	Reino Unido	Japão	Alemanha	França
1960	1,50	1,00	3,60	1,50	3,60
1961	1,10	3,40	5,40	2,30	3,40
1962	1,10	4,50	6,70	4,50	4,70
1963	1,20	2,50	7,70	3,00	4,80
1964	1,40	3,90	3,90	2,30	3,40
1965	1,60	4,60	6,50	3,40	2,60
1966	2,80	3,70	6,00	3,50	2,70
1967	2,80	2,40	4,00	1,50	2,70
1968	4,20	4,80	5,50	1,80	4,50
1969	5,00	5,20	5,10	2,60	6,40
1970	5,90	6,50	7,60	3,70	5,50

Fonte: GUJARATI (2005:11-12)

Modelos de Regressão Simples e Múltipla

- Técnicas de análise mais utilizadas em diversos campos do conhecimento
  - Retorno de um ativo em relação a variáveis de mercado
  - Custo em função da quantidade produzida e das horas trabalhadas
- Fenômeno de interesse: variável métrica

Modelos de Regressão Simples e Múltipla

- **Objetivo**:
  - Analisar como se comportam as relações entre variável **dependente métrica** e um conjunto de variáveis **explicativas**, **métricas** ou **categóricas binárias**
- As relações entre as variáveis são definidas com base na teoria e na experiência do pesquisador

MODELOS LINEARES DE REGRESSÃO

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

- Y: fenômeno em estudo (variável dependente quantitativa métrica)
- β_0: intercepto (constante, ou coeficiente linear)
- β_j: coeficientes angulares
- X_j: variáveis explicativas (métricas ou *dummies*)
- u: termo de erro (diferença entre o valor real de Y e o valor previsto de Y por meio do modelo, para cada observação)
- $i = 1, 2, 3, \ldots, n$ representa cada uma das observações da amostra

Modelo de **regressão múltipla**

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

Modelo de **regressão simples**

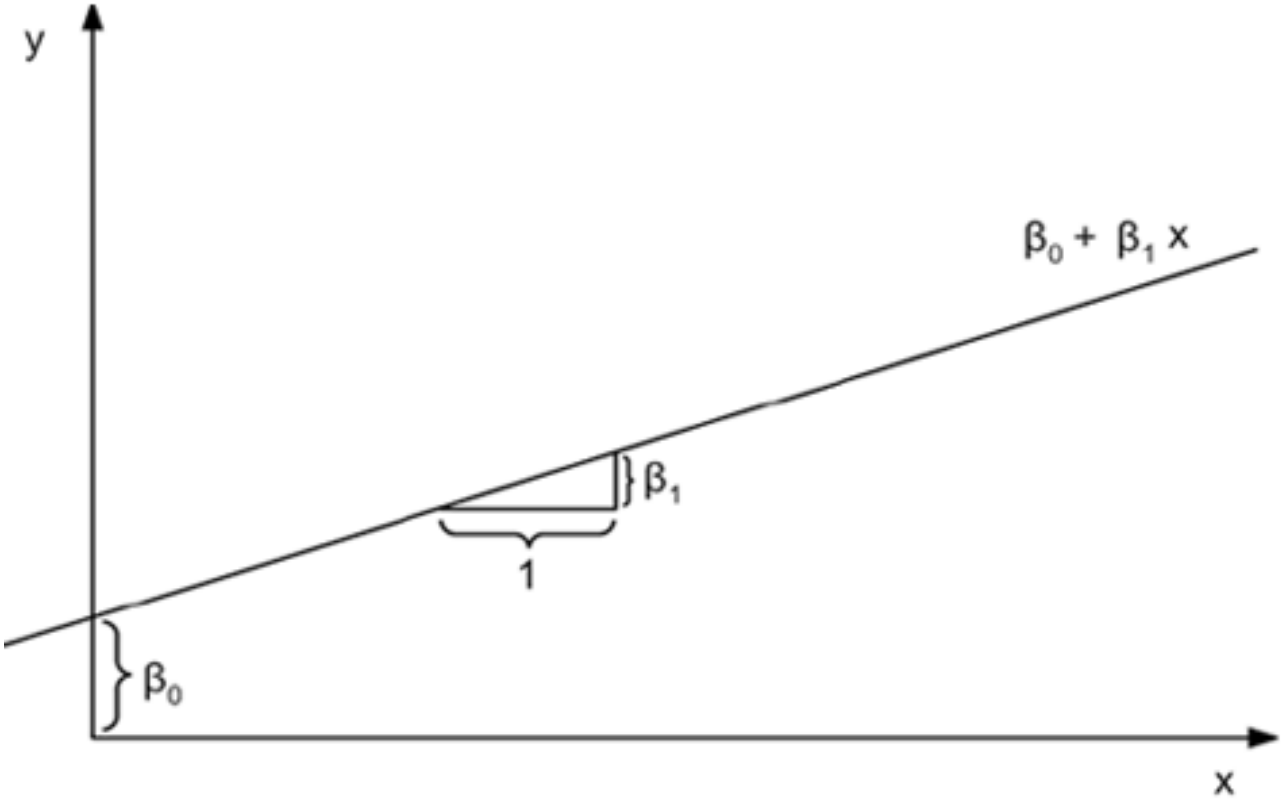
$$Y_i = \beta_0 + \beta_1 . X_i + u_i$$

Termo de erro u, chamado de **resíduo**:

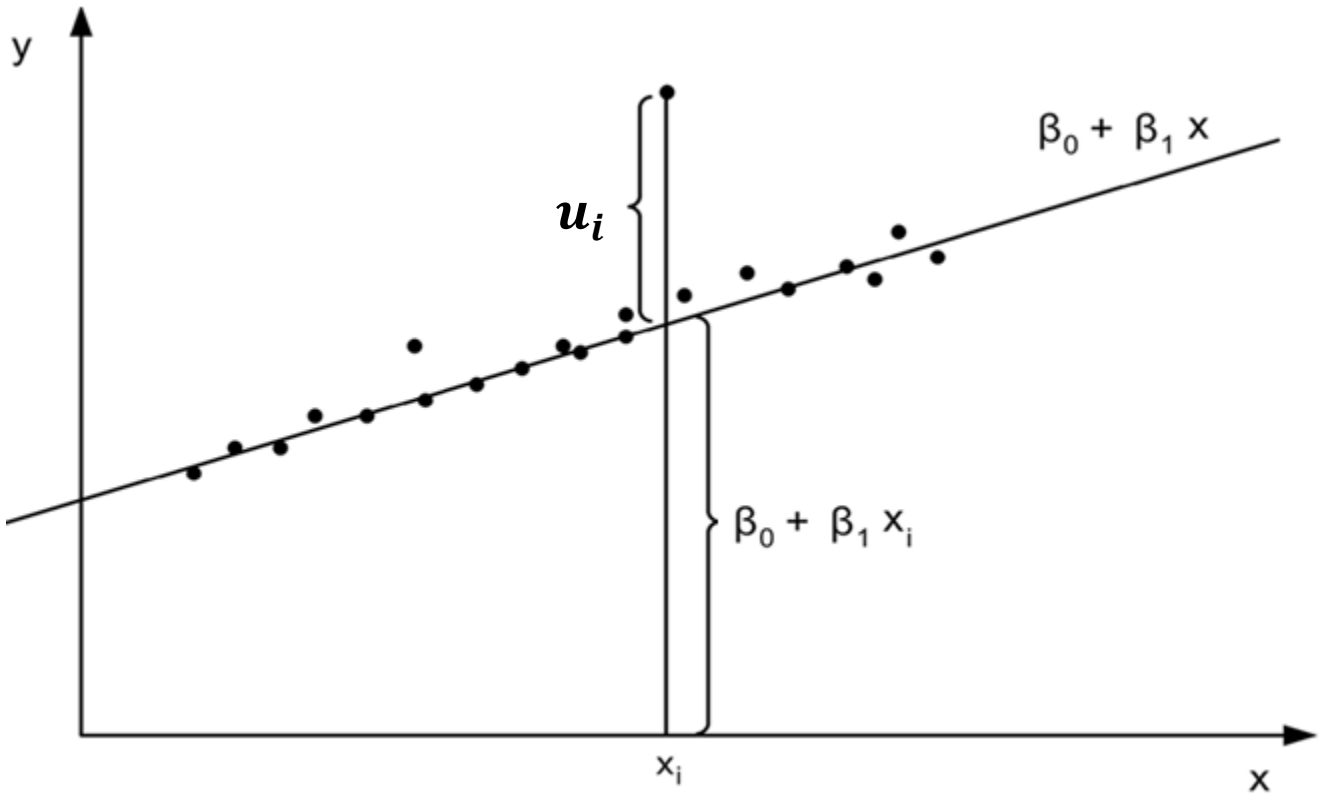
- Nenhuma relação entre Y e X se apresentará de maneira perfeita:
  - Falhas na especificação do modelo (variáveis omitidas e forma funcional errada)

$$u_i = Y_i - \beta_0 - \beta_1 . X_i$$

ESTIMAÇÃO DO MODELO DE REGRESSÃO LINEAR



ESTIMAÇÃO DO MODELO DE REGRESSÃO LINEAR



ESTIMAÇÃO DO MODELO DE REGRESSÃO LINEAR

Objetivo principal:

- Avaliar como se comporta uma variável Y em função de uma ou mais variáveis X, sem que necessariamente ocorra uma relação de causa e efeito

