

Análise de Regressão - Aula Prática 1

R, SAS, Minitab e SPSS

Danilo Silva
Cibele Russo
Mariana Cúri

19 de Março de 2013

- 1 Introdução
- 2 Ajuste em R
- 3 Ajuste em SAS
- 4 Ajuste em Minitab
- 5 Ajuste em SPSS

Sumário

- 1 Introdução
- 2 Ajuste em R
- 3 Ajuste em SAS
- 4 Ajuste em Minitab
- 5 Ajuste em SPSS

Conjunto de Dados

Um pesquisador acredita que o tempo de duração (Y) de um determinado anestésico está associado à idade (X) dos indivíduos. Para investigar tal desconfinança, o pesquisador coletou dados em 20 indivíduos e obteve os seguintes resultados:

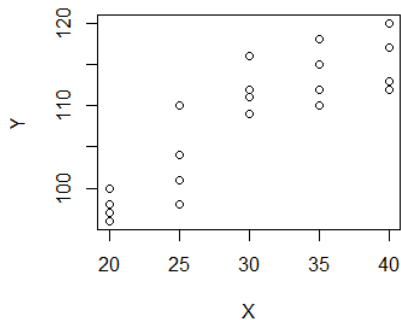
i	Y	X
1	96	20
2	97	20
3	98	20
4	100	20
5	98	25
6	104	25
7	110	25
8	101	25
9	116	30
10	112	30
11	109	30
12	111	30
13	112	35
14	110	35
15	118	35
16	115	35
17	113	40
18	112	40
19	120	40
20	117	40

Sumário

- 1 Introdução
- 2 Ajuste em R**
- 3 Ajuste em SAS
- 4 Ajuste em Minitab
- 5 Ajuste em SPSS

Gráfico de Dispersão

```
> dados <- read.table("http://wiki.icmc.usp.br/images/7/7d/AulaPratica1DadosAnestesico.txt", header=TRUE)
> attach(dados)
> plot(X,Y)
```



Ajuste do Modelo

```
> fit <- lm(Y ~ X)
> summary(fit)
```

```
Call:
lm(formula = Y ~ X)

Residuals:
    Min       1Q   Median       3Q      Max
-5.85  -2.90  -0.25   2.40   7.55

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  80.8500     3.7927   21.317 3.19e-14 ***
X              0.9200     0.1231    7.477 6.34e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.891 on 18 degrees of freedom
Multiple R-squared:  0.7564, Adjusted R-squared:  0.7429
F-statistic: 55.9 on 1 and 18 DF, p-value: 6.337e-07
```

ANOVA

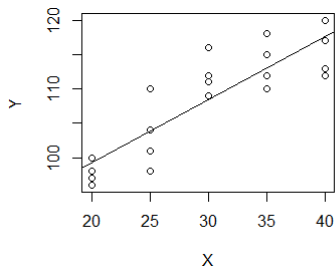
```
> anova(fit)
> plot(X,Y)
> abline(fit)
```

Analysis of Variance Table

Response: Y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
x	1	846.40	846.40	55.899	6.337e-07 ***
Residuals	18	272.55	15.14		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1



Sumário

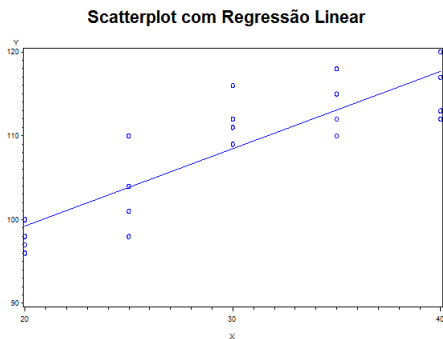
- 1 Introdução
- 2 Ajuste em R
- 3 Ajuste em SAS**
- 4 Ajuste em Minitab
- 5 Ajuste em SPSS

Leitura dos Dados

```
proc import datafile = 'C:\Monitoria\AulaPratica1DadosAnestesico.xls'  
    dbms = excel replace  
    out = pratical;  
    sheet = 'Plan1';  
    getnames = yes;  
run;  
  
proc insight data = pratical;  
run;
```

