

Tutorial Prático II

Correlação e Regressão

Felipe Figueiredo

Instituto Nacional de Traumatologia e Ortopedia

1 Associação

- Dados
- Gráfico de Dispersão
- Sugestão
- Correlação
- Regressão

1

Associação

- Dados
- Gráfico de Dispersão
- Sugestão
- Correlação
- Regressão

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

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

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

% C20–22 Polyunsaturated Fatty Acids	Insulin Sensitivity (mg/m ² /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



Tutorial Prático II

Felipe
Figueiredo

Associação

Dados

Gráfico de Dispersão

Sugestão

Correlação

Regressão

BioEstat 5.3

Arquivo Editar Estatísticas Gráficos Sugestões Configurar Ajuda

Escolha um teste

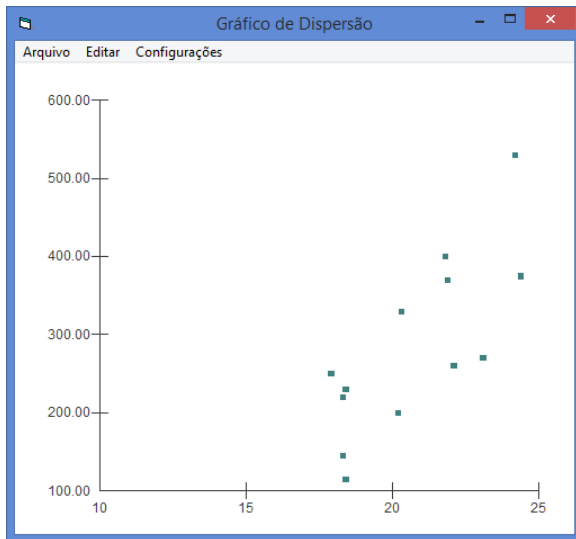
Dados 1

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								

- 1 Associação
 - Dados
 - Gráfico de Dispersão
 - Sugestão
 - Correlação
 - Regressão

Gráfico de Dispersão



1

Associação

- Dados
- Gráfico de Dispersão
- **Sugestão**
- Correlação
- Regressão

Sugestão do BioEstat

Associação

Dados

Gráfico de Dispersão

Sugestão

Correlação

Regressão

 **Testes de hipótese - Sugestões** ✕

AMOSTRAS	VARIÁVEIS	DADOS	TESTE SUGERIDO
1 amostra			<div>Sair</div>
2 amostras independentes			
2 amostras pareadas			
k amostras independentes			
k amostras relacionadas			

Testes de hipótese - Sugestões

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

Associação

Dados

Gráfico de Dispersão

Sugestão



Correlação

Regressão

Sugestão do BioEstat



The screenshot shows a software window titled "Testes de hipótese - Sugestões" with a red close button in the top right corner. The window is divided into four main panels. The top-left panel, labeled "AMOSTRAS", contains a yellow button labeled "1 amostra". The top-middle panel, labeled "VARIÁVEIS", contains a yellow button labeled "2 variáveis". The top-right panel, labeled "DADOS", contains three buttons: "Categóricos", "Ordenativos", and "Numéricos", along with a button labeled "<< Variáveis". The bottom-right panel, labeled "TESTE SUGERIDO", contains a button labeled "Sair". Below these four panels are two large, empty white rectangular areas.

 Testes de hipótese - Sugestões 

AMOSTRAS	VARIÁVEIS	DADOS	TESTE SUGERIDO
1 amostra			Ajustamento de Curvas
	2 variáveis		Correlação Linear (Pearson)
		Numéricos	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.

$$N = \text{número de pares}$$
$$r = \frac{\sum xy}{\sqrt{\sum x^2} \sqrt{\sum y^2}}$$
$$gl = N - 2$$
$$t = r\sqrt{N-2} / \sqrt{1-r^2}$$