ESTIMAÇÃO DO MODELO DE REGRESSÃO LINEAR

Função de regressão **populacional**

$$Y_i = \beta_0 + \beta_1 . X_i + u_i$$

Função de regressão **amostral**

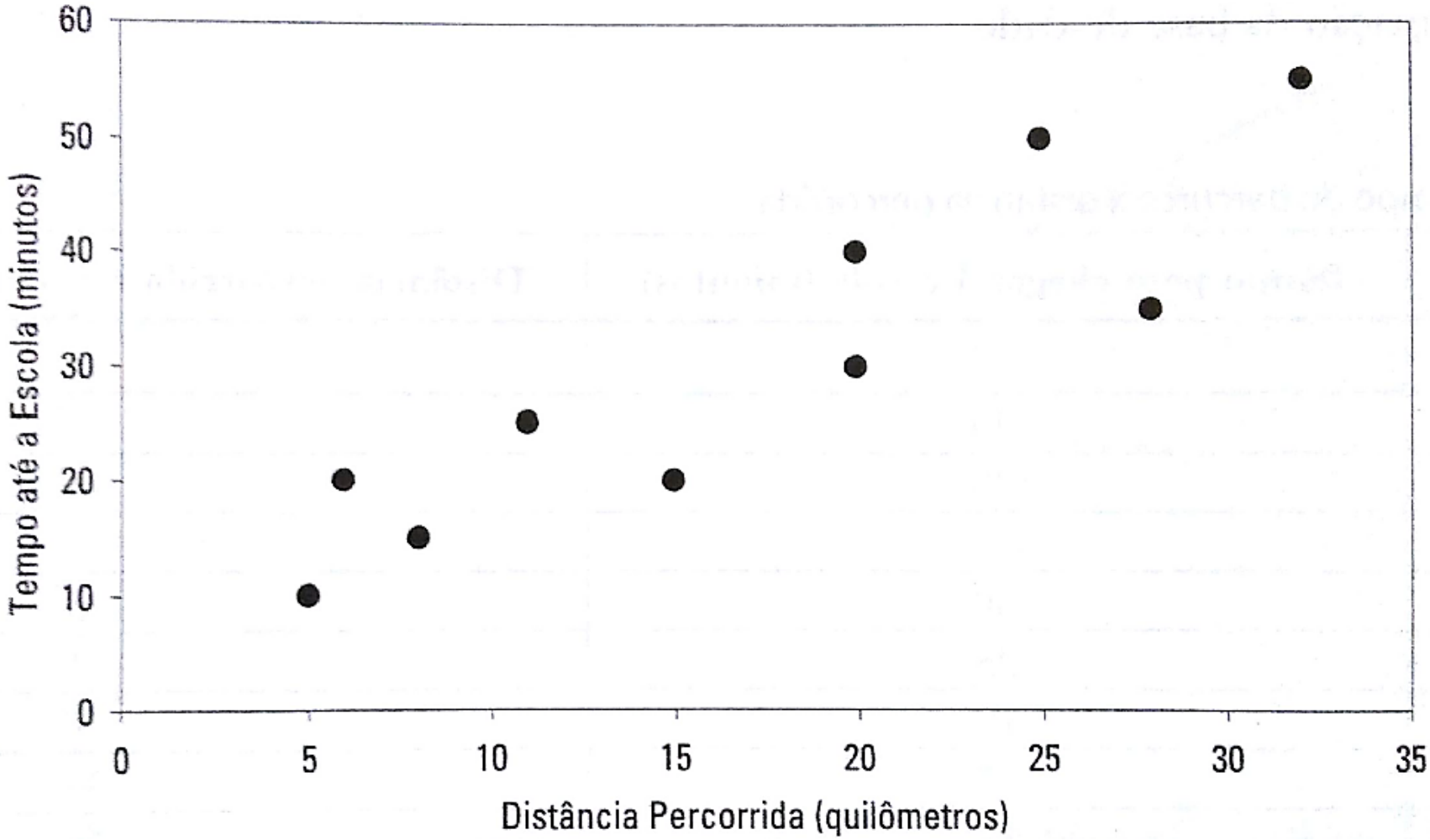
$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 . X_i + \hat{u}_i$$

- $\hat{\beta}_0, \hat{\beta}_1$: estimadores de β_0, β_1

EXEMPLO

	estudante	tempo	dist	sem	per	perfil
1	Gabriela	15	8	0	manha	calmo
2	Dalila	20	6	1	manha	moderado
3	Gustavo	20	15	0	manha	moderado
4	Leticia	40	20	1	tarde	agressivo
5	Luiz	50	25	2	tarde	agressivo
6	Leonor	25	11	1	manha	moderado
7	Ana	15	5	0	tarde	calmo
8	Antonio	55	32	3	tarde	calmo
9	Julia	35	28	1	manha	moderado
10	Mariana	30	20	1	manha	moderado

ESTIMAÇÃO DO MODELO DE REGRESSÃO LINEAR



Exemplo

$$tempo = f(dist)$$

O que gostaríamos de obter:

$$tempo = \beta_0 + \beta_1 . dist_i + u_i$$

O que vamos obter:

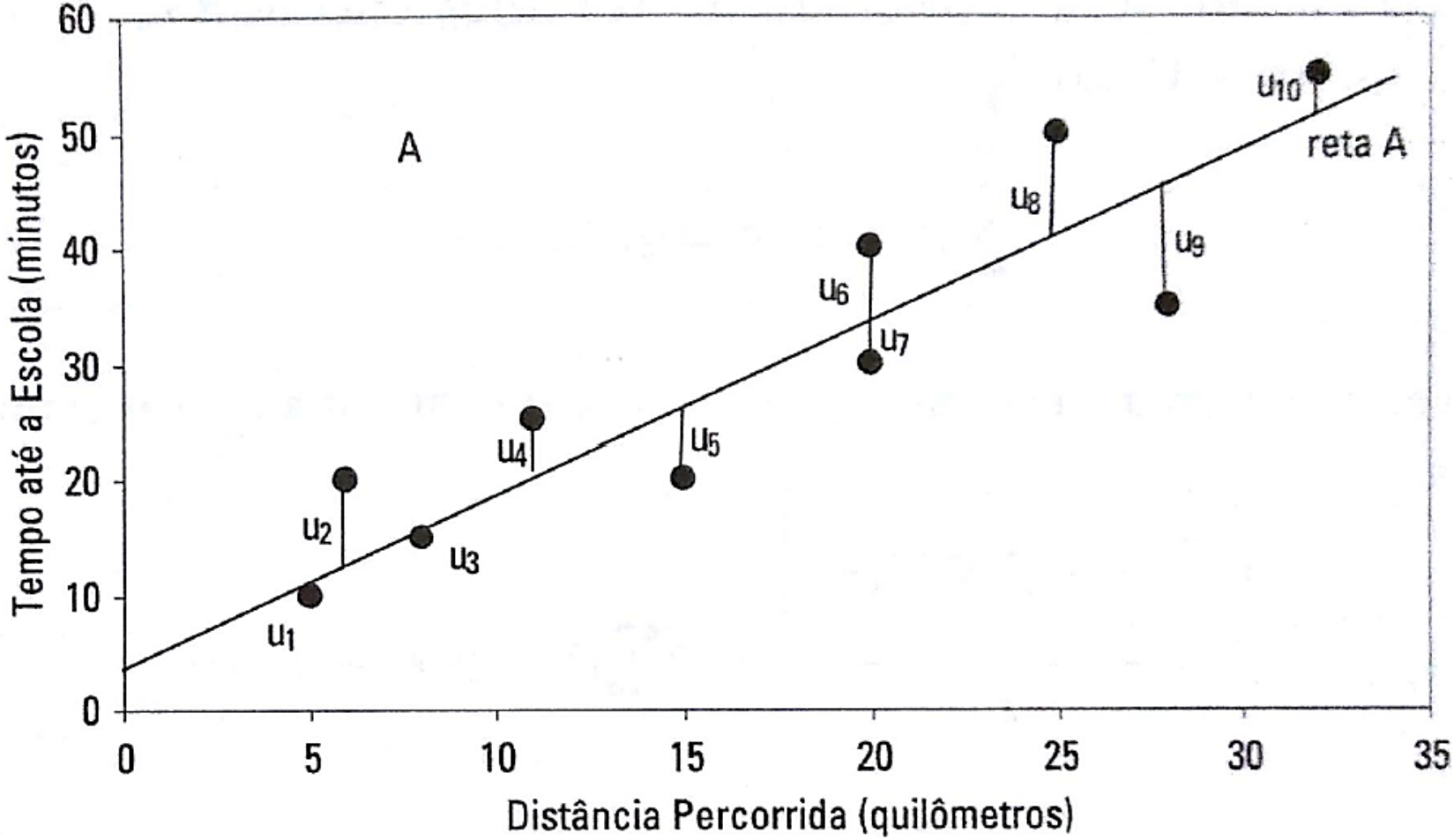
$$tempo = \hat{\beta}_0 + \hat{\beta}_1 . dist_i + \hat{u}_i$$

ESTIMAÇÃO DO MODELO DE REGRESSÃO LINEAR

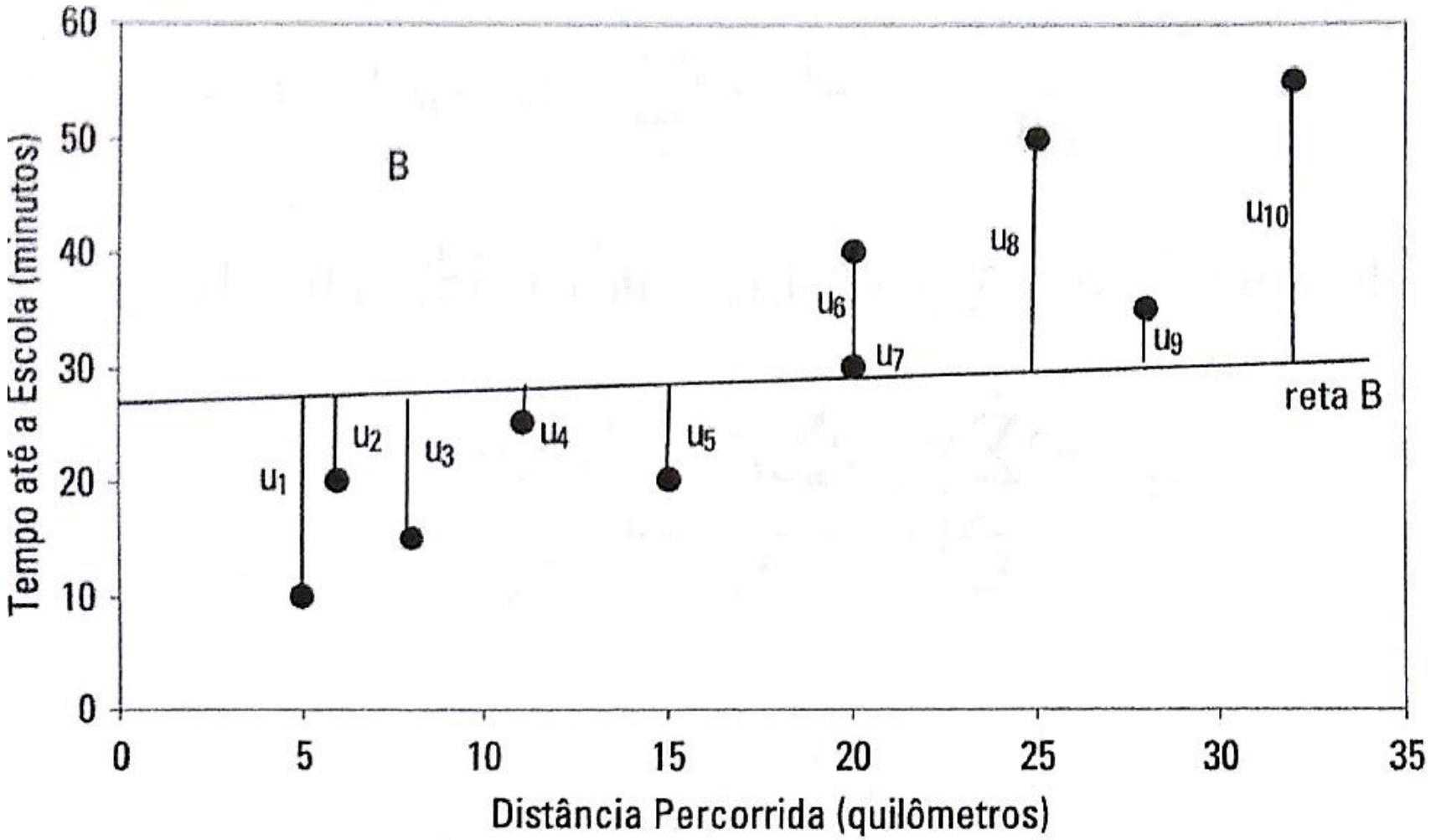
- 1º critério para estimar β_0 e β_1
  - Encontrar $\hat{\beta}_0$ e $\hat{\beta}_1$ tal que:

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

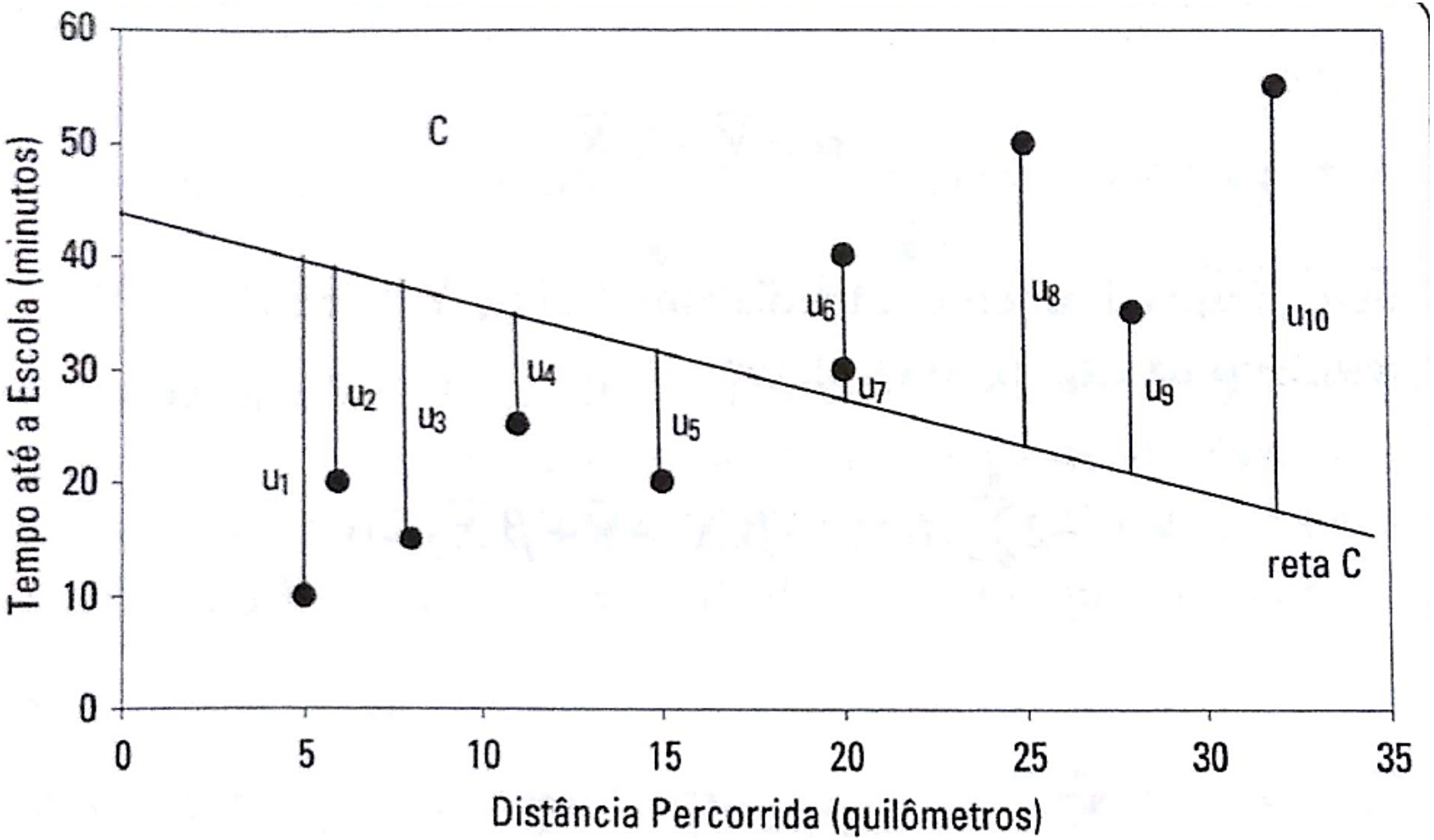
ESTIMAÇÃO DO MODELO DE REGRESSÃO LINEAR



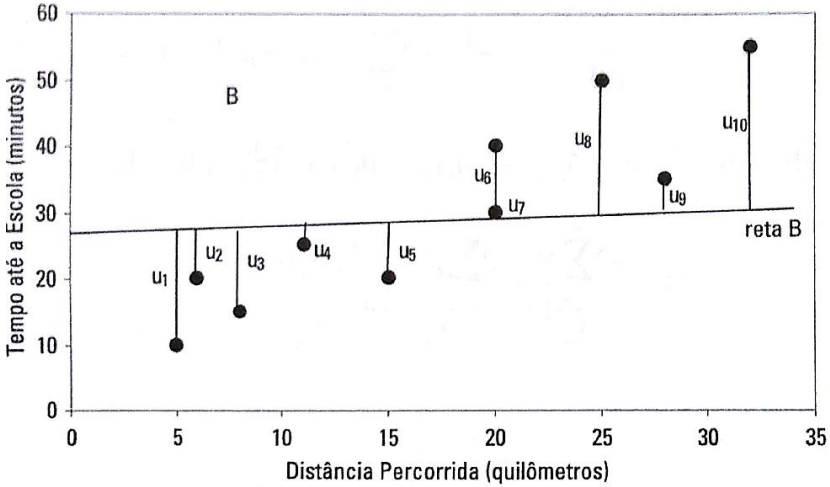
ESTIMAÇÃO DO MODELO DE REGRESSÃO LINEAR

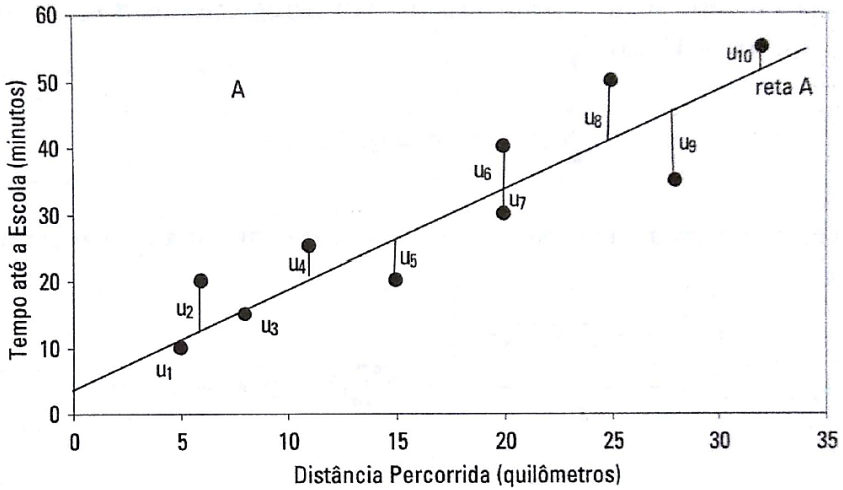


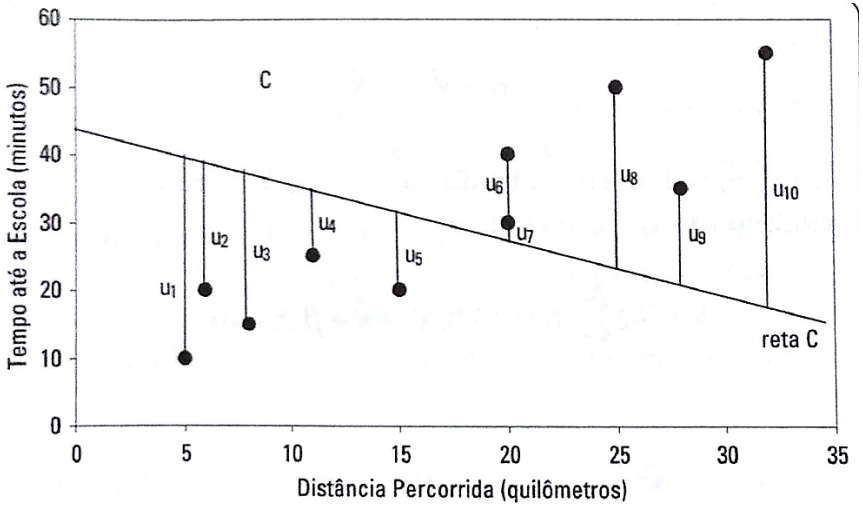
ESTIMAÇÃO DO MODELO DE REGRESSÃO LINEAR



ESTIMAÇÃO DO MODELO DE REGRESSÃO LINEAR







ESTIMAÇÃO DO MODELO DE REGRESSÃO LINEAR

- Para uma mesma base de dados, diversas equações da reta podem ser encontradas de modo que a soma dos resíduos seja zero
- 2º critério para estimar β_0 e β_1
  - Encontrar $\hat{\beta}_0$ e $\hat{\beta}_1$ tal que:

$$\sum_{i=1}^{n} \hat{u}_i^2 = \text{mín.}$$

ESTIMADORES DE MÍNIMOS QUADRADOS ORDINÁRIOS

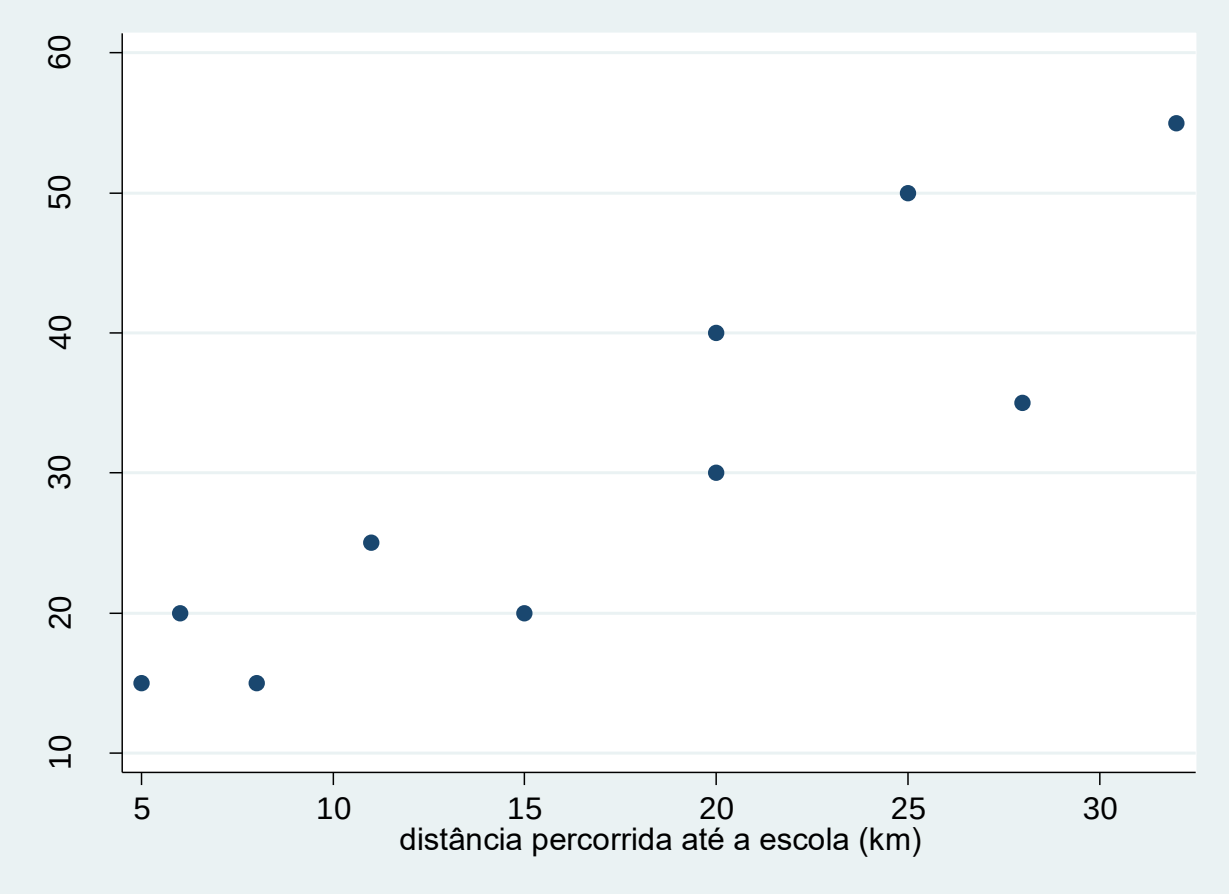
- Encontrar $\hat{\beta}_0$ e $\hat{\beta}_1$ que minimizem $\sum_{i=1}^{n} \hat{u}_i^2 = \{y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i)\}^2$

- Tomamos $\frac{\partial \sum_{i=1}^{n} \hat{u}_i^2}{\partial \hat{\beta}_0} = 0$ e $\frac{\partial \sum_{i=1}^{n} \hat{u}_i^2}{\partial \hat{\beta}_1} = 0$ para encontrar:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (X_i - \bar{X}).(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2}$$

EXEMPLO

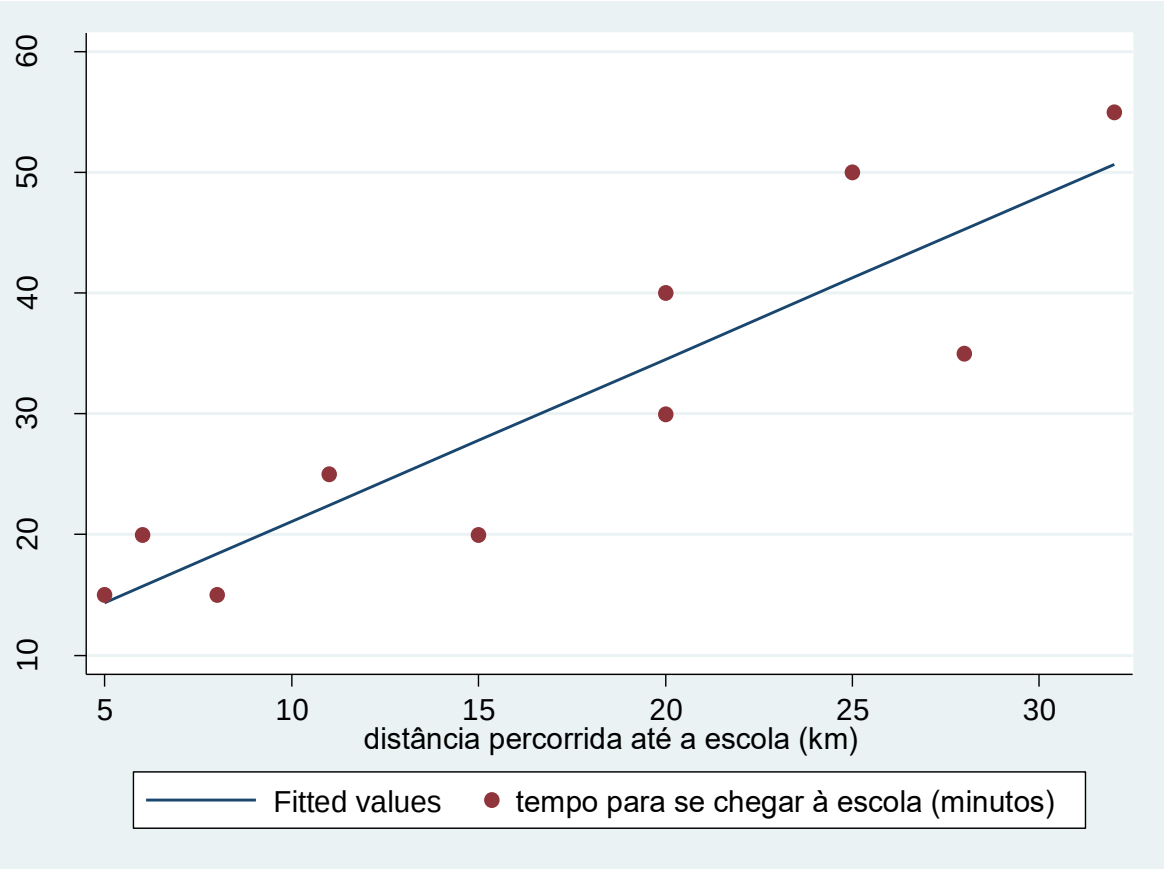

60
50
40
30
20
10
5
10
15
20
25
30
distância percorrida até a escola (km)

EXEMPLO

	estudante	tempo	dist	sem	per	perfil
1	Gabriela	15	8	0	manha	calmo
2	Dalila	20	6	1	manha	moderado
3	Gustavo	20	15	0	manha	moderado
4	Leticia	40	20	1	tarde	agressivo
5	Luiz	50	25	2	tarde	agressivo
6	Leonor	25	11	1	manha	moderado
7	Ana	15	5	0	tarde	calmo
8	Antonio	55	32	3	tarde	calmo
9	Julia	35	28	1	manha	moderado
10	Mariana	30	20	1	manha	moderado