Gráfico de Dispersão

```
SYMBOL1 V=circle C=blue I=r;  
title 'Scatterplot com Regressão Linear';  
  
proc gplot data = pratica1;  
    plot Y * X ;  
run;
```



Ajuste do Modelo e ANOVA

```
proc reg data = pratica1;
    model Y = X ;
run;
```

The REG Procedure
 Model: MODEL1
 Dependent Variable: Y Y

Number of Observations Read 20
 Number of Observations Used 20

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	846.40000	846.40000	55.90	<.0001
Error	18	272.55000	15.14167		
Corrected Total	19	1118.95000			

Root MSE 3.89123 R-Square 0.7564
 Dependent Mean 108.45000 Adj R-Sq 0.7429
 Coeff Var 3.58804

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	80.85000	3.79270	21.32	<.0001
X	X	1	0.92000	0.12305	7.48	<.0001

Sumário

- 1 Introdução
- 2 Ajuste em R
- 3 Ajuste em SAS
- 4 Ajuste em Minitab**
- 5 Ajuste em SPSS

Gráfico de Dispersão

The screenshot shows the Minitab software interface. The 'Graph' menu is open, and 'Scatterplot...' is selected. The background worksheet contains the following data:

	C1	C2
	Y	X
1	96	20
2	97	20
3	98	20
4	100	20
5	98	25
6	104	25
7	110	25
8	101	25
9	116	30
10	112	30
11	109	30
12	111	30

Gráfico de Dispersão

Minitab - Untitled

File Edit Data Calc Stat Graph Editor Tools Window Help Assistant

Session

Welcome to Minitab, press F1 for help.

Worksheet1 ***

	C1	C2	C3	C4	C5	C6
	Y	X				
1	96	20				
2	97	20				
3	98	20				
4	100	20				
5	98	25				
6	104	25				
7	110	25				
8	101	25				
9	116	30				
10	112	30				
11	109	30				
12	111	30				
13	112	35				
14	110	35				

Scatterplots

Simple With Groups With Regression With Regression and Groups

With Connect Line With Connect and Groups

Help OK Cancel

Gráfico de Dispersão

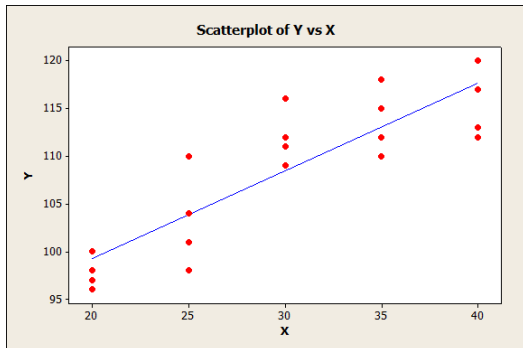
The screenshot shows the Minitab interface with a 'Scatterplot - With Regression' dialog box open. The background worksheet contains the following data:

	C1	C2	C3	C4	C5
	Y	X			
1	96	20			
2	97	20			
3	98	20			
4	100	20			
5	98	25			
6	104	25			
7	110	25			
8	101	25			
9	116	30			
10	112	30			
11	109	30			
12	111	30			
13	112	35			
14	110	35			
15	110	35			

The 'Scatterplot - With Regression' dialog box is configured as follows:

- Y variables:** C1 Y
- X variables:** C2 X
- Buttons:** Scale..., Labels..., Data View..., Multiple Graphs..., Data Options..., Select, Help, OK (circled in red), Cancel.

Gráfico de Dispersão



Ajuste do Modelo

Minitab - Untitled

File Edit Data Calc Stat Graph Editor Tools Window Help Assistant

Basic Statistics
Regression
ANOVA
DOE
Control Charts
Quality Tools
Reliability/Survival
Multivariate
Time Series
Tables
Nonparametrics
EDA
Power and Sample Size

Regression...
General Regression...
Stepwise...
Best Subsets...
Fitted Line Plot...
Nonlinear Regression...
Orthogonal Regression...
Partial Least Squares...
Binary Logistic Regression...
Ordinal Logistic Regression...
Nominal Logistic Regression...

