Gabarito – Prova P2 – GET 170

Gilberto Pereira Sassi

7 de janeiro de 2017

Exercício 1

Gráfico de dispersão (1,5 pontos)

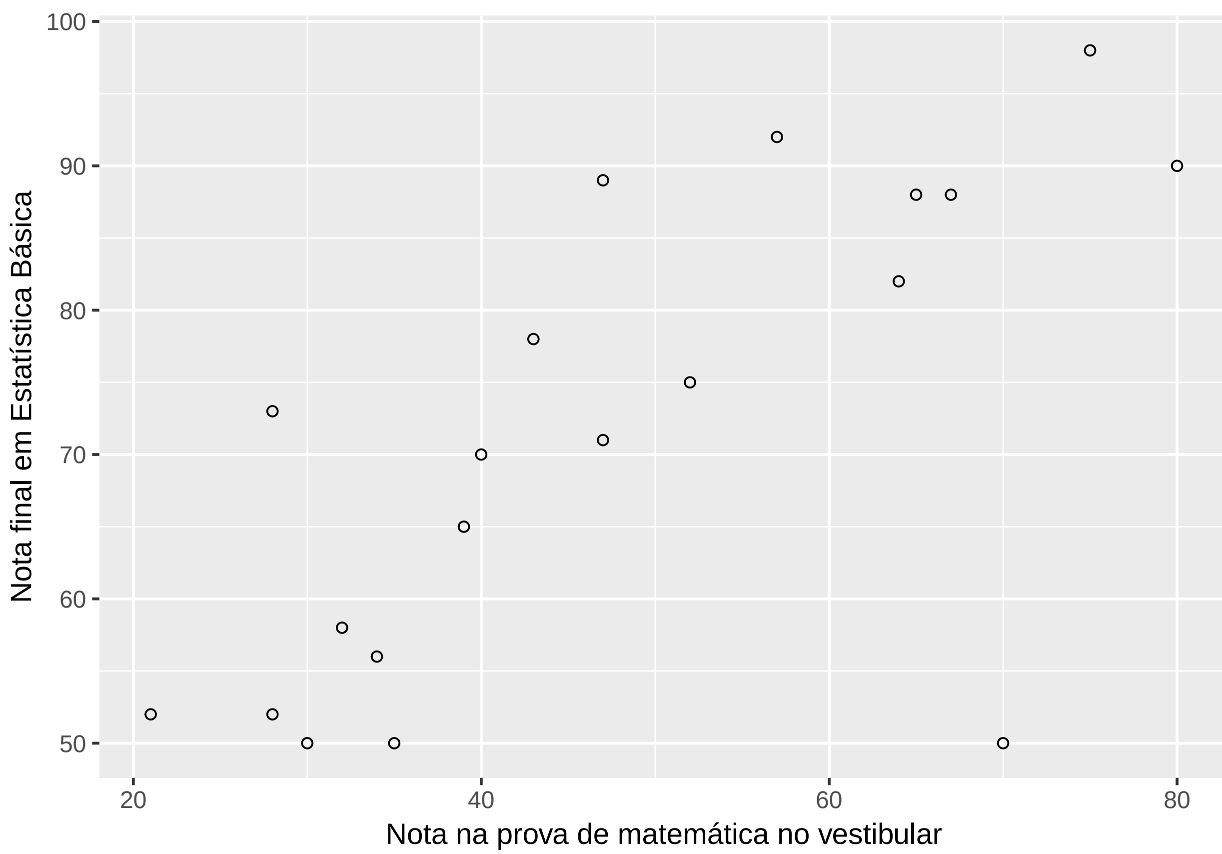

Figura 1: Gráfico de Dispersão.

Comentário: Na figura 1, notamos que quando maior a nota na prova de matemática no vestibular maior a nota final em Estatística Básica, apontando uma associação positiva entre as duas variáveis.

Coeficiente de correlação de Pearson (1,5 pontos)

$$\begin{aligned} corr(x,y) &= \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{(\sum x^2 - n\bar{x}^2)(\sum y^2 - n\bar{y}^2)}} \\ &= \frac{71869 - 20 \cdot 95,5 \cdot 14,27}{\sqrt{(51390 - 20 \cdot 95,4^2)(10693 - 20 \cdot 14,27^2)}} \\ &= 0,69 \end{aligned}$$

Como o valor do coeficiente de correlação linear de Pearson está próximo de 1, podemos dizer que as variáveis estão positivamente associadas.

