Regressão Múltipla (Seção 10.5)

Período	CIF	HMOD	HM
1	350	4	10
2	400	8	14
3	470	12	16
4	550	10	26
5	620	15	31
6	380	7	12
7	290	6	13
8	490	10	21
9	580	11	26
10	610	13	24
11	560	12	23
12	420	8	12
13	450	11	19
14	510	12	19
15	380	5	11

O modelo de regressão Múltipla é uma extensão lógica do Modelo de regressão Simples.

Em geral:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n$$

No caso do exemplo da tabela ao lado:

$$C\hat{I}F = b_0 + b_1\ HMOD + b_2\ HM$$

Resolução pelo Excel:

Ferramentas – Análise de Dados – **Correlação**

Ferramentas – Análise de Dados – **Regressão**

Modelo Clássico de Regressão Linear

Desenvolvimento matricial para regressão linear múltipla.

$$\underset{[n\times 1]}{Y} = \underset{[n\times(r+1)]}{X}\ \underset{[(r+1)\times 1]}{\beta} + \underset{[n\times 1]}{\varepsilon}$$

Modelo Clássico de Regressão Linear

O modelo de regressão:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_r x_r + \varepsilon$$

n observações independentes:

$$\begin{gathered}
Y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \ldots + \beta_r x_{1r} + \varepsilon_1 \\
Y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \ldots + \beta_r x_{2r} + \varepsilon_2 \\
\vdots \\
Y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \ldots + \beta_r x_{nr} + \varepsilon_n
\end{gathered}$$

Na forma matricial

$$Y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \ldots + \beta_r x_{1r} + \varepsilon_1$$
$$Y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \ldots + \beta_r x_{2r} + \varepsilon_2$$
$$\vdots$$
$$Y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \ldots + \beta_r x_{nr} + \varepsilon_n$$

$$\Downarrow$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1r} \\ 1 & x_{21} & x_{22} & \ldots & x_{2r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \ldots & x_{nr} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_r \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$\underset{[n\times 1]}{Y} = \underset{[n\times(r+1)]}{X}\;\underset{[(r+1)\times 1]}{\beta} + \underset{[n\times 1]}{\varepsilon}$$

Estimativa de mínimos quadrados LS (*Least Squares*)

Solução:

$$\beta = (Z'Z)^{-1}(Z'Y)$$

$$Y = X\beta + \varepsilon$$

Z' Transposta de Z

Exemplo:

y	x1	x2
6,1	0	2
4,9	1	1
13	2	3
8,9	3	1

Exemplo:

y	x1	x2
6,1	0	2
4,9	1	1
13	2	3
8,9	3	1

Solução: $$\beta = (Z'Z)^{-1}(Z'Y)$$

$$Z = \begin{vmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 1 \end{vmatrix}$$

$$Z' = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 2 & 1 & 3 & 1 \end{vmatrix}$$

$$(Z'Z) = \begin{vmatrix} 4 & 6 & 7 \\ 6 & 14 & 10 \\ 7 & 10 & 15 \end{vmatrix}$$

$$Y = \begin{vmatrix} 6{,}1 \\ 4{,}9 \\ 13 \\ 8{,}9 \end{vmatrix}$$

$$(Z'Z)^{-1} = \begin{vmatrix} 2{,}0370 & -0{,}3704 & -0{,}7037 \\ -0{,}3704 & 0{,}2037 & 0{,}0370 \\ -0{,}7037 & 0{,}0370 & 0{,}3704 \end{vmatrix}$$

$$(Z'Y) = \begin{vmatrix} 32{,}9 \\ 57{,}6 \\ 65{,}0 \end{vmatrix}$$

$$\beta = \begin{vmatrix} -0{,}06 \\ 1{,}96 \\ 3{,}06 \end{vmatrix}$$

Exemplo:

y	x1	x2
6,1	0	2
4,9	1	1
13	2	3
8,9	3	1

$$Y = -0,06 + 1,96\,x_1 + 3,06\,x_2$$

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0,9999
R Square	0,9998
Adjusted R Square	0,9995
Standard Error	0,0816
Observations	4

Solução no Excel:

Solução:

$$\beta = \begin{vmatrix} -0,06 \\ 1,96 \\ 3,06 \end{vmatrix}$$

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	38,821	19,410	2911,56	0,01
Residual	1	0,007	0,007		
Total	3	38,8275			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-0,06	0,12	-0,48	0,72	-1,54	1,43
x1	1,96	0,04	53,07	0,01	1,49	2,42
x2	3,06	0,05	61,49	0,01	2,42	3,69

Regressão Múltipla (Seção 10.5)

RESUMO DOS RESULTADOS

$$\hat{CIF} = b_0 + b_1\,HMOD + b_2\,HM$$

Estatística de regressão	
R múltiplo	0,941
R-Quadrado	0,885
R-quadrado ajustado	0,866
Erro padrão	36,827
Observações	15

$$\hat{CIF} = 184{,}884 + 11{,}746\,HMOD + 9{,}369\,HM$$

ANOVA

	gl	*SQ*	*MQ*	*F*	*F de sig.*
Regressão	2	125218,94	62609,47	46,165	0,000
Resíduo	12	16274,39	1356,20		
Total	14	141493,33			

	Coeficientes	*Erro padrão*	*Stat t*	*valor-P*	*95% inf.*	*95% sup.*
Interseção	184,884	31,762	5,821	0,000	115,68	254,09
HMOD	11,746	5,835	2,013	0,067	-0,97	24,46
HM	9,369	2,825	3,317	0,006	3,21	15,52

Regressão Múltipla (Seção 10.5)

Exemplo do livro FIPECAFI, 2007, p161. Utilização do Software SPSS

Base de dados: Exame – 500 Melhores & Maiores, ano 2001

Variáveis - Indicadores Financeiros:

RENTAT:	Rentabilidade do Ativo
RENTPL:	Rentabilidade do Patrimônio Líquido
ALOPER:	Alavancagem Operacional
MARVEN:	Margem de Vendas
ALFIN:	Alavancagem Financeira
LUPRE:	Lucro ou Prejuízo

Variável dependente **RENTPL**

Objetivo: Descobrir as relações mais relevantes através de correlações e regressões (pois graficamente fica impraticável)

No SPSS: **Correlação**: Correlate – Bivariate – Pearson
Correlação parcial: Correlate – Partial
Regressão: Regression – Linear (Method: stepwise)