EXEMPLO

$$Tempo = 7{,}63 + 1{,}35.distancia$$



REGRESSÃO LINEAR

- No Stata, arquivo Tempodistsemperperfil.dta

desc *(comando que descreve as variáveis)*

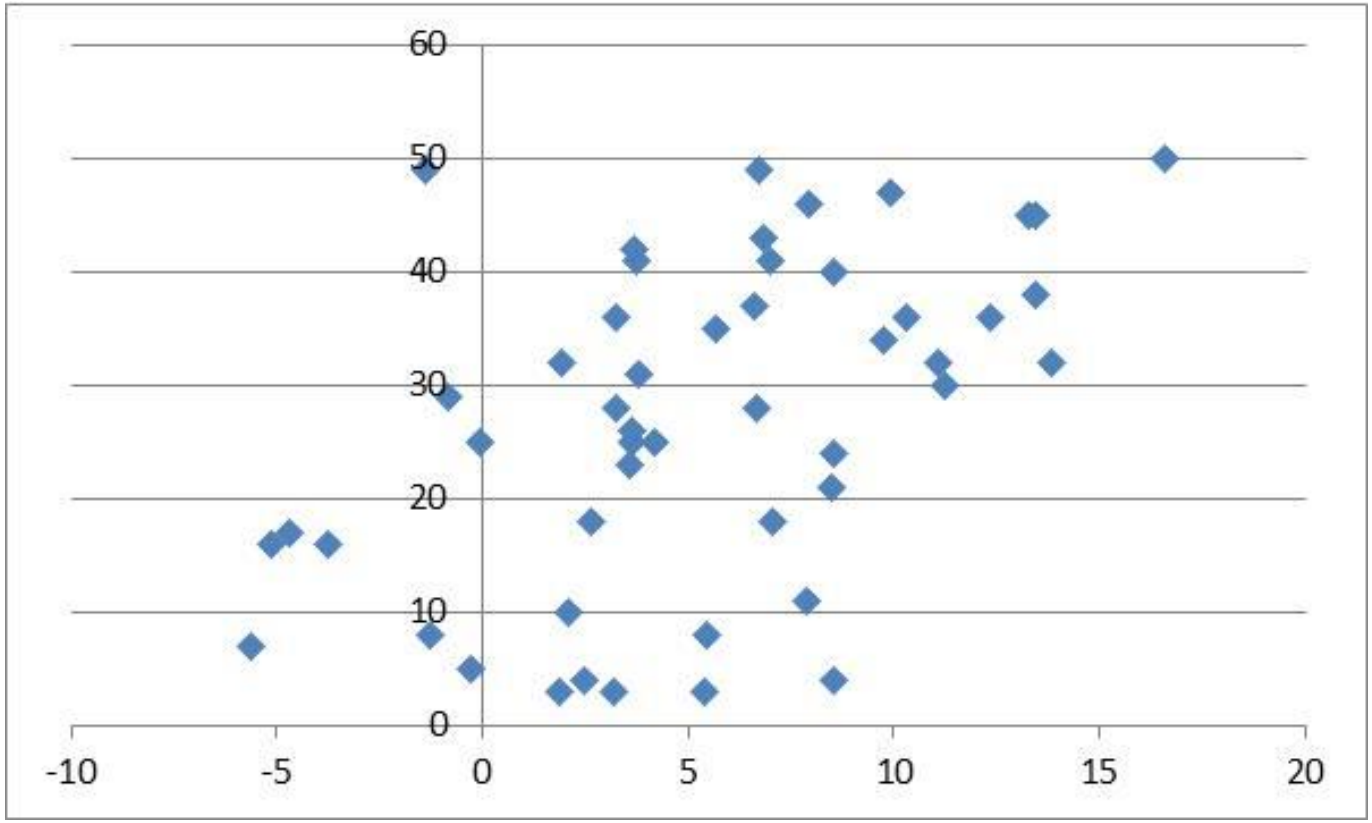
reg tempo dist *(comando que faz a regressão de tempo contra dist)*

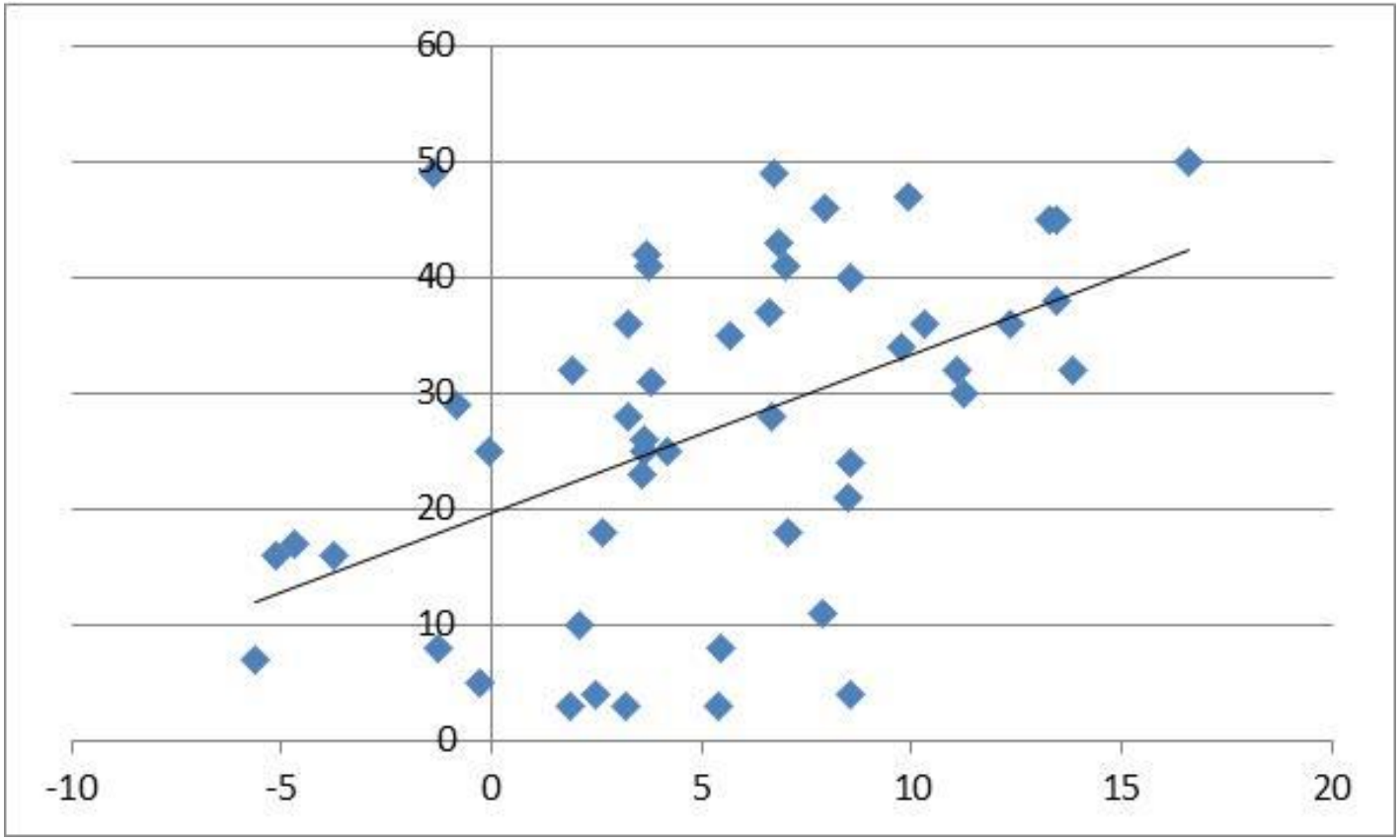
predict yhat *(comando que calcula o y previsto)*

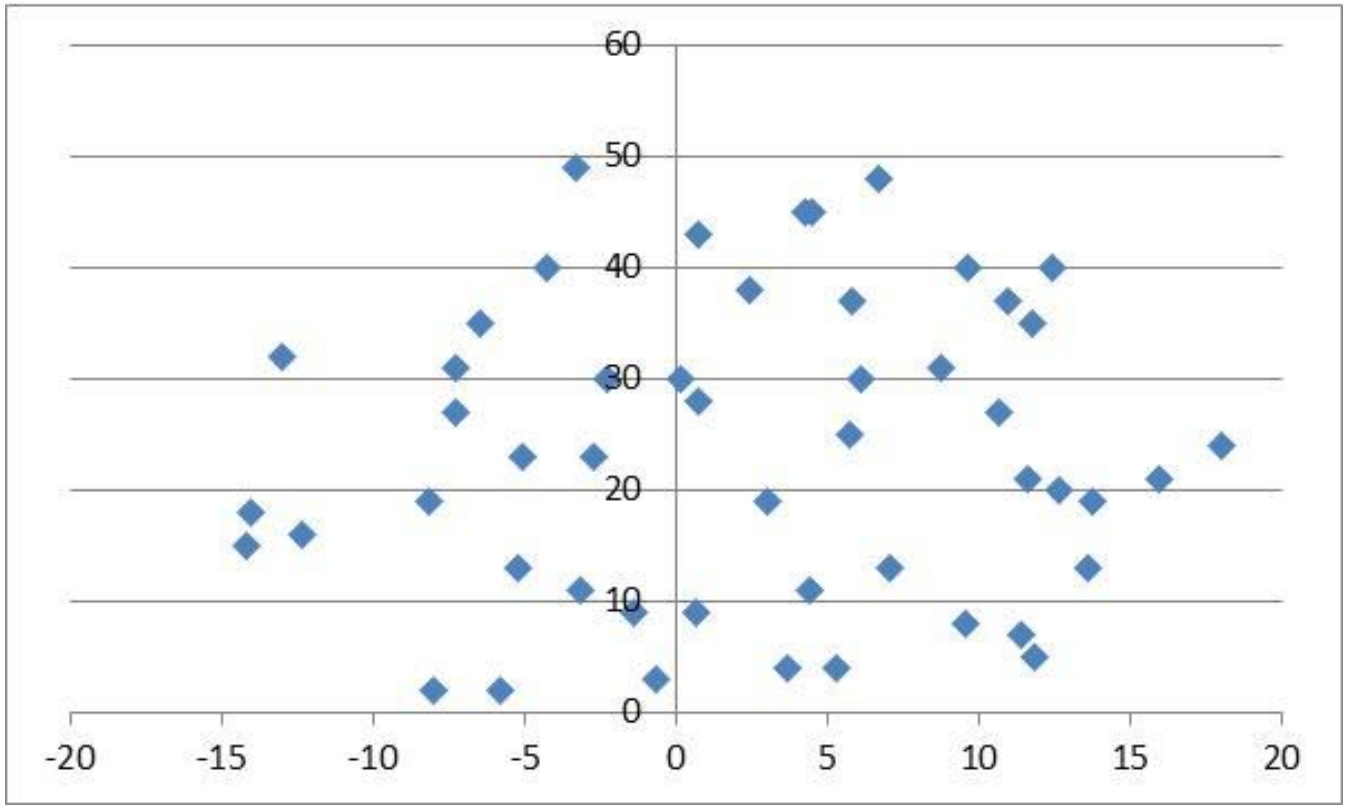
twoway (line yhat dist, sort) (scatter tempo dist) *comando que faz o gráfico de duas variáveis*