Session

Welcome to Minitab,

Worksheet 1 ***

	C1	C2
	Y	
1	96	
2	97	
3	98	20
4	100	20
5	98	25
6	104	25
7	110	25

Ajuste do Modelo

Minitab - Untitled

File Edit Data Calc Stat Graph Editor Tools Window Help Assistant

Session

Welcome to Minitab, press F1 for help.

Worksheet 1 ***

	C1	C2	C3	C4	C5	C6
	Y	X				
1	96	20				
2	97	20				
3	98	20				
4	100	20				
5	98	25				
6	104	25				
7	110	25				
8	101	25				
9	116	30				
10	112	30				
11	109	30				
12	111	30				
13	112	35				
14	110	35				
15	118	35				
16	115	35				

Regression

C1 Y
C2 X

Response: Y

Predictors: X

Select

Help

Graphs... Options...
Results... Storage...
OK Cancel

Ajuste do Modelo

Minitab - Untitled

File Edit Data Calc Stat Graph Editor Tools Window Help Assistant

Session

Regression Analysis: Y versus X

The regression equation is
 $Y = 80,9 + 0,920 X$

Predictor	Coef	SE Coef	T	P
Constant	80,850	3,793	21,32	0,000
X	0,9200	0,1231	7,48	0,000

S = 3,89123 R-Sq = 75,6% R-Sq(adj) = 74,3%

Worksheet1 ***

↓	C1	C2	C3	C4	C5	C6	C7
	Y	X					
1	96	20					
2	97	20					
3	98	20					
4	100	20					
5	98	25					

ANOVA

The screenshot displays the Minitab interface. The top window, titled 'Session', shows the results of an Analysis of Variance (ANOVA). Below this, the 'Worksheet1 ***' window shows a data table with columns C1 through C7 and rows 1 through 5.

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	846,40	846,40	55,90	0,000
Residual Error	18	272,55	15,14		
Total	19	1118,95			

	C1	C2	C3	C4	C5	C6	C7
	Y	X					
1	96	20					
2	97	20					
3	98	20					
4	100	20					
5	98	25					

Sumário

- 1 Introdução
- 2 Ajuste em R
- 3 Ajuste em SAS
- 4 Ajuste em Minitab
- 5 Ajuste em SPSS**

Configurações iniciais

SPSS Statistics Data Editor - Variable View

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	Y	Numeric	8	2		None	None	8	Right	Scale
2	X	Numeric	8	2		None	None	8	Right	Scale
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										

Data View Variable View

Gráfico de Dispersão

The screenshot shows the SPSS Data Editor interface. The main window displays a data table with two columns, Y and X, and 26 rows of data. The 'Graphs' menu is open, and the 'Scatter Plot' option is highlighted with a red circle. The data table is as follows:

	Y	X
1	96,00	20,00
2	97,00	20,00
3	98,00	20,00
4	100,00	20,00
5	98,00	25,00
6	104,00	25,00
7	110,00	25,00
8	101,00	25,00
9	116,00	30,00
10	112,00	30,00
11	109,00	30,00
12	111,00	30,00
13	112,00	35,00
14	110,00	35,00
15	118,00	35,00
16	115,00	35,00
17	113,00	40,00
18	112,00	40,00
19	120,00	40,00
20	117,00	40,00
21		
22		
23		
24		
25		
26		

Gráfico de Dispersão

SPSS Data Editor window showing a dataset with columns Y and X. The 'Scatter/Dot' dialog box is open, with 'Simple Scatter' selected and the 'Define' button highlighted.

