



SME0803

ANÁLISE DE REGRESSÃO

1º/2011

**Cibele Russo**

cibele@icmc.usp.br

<http://www.icmc.usp.br/~cibele>

Sala 3-160, ramal 8174

**CoteiaWIKI:**

<http://wiki.icmc.usp.br/index.php/SME-260>



## OBJETIVO DO CURSO

Introduzir as principais técnicas de  
Análise de Regressão

Trabalhar em aplicações de Análise de  
Regressão com e sem o auxílio de  
pacotes estatísticos



# PROGRAMA RESUMIDO DO CURSO

1. Regressão Linear Simples.
2. Regressão Linear Múltipla.
3. Análise de Resíduos
4. Métodos de Diagnóstico
5. Regressão polinomial
6. Variáveis "dummy"
7. Colinearidade
8. Seleção de variáveis
9. Modelos lineares generalizados (tópico adicional)
10. Regressão não linear (tópico adicional)



# CRITÉRIO DE AVALIAÇÃO

$$\text{Média Final MF} = \frac{4 P1 + 5 P2 + NL}{10}$$

P1 Prova 1: 09/05/2011

P2 Prova 2: 30/06/2011

NL Nota de Listas de Exercícios;

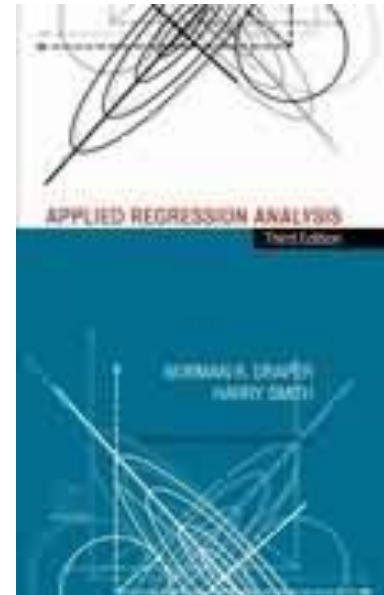
**Não haverá prova substitutiva.**

Critério de Recuperação: ver JúpiterWeb.

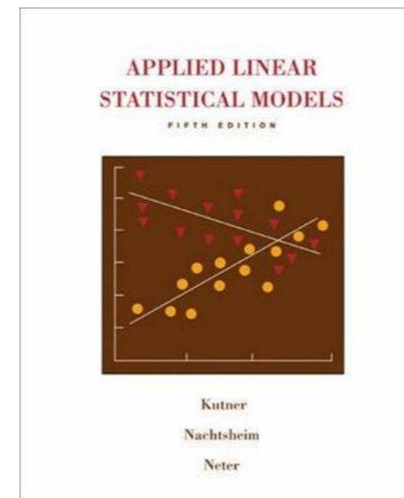


# BIBLIOGRAFIA PRINCIPAL

Draper, N.R., Smith, H. (1998).  
Applied Regression Analysis,  
3rd.ed., John Wiley.  
(disponível na biblioteca)

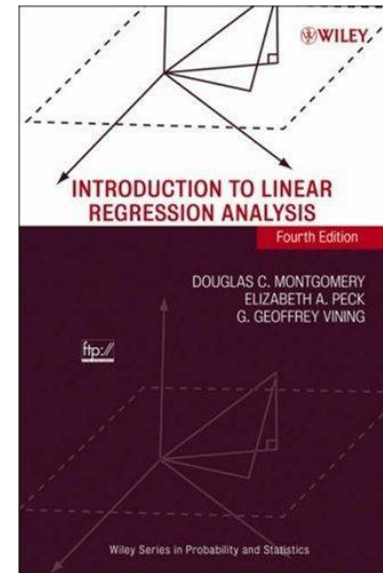


Neter, J. Kutner, M. H.  
Nachtsheim, C. J. and  
Wasserman, W. Applied Linear  
Statistical Models, IRWIN  
1996.  
(disponível na biblioteca a  
edição de 1985)