Regressão linear simples (1,5 pontos)

Y = a + b X

$$\begin{aligned} b &= \frac{n \cdot \sum xy - \left(\sum x\right)\left(\sum y\right)}{n\sum x^2 - \left(\sum x\right)^2} \\ &= \frac{20 \cdot 71689 - 954 \cdot 1427}{20 \cdot 51390 - 954^2} \\ &= 0,65 \\ a &= \frac{\sum y - b \cdot \sum x}{n} \\ &= \frac{1427 - 0,65 \cdot 954}{20} \\ &= 40,53 \end{aligned}$$

Interpretação: Podemos estimar a nota final de um aluno de Estatística Básica (y) a partir da nota de matemática do vestibular (x) atráves da equação $y = 40,53 + 0.65 \cdot x$. Por exemplo, um aluno com nota 60 terá aproximadamente uma nota final $79,53$ no curso de Estatística Básica.

Nota final de um aluno de Estatística Básica com nota 60 no prova de matemática do vestibular

$$\begin{aligned} \hat{y} &= 40,53 + 0,65 \cdot 60 \\ &= 79,53 \end{aligned}$$

Exercício 2

Gráfico de dispersão (1,5 pontos)

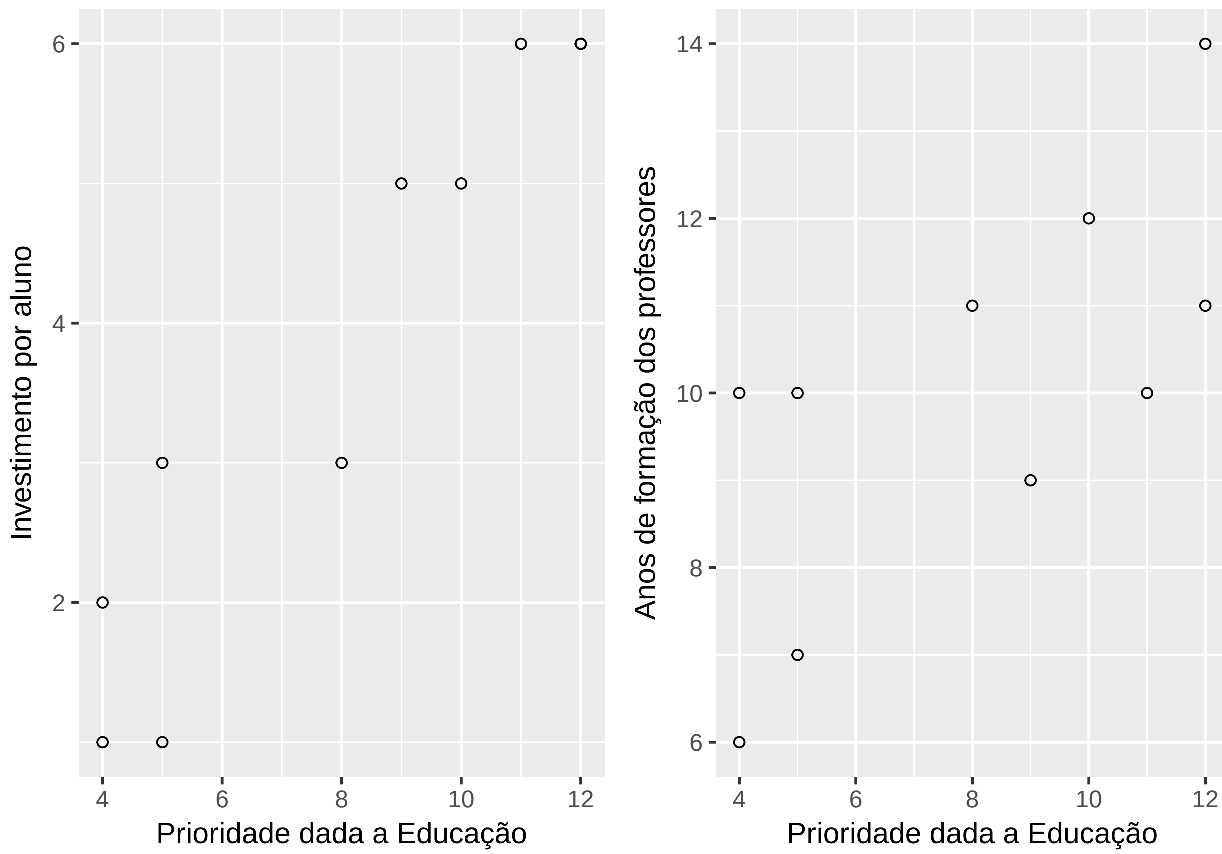

Figura 2: Gráficos de dispersão X e Y (figura A) e entre X e Z (figura B).