	Y	X	var	var	var	var	var	var	var	var
1	96,00	20,00								
2	97,00	20,00								
3	98,00	20,00								
4	100,00	20,00								
5	98,00	25,00								
6	104,00	25,00								
7	110,00	25,00								
8	101,00	25,00								
9	116,00	30,00								
10	112,00	30,00								
11	109,00	30,00								
12	111,00	30,00								
13	112,00	35,00								
14	110,00	35,00								
15	118,00	35,00								
16	115,00	35,00								
17	113,00	40,00								
18	112,00	40,00								

Gráfico de Dispersão

*Untitled1 [DataSet0] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1:

	Y	X	var	var
1	96,00	20,00		
2	97,00	20,00		
3	98,00	20,00		
4	100,00	20,00		
5	98,00	25,00		
6	104,00	25,00		
7	110,00	25,00		
8	101,00	25,00		
9	116,00	30,00		
10	112,00	30,00		
11	109,00	30,00		
12	111,00	30,00		
13	112,00	35,00		
14	110,00	35,00		
15	118,00	35,00		
16	115,00	35,00		
17	113,00	40,00		
18	112,00	40,00		
19	120,00	40,00		
20	117,00	40,00		
21				
22				
23				
24				
25				

Simple Scatterplot

Y Axis: Y

X Axis: X

Set Markers by:

Label Cases by:

Panel by

Rows:

Columns:

Template

Use chart specifications from:

File...

OK Paste Reset Cancel Help

Gráfico de Dispersão

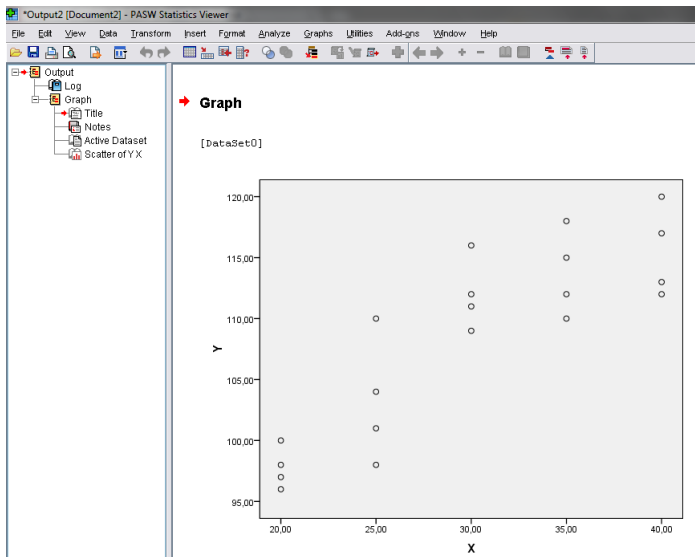


Gráfico de Dispersão

Duplo clique em cima do gráfico.