1 Associação

- Dados
- Gráfico de Dispersão
- Sugestão
- **Correlação**
- Regressão

Correlação

Associação

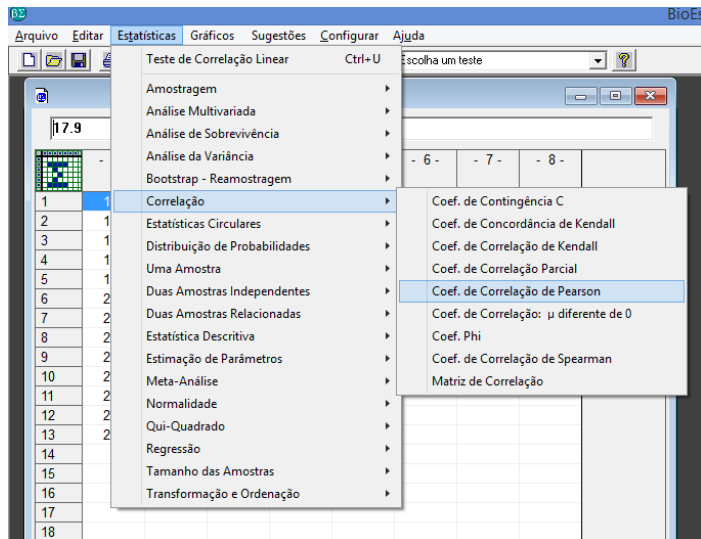
Dados

Gráfico de Dispersão

Sugestão

Correlação

Regressão



The screenshot shows the BioEs software interface. The 'Estatísticas' menu is open, and the 'Correlação' option is selected. A sub-menu is displayed, listing the following correlation coefficients:

- 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: μ diferente de 0
- Coef. Phi
- Coef. de Correlação de Spearman
- Matriz de Correlação

The background shows a data table with columns 1 through 18. The first column contains values 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18. The second column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The third column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The fourth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The fifth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The sixth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The seventh column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The eighth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The ninth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The tenth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The eleventh column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The twelfth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The thirteenth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The fourteenth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The fifteenth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The sixteenth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The seventeenth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2. The eighteenth column contains values 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2.

Correlação



Tutorial
Prático II

Felipe
Figueiredo

Associação

Dados

Gráfico de Dispersão

Sugestão

Correlação

Regressão

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

1

Associação

- Dados
- Gráfico de Dispersão
- Sugestão
- Correlação
- **Regressão**

Regressão

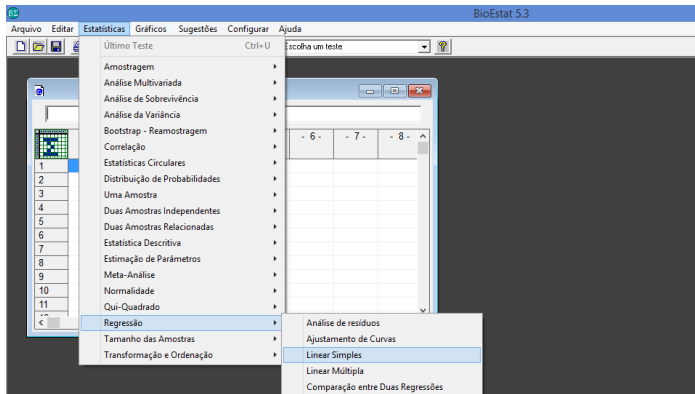


Tutorial Prático II

Felipe
Figueiredo

Associação

Dados
Gráfico de Dispersão
Sugestão
Correlação
Regressão



Regressão



Tutorial
Prático II

Felipe
Figueiredo

Associação

Dados

Gráfico de Dispersão

Sugestão

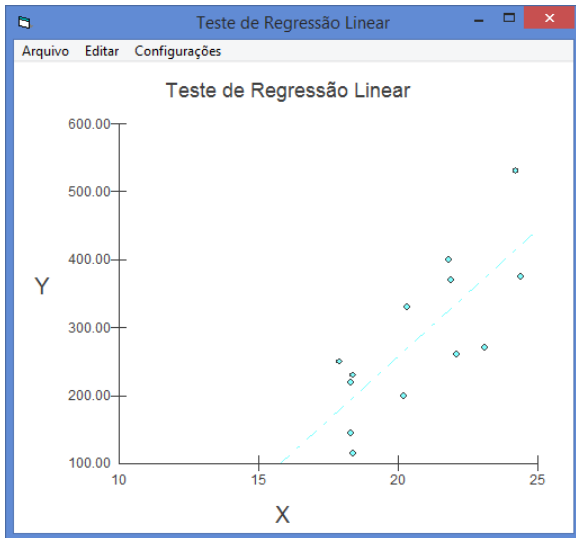
Correlação

Regressão

Teste de Regressão Linear			
Arquivo Editar Gráfico Estimar Y			
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



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

- 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

Associação

Dados

Gráfico de Dispersão

Sugestão

Correlação

Regressão

Teste de Regressão Linear			
Arquivo Editar Gráfico Estimar Y			
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

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

% C20–22 Polyunsaturated Fatty Acids	Insulin Sensitivity (mg/m ² /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 $X = 19.5$?

Cálculo da Estimação de Y

	Digite
X 1	19.5

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

$$Y' = 239.0082$$

Estimar Y

Cancela