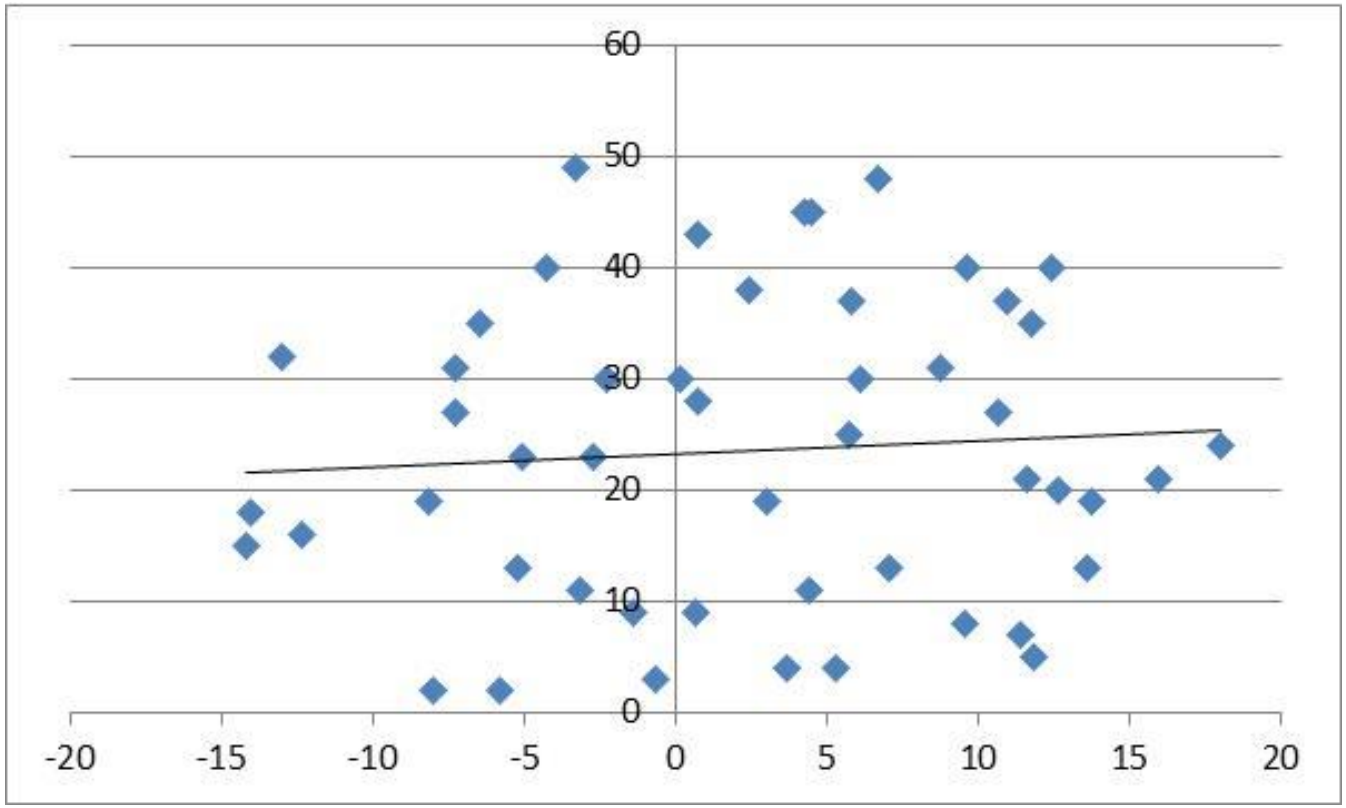
ANÁLISE DOS RESULTADOS DA REGRESSÃO

- O modelo é válido?
- Os coeficientes são estatisticamente significativos?
- Os sinais dos coeficientes estimados concordam com as expectativas teóricas?
- Qual o grau de ajuste do modelo?
- Os pressupostos da regressão foram atendidos?


60
50
40
30
20
10
0
-10
-5
0
5
10
15
20


60
50
40
30
20
10
0
-10
-5
0
5
10
15
20


60
50
40
30
20
10
0
-20
-15
-10
-5
0
5
10
15
20


60
50
40
30
20
10
0
-20
-15
-10
-5
0
5
10
15
20

Significância geral do modelo de regressão
O modelo é válido?

- É importante analisar a significância estatística geral do modelo estimado

- No caso de um modelo de regressão simples

$$Y_i = \beta_0 + \beta_1 . X_i + u_i$$

- Queremos avaliar se:

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

Significância geral do modelo de regressão
O modelo é válido?

- É importante analisar a significância estatística geral do modelo estimado
- No caso de um modelo de regressão múltipla

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

- Queremos avaliar se:

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0$$

$$H_1 : \text{existe um } \beta_j \neq 0$$

Significância geral do modelo de regressão
O modelo é válido?

Modelo de regressão estatisticamente significante:

- Fornece evidências de que existe relação linear entre a variável dependente Y e pelo menos uma das variáveis explicativas X_i
- A fração da variância explicada é significativamente maior que a fração da variância não explicada

$$\text{Estatística de teste} = \frac{variancia\ explicada}{variancia\ nao\ explicada}$$

TESTE F

- Um teste F é qualquer teste de hipóteses em que a estatística de teste segue uma distribuição F sob a hipótese nula

- É frequentemente usado na comparação de **modelos que tenham sido ajustados a um conjunto de dados**

TESTE F

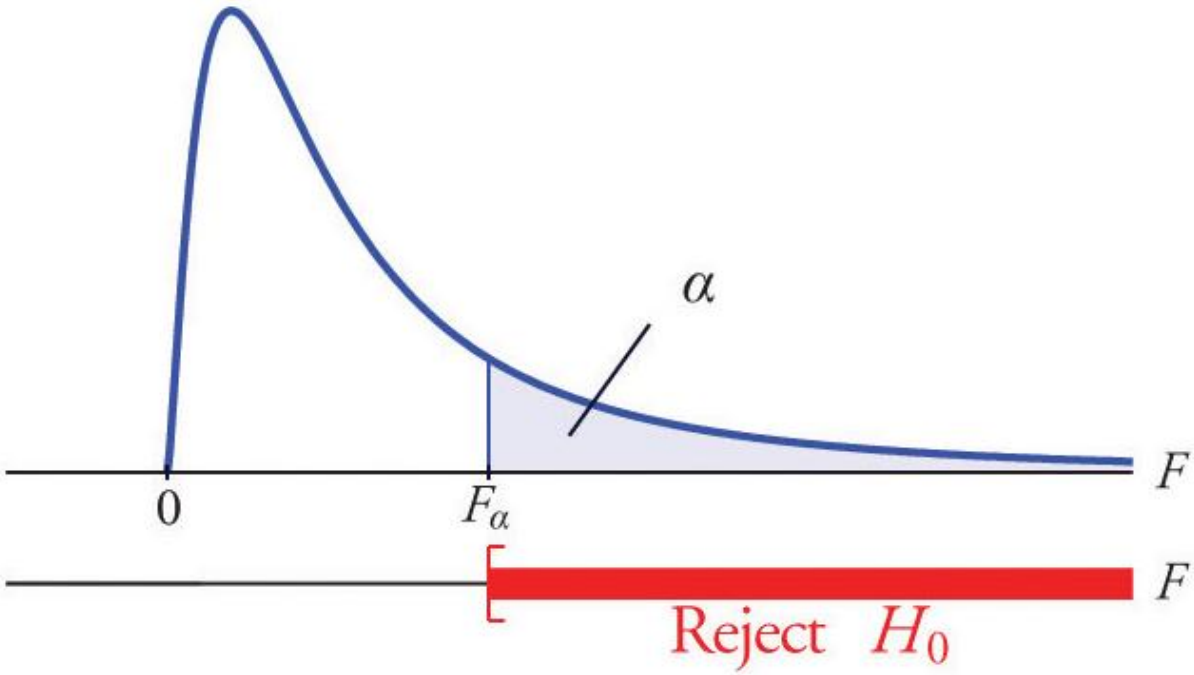
- A estatística de teste F é:

$$F = \frac{variancia\ explicada}{variancia\ nao\ explicada}$$

No caso da análise de regressão:

$$F = \frac{\dfrac{\sum_{i=1}^{n}\left(\hat{Y}_i - \bar{Y}\right)^2}{k-1}}{\dfrac{\sum_{i=1}^{n}(u_i)^2}{n-k}} = \frac{\dfrac{SQR}{k-1}}{\dfrac{SQU}{n-k}}$$

TESTE F



Significância geral do modelo de regressão
O modelo é válido?

```
. reg tempo dist

      Source |       SS       df       MS              Number of obs =      10
-------------+------------------------------           F(  1,     8) =   33.72
       Model |  1473.00369     1  1473.00369           Prob > F      =  0.0004
    Residual |  349.496314     8  43.6870393           R-squared     =  0.8082
-------------+------------------------------           Adj R-squared =  0.7843
       Total |     1822.5      9     202.5              Root MSE      =  6.6096
```

Pesos

Queremos avaliar se:

$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$

H_1: existe um $\beta_j \neq 0$

$$F = \frac{1473}{43{,}69} = 33{,}72$$

- Como $Prob > F$ é menor que o nível de significância α, há indícios de que pelo menos um $\beta_j \neq 0$

Significância geral do modelo de regressão
O modelo é válido?

```
. reg tempo dist

      Source |       SS       df       MS              Number of obs =      10
-------------+------------------------------           F(  1,     8) =   33.72
       Model |  1473.00369     1  1473.00369           Prob > F      =  0.0004
    Residual |  349.496314     8  43.6870393           R-squared     =  0.8082
-------------+------------------------------           Adj R-squared =  0.7843
       Total |      1822.5     9       202.5           Root MSE      =  6.6096
```

- Como a significância observada é 0,0004, podemos afirmar que o modelo de regressão é estatisticamente significante a 1% (uma vez que 0,0003 < 0,01)

Significância dos parâmetros
Os coeficientes são estatisticamente significantes?

- No caso de um modelo de regressão múltipla

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

- Queremos avaliar se:

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k = 0$$

$$H_1 : \text{existe um } \beta_j \neq 0$$

- Caso o teste F aponte para a rejeição de H_0, ainda assim não saberemos quais das variáveis explicativas X_i apresentarão coeficientes estatisticamente significativos

Significância dos parâmetros
Os coeficientes são estatisticamente significantes?

- Assim, é preciso realizar testes de hipóteses individuais para cada um dos coeficientes estimados $\beta_1, \beta_2, \dots, \beta_k$
- Ou seja, precisamos testar se:

$$H_0: \beta_k = 0$$

$$H_1: \beta_j \neq 0$$

- A estatística de teste para cada coeficiente é

$$t_{\beta_j} = \frac{\beta_j - 0}{e.p.(\beta_j)}$$

- Onde e.p. é o erro padrão do estimador de β_j

Significância dos parâmetros
Os coeficientes são estatisticamente significantes?

```
. reg tempo dist

      Source |       SS       df       MS              Number of obs =      10
-------------+------------------------------           F(  1,     8) =   33.72
       Model |  1473.00369     1  1473.00369           Prob > F      =  0.0004
    Residual |  349.496314     8  43.6870393           R-squared     =  0.8082
-------------+------------------------------           Adj R-squared =  0.7843
       Total |      1822.5     9       202.5           Root MSE      =  6.6096

------------------------------------------------------------------------------
       tempo |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        dist |   1.345209    .231667     5.81   0.000     .8109839    1.879434
       _cons |    7.63145   4.458611     1.71   0.125    -2.650126    17.91303
------------------------------------------------------------------------------
```

- O que podemos concluir a respeito da significância dos coeficientes?