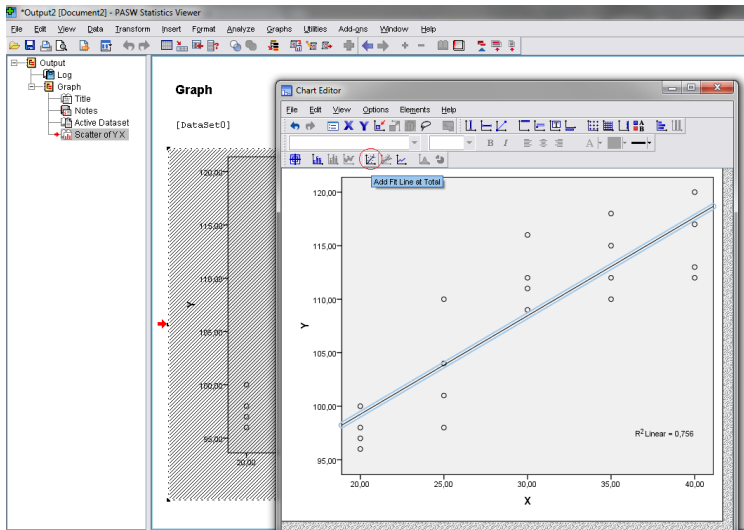
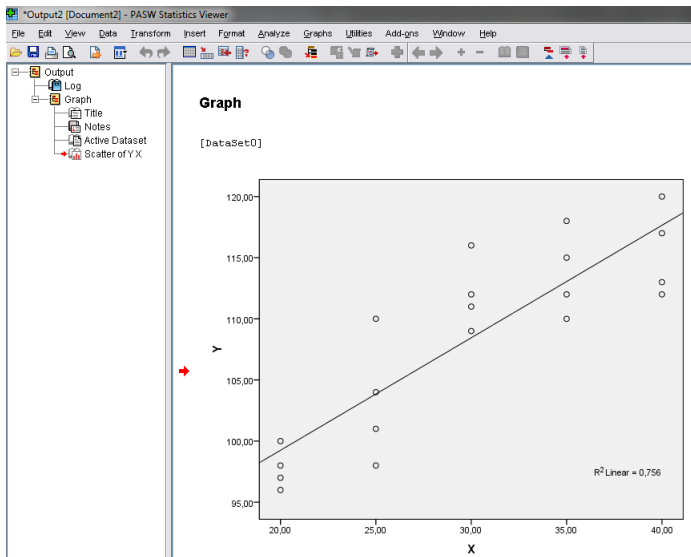
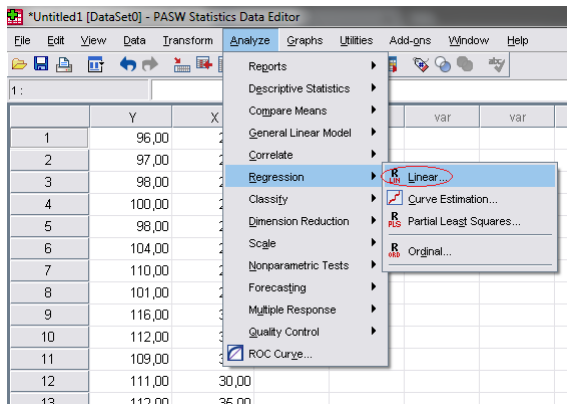


Gráfico de Dispersão



Ajuste do Modelo e ANOVA



The screenshot shows the SPSS interface with the 'Analyze' menu open. The 'Regression' option is selected, and its sub-menu is displayed, with 'Linear...' highlighted. The background shows a data editor window with columns Y and X, and a list of data points.

	Y	X
1	96,00	
2	97,00	
3	98,00	
4	100,00	
5	98,00	
6	104,00	
7	110,00	
8	101,00	
9	116,00	
10	112,00	
11	109,00	
12	111,00	30,00
13	112,00	35,00

Ajuste do Modelo e ANOVA

SPSS Statistics Data Editor window showing a dataset with 22 rows and 12 columns. The first two columns are labeled 'Y' and 'X'. The data is as follows:

	Y	X	var	var	var	var	var	var	var	var	var
1	96,00	20,00									
2	97,00	20,00									
3	98,00	20,00									
4	100,00	20,00									
5	98,00	25,00									
6	104,00	25,00									
7	110,00	25,00									
8	101,00	25,00									
9	116,00	30,00									
10	112,00	30,00									
11	109,00	30,00									
12	111,00	30,00									
13	112,00	35,00									
14	110,00	35,00									
15	118,00	35,00									
16	115,00	35,00									
17	113,00	40,00									
18	112,00	40,00									
19	120,00	40,00									
20	117,00	40,00									
21											
22											

The Linear Regression dialog box is open, showing the following configuration:

- Dependent: Y
- Independent(s): X
- Method: Enter
- Selection Variable: (empty)
- Case Labels: (empty)
- WLS Weight: (empty)

The OK button is highlighted with a red circle.

Ajuste do Modelo e ANOVA

SPSS Statistics Viewer - *Output3 [Document3] - PASW Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help

Output Log Regression Title Notes Variables Entered/Removed Model Summary ANOVA Coefficients

Regression

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	X ^a	.	Enter

a. All requested variables entered.
b. Dependent Variable: Y

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,870 ^a	,756	,743	3,89123

a. Predictors: (Constant), X

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	846,400	1	846,400	55,899	,000 ^a
	Residual	272,550	18	15,142		
	Total	1118,950	19			

a. Predictors: (Constant), X
b. Dependent Variable: Y

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	80,850	3,793		21,317	,000
	X	,920	,123	,870	7,477	,000

a. Dependent Variable: Y