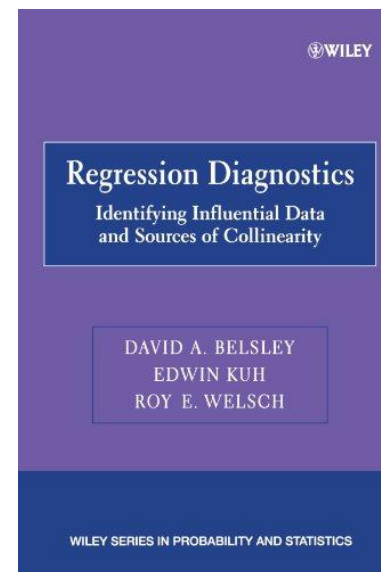


# BIBLIOGRAFIA COMPLEMENTAR

Montgomery, D. C., Peck, E. A. and Vining, G. G., Introduction to Linear Regression Analysis, 4th ed. New York. John Wiley, 2006.  
(disponível na biblioteca)



Belsley, D. A., Kuh, E., Welsch, R. E. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity, Wiley Series in Probability and Statistics, 2004.  
(disponível na biblioteca)

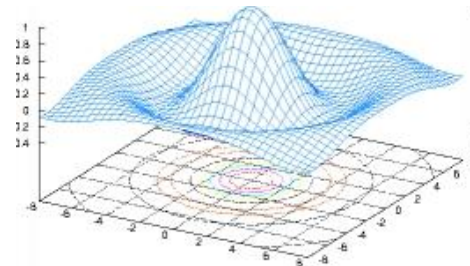
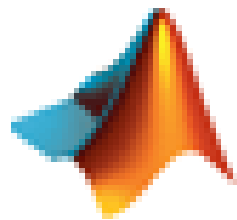
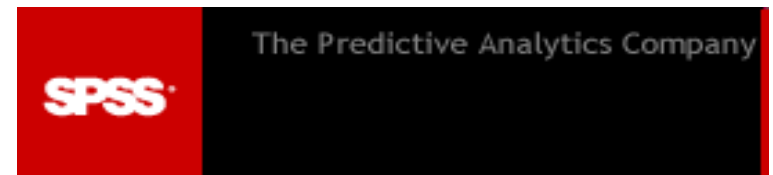


# APOIO COMPUTACIONAL

Excel.

R ([www.R-project.org](http://www.R-project.org)).

Minitab, S-PLUS, SAS, SPSS, Statistica, MatLab, Octave, Scilab, etc.



## O QUE É ESTATÍSTICA?

**Estatística** é um conjunto de técnicas e métodos de pesquisa que envolve o planejamento do experimento a ser realizado, a coleta qualificada dos dados, a inferência, o processamento, a análise e a disseminação das informações  
(ENCE - Escola Nacional de Ciências Estatísticas)



## HISTÓRIA DA ESTATÍSTICA

Desenvolvimento da Estatística, principalmente no Brasil:

<http://www.redeabe.org.br/historia.htm>

Memória, J. M. P., Breve História da Estatística. EMBRAPA, Brasília, 2004.

[http://www.embrapa.br/publicacoes/tecnico/folderTextoDiscussao/arquivos-pdf/Miolo\\_21.pdf/view](http://www.embrapa.br/publicacoes/tecnico/folderTextoDiscussao/arquivos-pdf/Miolo_21.pdf/view)

Materials for the History of Statistics:

<http://www.york.ac.uk/depts/maths/histstat/>

História da Estatística (por Prof. Gauss Cordeiro):

[http://www.des.uem.br/uploads/arquivos\\_professor/1125192908.pdf](http://www.des.uem.br/uploads/arquivos_professor/1125192908.pdf)

(Arquivo resumido)

[http://wiki.icmc.usp.br/images/2/20/Hist%C3%B3ria\\_da\\_Estat%C3%ADstica\\_-\\_Resumida.PDF](http://wiki.icmc.usp.br/images/2/20/Hist%C3%B3ria_da_Estat%C3%ADstica_-_Resumida.PDF)

# ALGUMAS DAS GRANDES ÁREAS DA ESTATÍSTICA

Probabilidades e Processos Estocásticos

Planejamento de experimentos

Amostragem

Modelos de Regressão

Séries Temporais

Métodos Bayesianos

Análise de Sobrevida

# ANÁLISE DE REGRESSÃO

O primeiro uso da palavra **regressão** surgiu no contexto estatístico por meio do antropologista e meteorologista Sir Francis Galton por volta de 1885, embora Legendre e Gauss já tivessem publicado trabalhos utilizando o método dos mínimos quadrados no início do século XIX.

Originalmente o termo utilizado por Galton foi **reversão**.

Os primeiros trabalhos de Galton buscavam explicar a altura dos filhos a partir da altura de seus pais.