Poder explicativo do modelo de regressão: R^2
Qual o grau de ajuste do modelo?

- Percentual da variabilidade de Y que é explicada pelo comportamento das variáveis explicativas incluídas no modelo

Poder explicativo do modelo de regressão: R^2
Qual o grau de ajuste do modelo?

- Conceitos importantes para analisar o poder explicativo do modelo:
  - Soma total dos quadrados (STQ)

$$\text{STQ} = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

  - Soma dos quadrados da regressão (SQE)

$$\text{SQR} = \sum_{i=1}^{n} \left(\hat{Y}_i - \bar{Y}\right)^2$$

  - Soma dos quadrados dos resíduos (SQU)

$$\text{SQU} = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2$$

Poder explicativo do modelo de regressão: R^2
Qual o grau de ajuste do modelo?

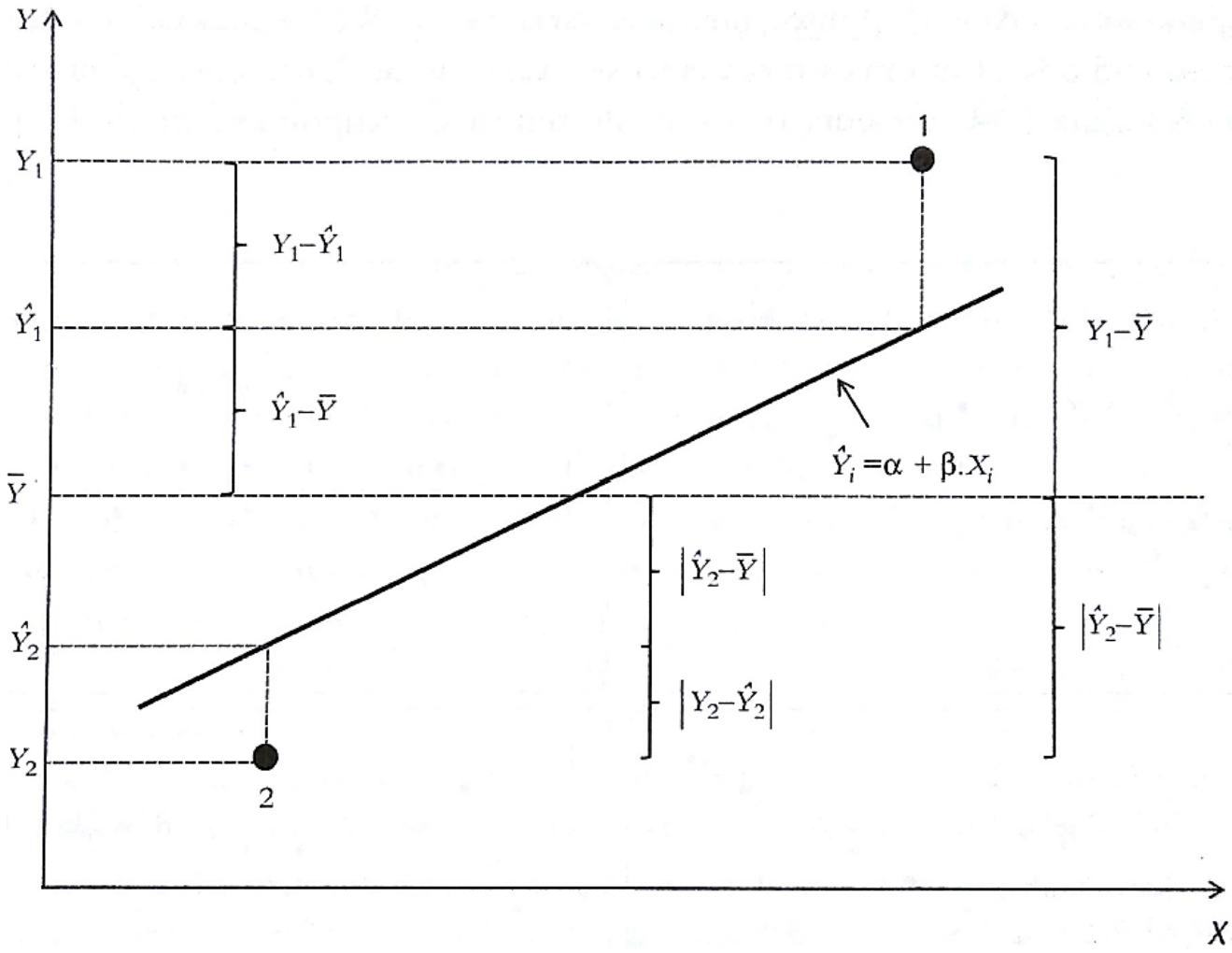
- Conceitos importantes para analisar o poder explicativo do modelo:

$$SQT = SQR + SQU$$

- De modo que

$$\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \sum_{i=1}^{n}\left(\hat{Y}_i - \bar{Y}\right)^2 + \sum_{i=1}^{n}(u_i)^2$$

Poder explicativo do modelo de regressão: R^2
Qual o grau de ajuste do modelo?

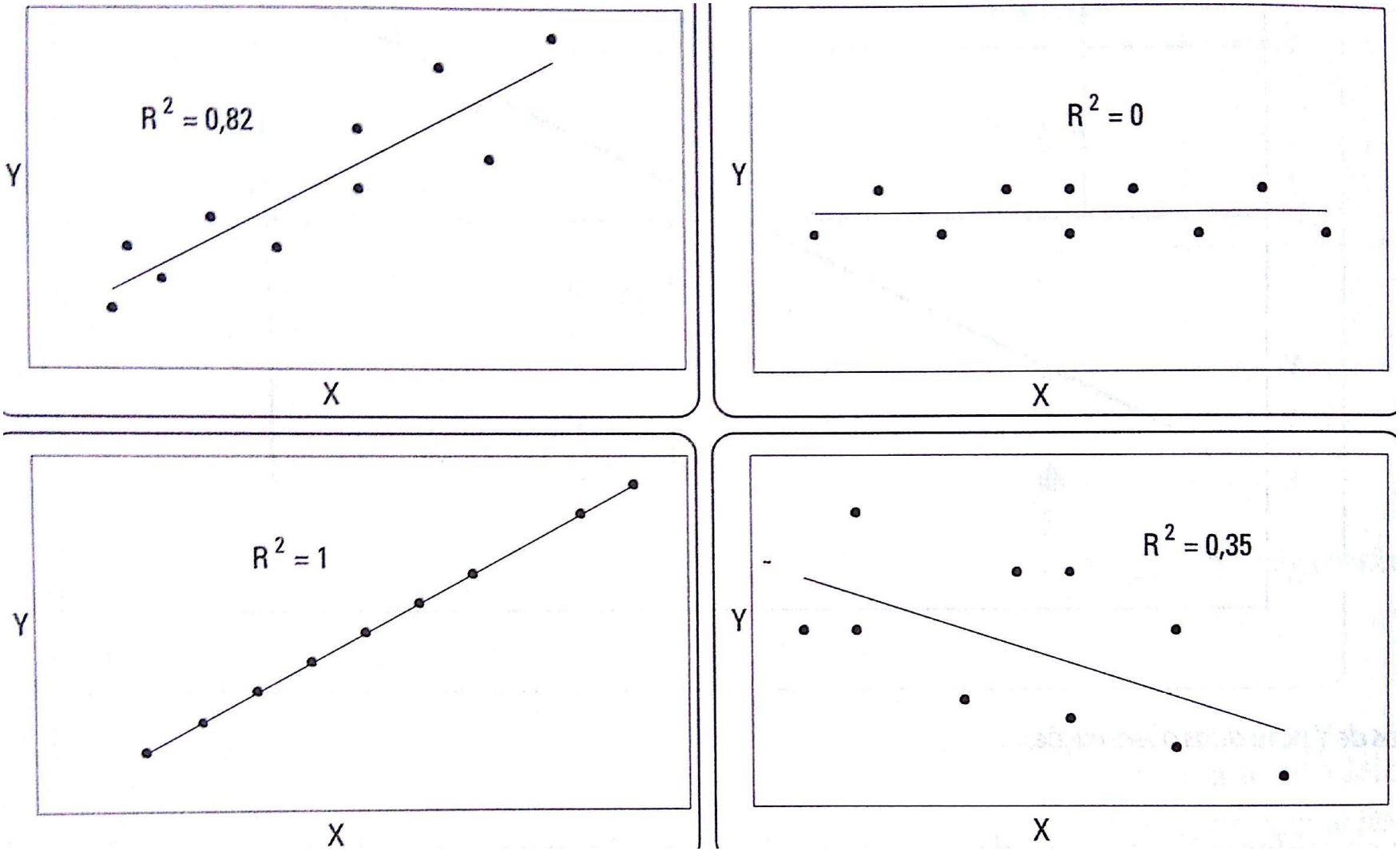


COEFICIENTE DE AJUSTE R^2

$$R^2 = \frac{SQR}{SQR + SQU} = \frac{SQR}{SQT}$$

- R^2
  - Pode variar entre 0 e 1 (0 a 100%)
  - Representa a fração da variância de Y explicada pelas variáveis explicativas
  - Representa o grau de ajuste do modelo aos dados
  - Não estamos nos referindo a causalidade

COEFICIENTE DE AJUSTE R^2



COEFICIENTE DE AJUSTE R^2

$$R^2 = \frac{SQR}{SQR + SQU} = \frac{SQR}{SQT}$$

```
. reg tempo dist

      Source |       SS       df       MS              Number of obs =      10
-------------+------------------------------           F(  1,     8) =   33.72
       Model |  1473.00369     1  1473.00369           Prob > F      =  0.0004
    Residual |  349.496314     8  43.6870393           R-squared     =  0.8082
-------------+------------------------------           Adj R-squared =  0.7843
       Total |     1822.5     9       202.5           Root MSE      =  6.6096
```

$$R^2 = \frac{1473}{1473 + 349{,}49} = \frac{1474}{1822{,}50} = 0{,}8082$$

COEFICIENTE DE AJUSTE R^2

O coeficiente de ajuste R^2 **não informa**:

- Se uma determinada variável explicativa é estatisticamente significante
- Se uma variável explicativa é a causa verdadeira de variação da variável dependente
- Se existem problemas de omissão de variáveis importantes
- Se existem variáveis irrelevantes incluídas no modelo

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

- Relaciona os retornos de um título e os retornos da carteira de mercado, usando dados históricos
  - Premissa: os retornos verificados no passado são previsivelmente repetidos no futuro.

