

## Tutorial Prático II

### Correlação e Regressão

Felipe Figueiredo

Instituto Nacional de Traumatologia e Ortopedia

## Sumário

## Dados

- Usaremos os dados do exemplo da aula de correlação (Motulsky, 1995)
- Objetivo: refazer as análises do autor no BioEstat

## Dados

**Table 17.1.** Correlation Between %C20–22 and Insulin Sensitivity

% C20–22 Polyunsaturated Fatty Acids	Insulin Sensitivity (mg/m <sup>2</sup> /min)
17.9	250
18.3	220
18.3	145
18.4	115
18.4	230
20.2	200
20.3	330
21.8	400
21.9	370
22.1	260
23.1	270
24.2	530
24.4	375

## Dados



BioEstat 5.3

Arquivo Editar Estatísticas Gráficos Sugestões Configurar Ajuda

Escolha um teste

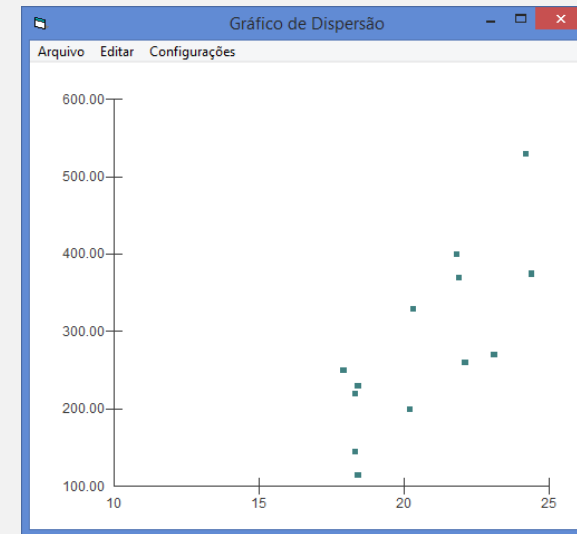
Dados 1

17.9

	- 1 -	- 2 -	- 3 -	- 4 -	- 5 -	- 6 -	- 7 -	- 8 -
1	17.900	250.000						
2	18.300	220.000						
3	18.300	145.000						
4	18.400	115.000						
5	18.400	230.000						
6	20.200	200.000						
7	20.300	330.000						
8	21.800	400.000						
9	21.900	370.000						
10	22.100	260.000						
11	23.100	270.000						
12	24.200	530.000						
13	24.400	375.000						
14								
15								
16								
17								
18								

Tutorial  
Prático II  
Felipe  
Figueiredo

## Gráfico de Dispersão



Tutorial  
Prático II  
Felipe  
Figueiredo

## Sugestão do BioEstat



Testes de hipótese - Sugestões

AMOSTRAS	VARIÁVEIS	DADOS	TESTE SUGERIDO
1 amostra			
2 amostras independentes			
2 amostras pareadas			
k amostras independentes			
k amostras relacionadas			

Sair

Tutorial  
Prático II  
Felipe  
Figueiredo

## Sugestão do BioEstat



Testes de hipótese - Sugestões

AMOSTRAS	VARIÁVEIS	DADOS	TESTE SUGERIDO
1 amostra	1 variável		
	2 variáveis		
	k variáveis		
<< Amostras			

Sair

Tutorial  
Prático II  
Felipe  
Figueiredo

Testes de hipótese - Sugestões

AMOSTRAS	VARIÁVEIS	DADOS	TESTE SUGERIDO
1 amostra	2 variáveis	Catégoricos Ordenativos Numéricos << Variáveis	Sair

Testes de hipótese - Sugestões

AMOSTRAS	VARIÁVEIS	DADOS	TESTE SUGERIDO
1 amostra	2 variáveis	Numéricos	Ajustamento de Curvas Correlação Linear (Pearson) Regressão Linear << Dados Sair

\* CORRELAÇÃO LINEAR DE PEARSON \*

Teste estatístico cuja finalidade é verificar a direção e o grau de associação entre duas variáveis (X e Y) de forma linear.  
Ex.: comprimento (X) e largura (Y) de folhas de uma árvore.

$N = \text{número de pares}$

$$r = \frac{\sum xy}{\sqrt{\sum x^2} \sqrt{\sum y^2}}$$

$$gl = N - 2$$

$$t = r\sqrt{N-2} / \sqrt{1-r^2}$$

BioEstat

Arquivo Editar Estatísticas Gráficos Sugestões Configurar Ajuda

Teste de Correlação Linear Ctrl+U Escolha um teste

- Amostragem
- Análise Multivariada
- Análise de Sobrevida
- Análise da Variância
- Bootstrap - Reamostragem
- Correlação
  - Coef. de Contingência C
  - Coef. de Concordância de Kendall
  - Coef. de Correlação de Kendall
  - Coef. de Correlação Parcial
  - Coef. de Correlação de Pearson
  - Coef. de Correlação:  $\mu$  diferente de 0
  - Coef. Phi
  - Coef. de Correlação de Spearman
  - Matriz de Correlação
- Estatísticas Circulares
- Distribuição de Probabilidades
- Uma Amostra
- Duas Amostras Independentes
- Duas Amostras Relacionadas
- Estatística Descritiva
- Estimação de Parâmetros
- Meta-Análise
- Normalidade
- Qui-Quadrado
- Regressão
- Tamanho das Amostras
- Transformação e Ordenação