O termo "Análise de Correlação" também é atribuído a Galton.



# O QUE É ANÁLISE DE REGRESSÃO?

Análise de regressão é uma **técnica estatística** para investigar e modelar a relação entre variáveis.

(Montgomery, Peck and Vining, 2006, pág. 1)



## EXEMPLO

Deseja-se investigar a relação entre o tempo que um indivíduo leva para reagir a um certo estímulo (em décimos de segundos) e algumas características como sexo, idade (em anos), acuidade visual (em %).

(Bussab, W., Análise de variância e regressão, 1986, Atual Editora)

# EXEMPLO

Ind	tempo	sexo	Idade	Acuidade_ visual
1	96	M	20	90
2	92	F	20	100
3	106	M	20	80
4	100	F	20	90
5	98	F	25	100
6	104	M	25	90
7	110	M	25	80
8	101	F	25	90
9	116	F	30	70
10	106	M	30	90
11	109	M	30	90
12	100	F	30	80
13	112	F	35	90
14	105	F	35	80
15	118	M	35	70
16	108	M	35	90
17	113	F	40	90
18	112	F	40	90
19	127	M	40	60
20	117	M	40	80



## EXEMPLO

Inicialmente vamos estudar a relação entre

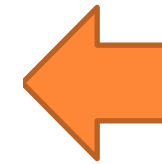
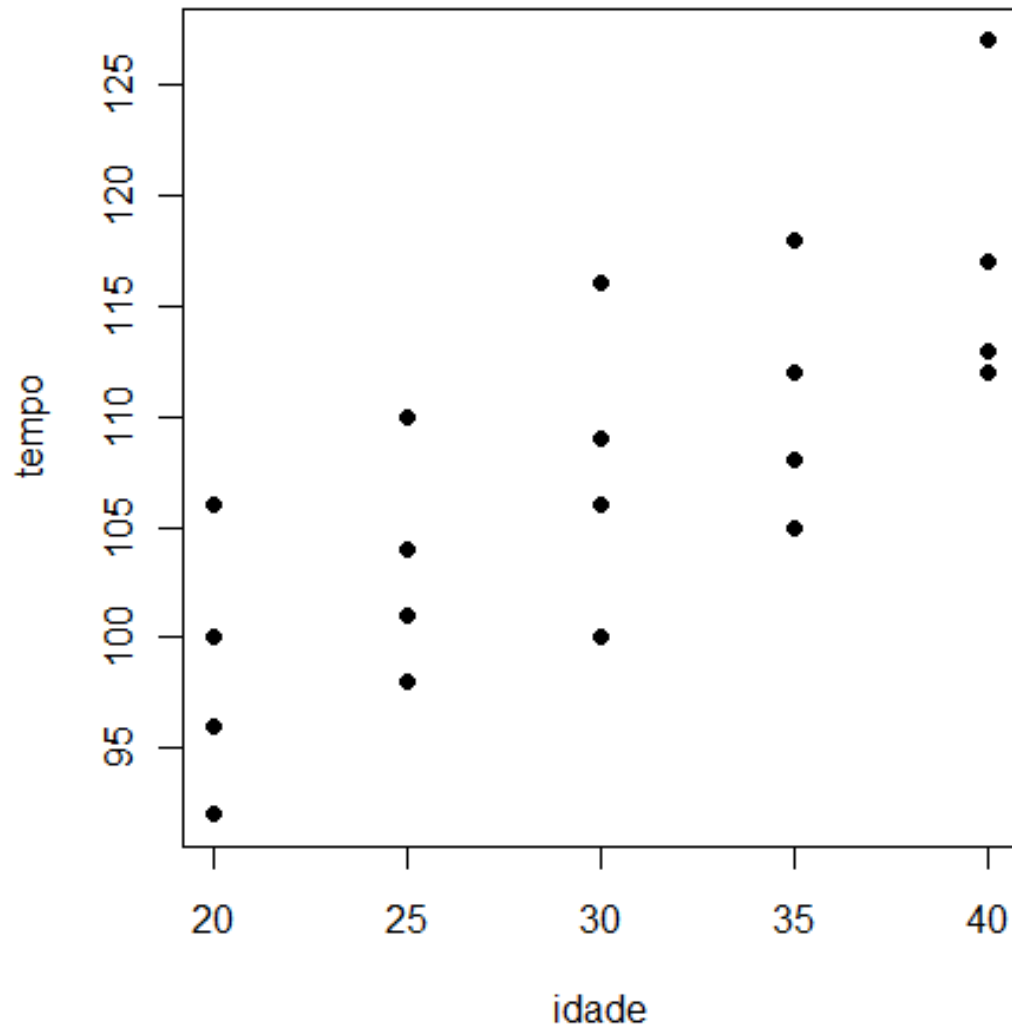
Y: tempo e X: idade

sem considerar as outras variáveis.