Na figura 2, notamos que x e y e x e z estão positivamente associadas. Ou seja, se o investimento por aluno ou tempo de formação dos professores aumenta a prioridade dada a educação aumenta.

Coeficiente de correlação linear de Pearson (1,5 pontos)

$$
\begin{aligned}
corr(x,y) &= \frac{\sum xy - n\cdot\bar{x}\cdot\bar{y}}{\sqrt{(\sum x^2 - n\cdot\bar{x}^2)\cdot(\sum y^2 - n\cdot\bar{y}^2)}}\\
&= \frac{361 - 10\cdot 8\cdot 3,8}{\sqrt{(736 - 10\cdot 8^2)\cdot(182 - 10\cdot 3,8^2)}}\\
&= 0,95\\
corr(x,z) &= \frac{\sum xz - n\cdot\bar{x}\cdot\bar{z}}{\sqrt{(\sum x^2 - n\cdot\bar{x}^2)\cdot(\sum z^2 - n\cdot\bar{z}^2)}}\\
&= \frac{848 - 10\cdot 8\cdot 10}{\sqrt{(736 - 10\cdot 8^2)\cdot(1048 - 10\cdot 10^2)}}\\
&= 0,71.
\end{aligned}
$$

Podemos afirmar que X e Y e X e Z estão positivamente associadas, pois as duas variáveis tem coeficiente de correlação de Pearson perto de 1. Contudo, o coeficiente de correlação de Pearson entre X e Y é maior e usaremos Y para estimar X através da regressão linear simples.

Regressão linear simples (1,5 pontos)

X = a + b Y

$$
\begin{aligned}
b &= \frac{n\cdot\sum xy - (\sum x)(\sum y)}{n\sum y^2 - (\sum y)^2}\\
&= \frac{10\cdot 361 - 80\cdot 30}{10\cdot 182 - 38^2}\\
&= 1,52\\
a &= \frac{\sum x - b\cdot\sum y}{n}\\
&= \frac{80 - 1,52\cdot 30}{10}\\
&= 2,24
\end{aligned}
$$

Interpretação Podemos estimar a prioridade dada a educação (x) usando a variável investimento por aluno (y) através da equação $X = 2,24 + 1,52\cdot Y$. Por exemplo, um município que investe 4 salários mínimos por aluno tem prioridade dada a Educação por $\hat{X} = 2,24 + 1,52\cdot 4 = 8,32$.

x	y	$\hat{x} = 2,24 + 1,52 \cdot y$	$(x - \hat{x})^2$	$(\hat{x} - \bar{x})^2$	$(x - \bar{x})^2$
4.00	1.00	3.76	0.06	18.02	16.00
5.00	1.00	3.76	1.55	18.02	9.00
4.00	2.00	5.27	1.62	7.45	16.00
5.00	3.00	6.79	3.19	1.47	9.00
8.00	3.00	6.79	1.47	1.47	0.00
9.00	5.00	9.82	0.67	3.31	1.00
10.00	5.00	9.82	0.03	3.31	4.00
11.00	6.00	11.34	0.11	11.12	9.00
12.00	6.00	11.34	0.44	11.12	16.00
12.00	6.00	11.34	0.44	11.12	16.00

Tabela 1: Tabela para calcular o coeficiente de determinação.

Coeficiente de determinação (1,5 pontos)

Então, temos que

$$SQREs = 9,59,$$
$$SQReg = 86,41,$$
$$SQTot = 96,$$

e coeficiente de determinação é $R^2 = \dfrac{86,41}{96} = 0,9.$

Interpretação: 90% da variação dos valores da prioridade dada a educação por ser explicada (x) por explicada pela variável investimento por aluno (y) através da equação $\hat{x} = 2,24 + 1,52 \cdot y$ e 10% da variação de x é devida a outros fatores.