Teste de Correlação Linear

Arquivo Editar Gráfico

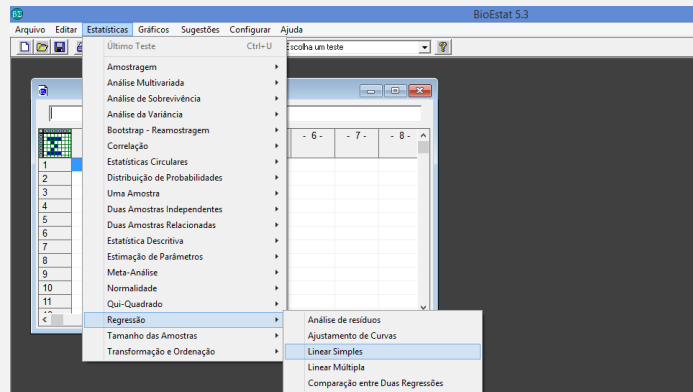
	Colunas 2 e 1
n (pares) =	13
r (Pearson) =	0.7700
IC 95% =	0.38 a 0.93
IC 99% =	0.20 a 0.95
R2 =	0.5929
t =	4.0026
GL =	11
(p) =	0.0021
Poder 0.05 =	0.9431
Poder 0.01 =	0.8161

# Regressão



Tutorial  
Prático II

Felipe  
Figueiredo



# Regressão



Tutorial  
Prático II

Felipe  
Figueiredo

Fontes de variação	GL	SQ	QM
Regressão	1	92280.9337	92280.9337
Erro	11	63361.3740	5760.1249
Total	12	155642.3077	---
F (regressão) =	16.0206	p = 0.0024	
Variável dependente =	Coluna 2		
Variável independente =	Coluna 1		
Média (X) =	20.7154		
Média (Y) =	284.2308		
Coef. de Determinação (R2) =	0.5929		
R2 (ajustado) =	0.5559		
Coefficiente de Correlação =	0.7700		
Intercepto (a) =	-486.5420	t = -2.5116	p = 0.0288
Coef. de Regressão (b) =	37.2077	t = 4.0026	p = 0.0021
IC 95% (a)	-912.911 a -60.173		
IC 95% (b)	16.747 a 57.668		
Equação	$Y' = a + bX$		

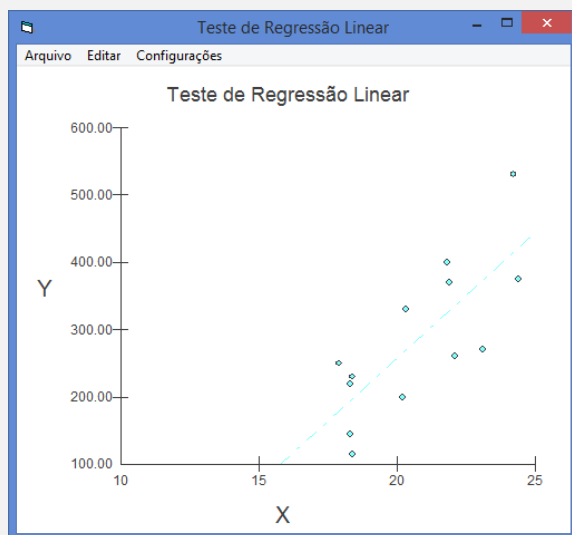
Selecionar “Gráfico” para visualizar a reta regressora

# Regressão



Tutorial  
Prático II

Felipe  
Figueiredo



# Interpolação



Tutorial  
Prático II

Felipe  
Figueiredo

- Agora que temos o modelo, podemos fazer estimativas para regiões onde não há dados!
- Basta selecionar “Estimar Y” e estipular um valor de X

# Regressão



Tutorial  
Prático II

Felipe  
Figueiredo

Teste de Regressão Linear			
Fontes de variação	GL	SQ	QM
Regressão	1	92280.9337	92280.9337
Erro	11	63361.3740	5760.1249
Total	12	155642.3077	---
F (regressão) =	16.0206	p = 0.0024	
Variável dependente =	Coluna 2		
Variável independente =	Coluna 1		
Média (X) =	20.7154		
Média (Y) =	284.2308		
Coef. de Determinação (R2) =	0.5929		
R2 (ajustado) =	0.5559		
Coefficiente de Correlação =	0.7700		
Intercepto (a) =	-486.5420	t = -2.5116	p = 0.0288
Coef. de Regressão (b) =	37.2077	t = 4.0026	p = 0.0021
IC 95% (a)	-912.911 a -60.173		
IC 95% (b)	16.747 a 57.668		
Equação	$Y' = a + bX$		

Selecionar “Estimar Y” e estipular um valor de X

# Dados



Tutorial  
Prático II

Felipe  
Figueiredo

**Table 17.1.** Correlation Between %C20–22 and Insulin Sensitivity

% C20–22 Polyunsaturated Fatty Acids	Insulin Sensitivity (mg/m <sup>2</sup> /min)
17.9	250
18.3	220
18.3	145
18.4	115
18.4	230
20.2	200
20.3	330
21.8	400
21.9	370
22.1	260
23.1	270
24.2	530
24.4	375

Qual seria o índice de insulina esperado, se  $X = 19.5$ ?

# Interpolação



Tutorial  
Prático II

Felipe  
Figueiredo

Cálculo da Estimação de Y

	Digite
X 1	19.5

$Y' = -486.5420 + (37.2077 \times 19.5000)$

---

$Y' = 239.0082$

Estimar Y

Cancela