Inicialmente podemos construir o gráfico de dispersão entre a idade (X) e o tempo (Y).



# EXEMPLO

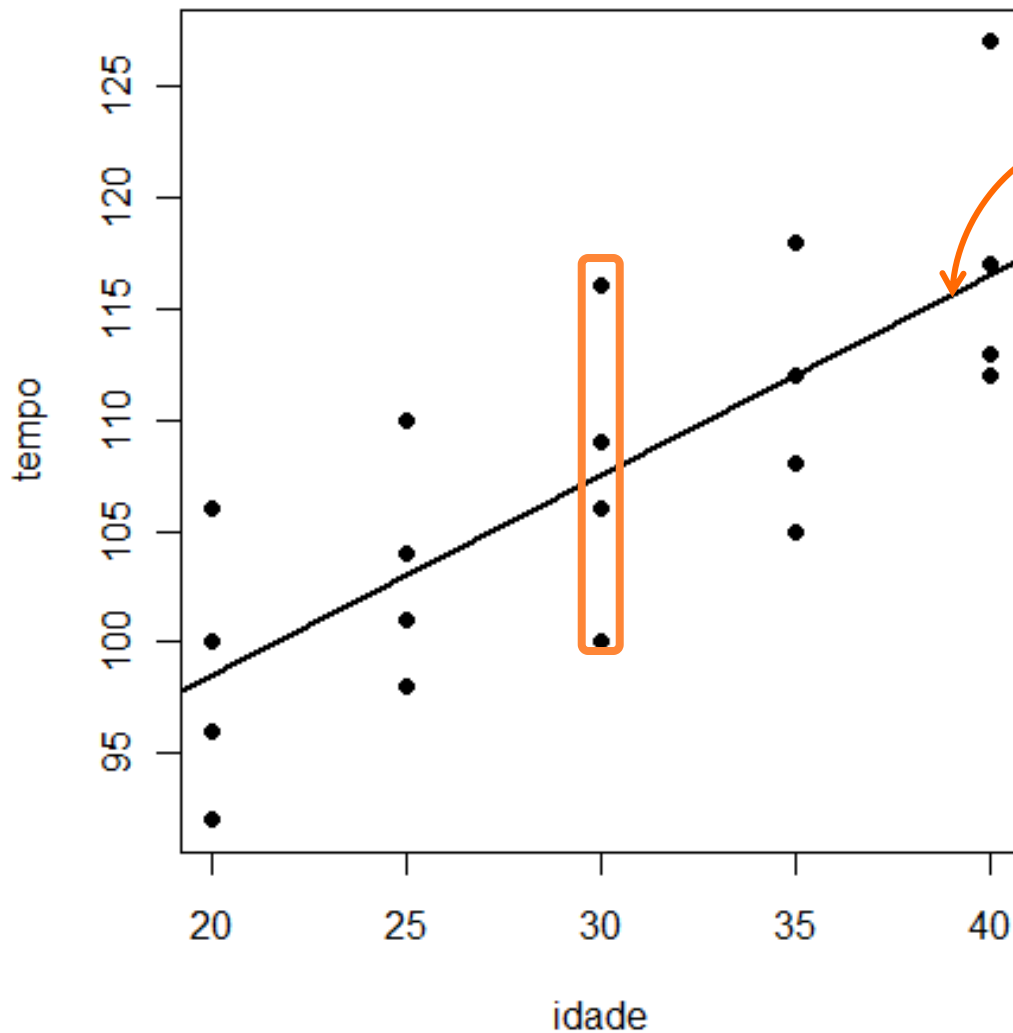


Pode-se notar uma tendência de aumento no tempo de reação com o aumento da idade.





# EXEMPLO



Reta de regressão

Observamos que para um valor fixado de  $X$ , temos diferentes valores de  $Y$ .  
ou seja, a variabilidade de  $Y$  não é devida somente a  $X$ .