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

$$r_{it} = \alpha_i + \beta_i r_{mt} + u_{it}$$

- r_{it}: retorno do ativo i no momento t
- r_{mt}: retorno da carteira de mercado no momento t
- α_i: intercepto, representa o retorno esperado em excesso de um ativo. Tanto no caso de ações quanto de portfolios, espera-se $\alpha = 0$.
- β_i: coeficiente angular
  - Representa o risco do ativo em relação ao mercado

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

$$\beta_i = \frac{COV_{r_i,r_M}}{VAR_{r_M}}$$

- A fórmula de β_i é a mesma do estimador de mínimos quadrados
- Exprime o risco de um ativo **em relação** à carteira de mercado, que tem seu risco minimizado pela diversificação
- Em geral usa-se dados mensais, em períodos de 3 a 5 anos, ou dados semanais em períodos de um ano para estimar o beta de uma ação

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

$$\beta_i = \frac{COV_{r_i,r_M}}{VAR_{r_M}}$$

- Ação com **alto** beta: relativamente **sensível** a movimentações do mercado
- Ação com **baixo** beta: relativamente **insensível** a movimentações do mercado
- $\beta_i > 1$: a ação é mais volátil que a carteira de mercado
- $\beta_i = 1$: a ação é igualmente volátil à carteira de mercado
- $\beta_i < 1$: a ação é menos volátil que a carteira de mercado

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

$$r_{it} = \alpha_i + \beta_i r_{mt} + u_{it}$$

- $\alpha_i < 0$: a ação resultou em um retorno inferior ao esperado para o seu nível de risco
- $\alpha_i = 0$: a ação resultou em um retorno igual ao esperado para o seu nível de risco
- $\alpha_i > 0$: a ação resultou em um retorno superior ao esperado para o seu nível de risco

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

Estimação dos parâmetros

$$r_{it} = \alpha_i + \beta_i r_{mt} + u_{it}$$

- α_i: retorno em excesso da ação, em relação ao retorno esperado de acordo com o seu nível de risco
- β_i: fator de risco da ação em relação ao risco da carteira de mercado
- Variância dos resíduos: medida do risco diversificável da ação
- R^2 da regressão: parte do risco da ação que é explicada pelas condições de mercado, o chamado risco sistemático (não diversificável)
- $1 - R^2$: fração do risco da ação que pode ser eliminada pela diversificação

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

Exercícios: arquivo Ret.acoes.ibov.dta

tsset t informa ao Stata que os dados são em série de tempo, e que *t* é a variável de tempo

reg vale ibov

```
. reg vale ibov

      Source |       SS       df       MS              Number of obs =     247
-------------+------------------------------           F(  1,   245) =  209.37
       Model |  .184492761     1  .184492761           Prob > F      =  0.0000
    Residual |  .215886272   245  .000881168           R-squared     =  0.4608
-------------+------------------------------           Adj R-squared =  0.4586
       Total |  .400379033   246  .001627557           Root MSE      =  .02968

------------------------------------------------------------------------------
        vale |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ibov |   1.594979   .1102287    14.47   0.000     1.377862    1.812096
       _cons |   -.000846   .0018926    -0.45   0.655    -.0045739    .0028818
------------------------------------------------------------------------------
```

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

```
. reg vale ibov

      Source |       SS       df       MS              Number of obs =     247
-------------+------------------------------           F(  1,   245) =  209.37
       Model |  .184492761     1  .184492761           Prob > F      =  0.0000
    Residual |  .215886272   245  .000881168           R-squared     =  0.4608
-------------+------------------------------           Adj R-squared =  0.4586
       Total |  .400379033   246  .001627557           Root MSE      =  .02968

------------------------------------------------------------------------------
        vale |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ibov |   1.594979   .1102287    14.47   0.000     1.377862    1.812096
       _cons |   -.000846   .0018926    -0.45   0.655    -.0045739    .0028818
------------------------------------------------------------------------------
```

O beta da Vale é 1,59

O beta da Vale é estatisticamente diferente de zero

O beta da Vale é estatisticamente diferente de 1

alfa da Vale é zero

```
. test ibov==1

 ( 1)  ibov = 1

       F(  1,   245) =   29.13
            Prob > F =    0.0000
```

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

- Risco total de uma ação: risco sistemático + risco não sistemático
- Variância total uma ação = variância explicada + variância não explicada

reg vale ibov

predict res_vale, res (gera os resíduos da regressão anterior)

sum vale res_vale (apresenta a estatística descritiva das variáveis vale e res_vale)

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

```
. reg vale ibov

      Source |       SS       df       MS              Number of obs =     247
-------------+------------------------------           F(  1,   245) =  209.37
       Model |  .184492761     1  .184492761           Prob > F      =  0.0000
    Residual |  .215886272   245  .000881168           R-squared     =  0.4608
-------------+------------------------------           Adj R-squared =  0.4586
       Total |  .400379033   246  .001627557           Root MSE      =  .02968

------------------------------------------------------------------------------
        vale |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ibov |   1.594979   .1102287    14.47   0.000     1.377862    1.812096
       _cons |   -.000846   .0018926    -0.45   0.655    -.0045739    .0028818
------------------------------------------------------------------------------

. sum vale res_vale

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
        vale |        247    .0008923     .040343     -.1205      .1135
    res_vale |        247    3.15e-11    .0296241  -.1150285   .0859826
```

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

- A regressão dos retornos da Vale contra a carteira de mercado produziu um $R^2 = 0{,}461$, e consequentemente $1 - R^2 = 0{,}539$
  - 46,1% da variância dos retornos da Vale é explicada pelas condições de mercado que afetam todas as ações (**risco sistemático, não diversificável**)
  - 53,9% da variância dos retornos da Vale é devido a fatores únicos da empresa (**risco não sistemático, diversificável**)

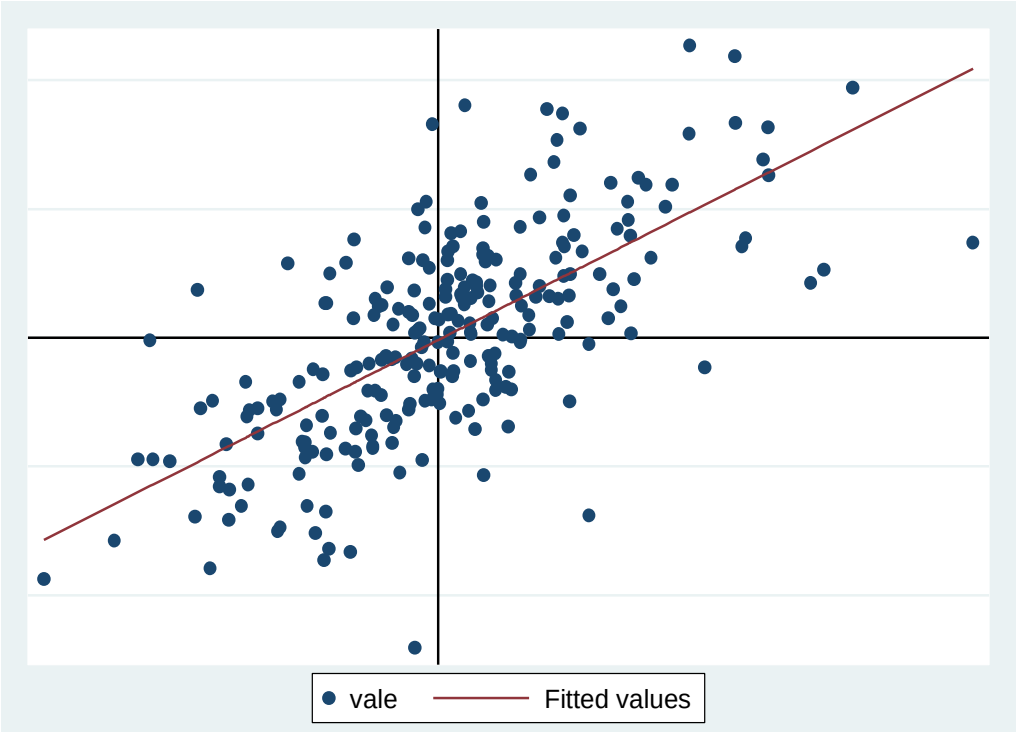
APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

- Síntese da análise da reta característica da ação da Vale (VALE5)
  - Seu beta é 1,59
  - Seu alfa é zero
  - 53,9% da sua volatilidade pode ser eliminada por meio da diversificação (46,1% da sua volatilidade é sistemática)
  - Sua reta característica é -0,0008 + 1,5950.IBOV

    (0,0019) (0,1102)

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

- Síntese da análise da reta característica da ação da Vale (VALE5)
  - O gráfico da reta característica é:



APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

```
reg vale ibov
predict yhat_vale
twoway (scatter vale ibov) (line yhat_vale ibov, sort), yscale(off) yline(0, lcolor(black)) xscale(off) xline(0, lcolor(black))
```

Explicação do comando gráfico

twoway (scatter vale ibov) (line yhat_vale ibov, sort) gera um gráfico de dispersão vale x ibov e um gráfico de linha vale previsto x ibov

yscale(off) yline(0, lcolor(black)) opção de omitir o eixo y e traçar uma reta em y=0, com linha na cor preta

xscale(off) xline(0, lcolor(black)) opção de omitir o eixo x e traçar uma reta em x=0, com linha na cor preta

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

Figura 1: Reta característica da ação VALE5

APLICAÇÃO: A RETA CARACTERÍSTICA DE UMA AÇÃO

- Encontre a reta característica das ações
  - ITAU
  - AMBEV
  - BRADESCO
- Analise os parâmetros alfa e beta
- Avalie os componentes risco sistemático e não sistemático
- Trace o gráfico da reta característica