

MAE116 - Noções de Estatística
Grupo B - 2º semestre de 2016
Gabarito da Lista de exercícios 3 ¹
Estatística Descritiva III - CASA

Exercício 1.

Uma amostra aleatória de 352 estudantes foi entrevistada, sendo anotado de cada participante o gênero (masculino; feminino), a disciplina que cursa (Estatística; Antropologia) e o resultado final da avaliação (reprovado (r); aprovado (a)). Os resultados estão apresentados no quadro a seguir.

Gênero e disciplina	Resultado da avaliação	
	Reprovado (r)	Aprovado (a)
Feminino e Estatística	24	72
Masculino e Estatística	48	48
Feminino e Antropologia	32	48
Masculino e Antropologia	32	48

- (a) Verifique se há indicação de associação entre resultado da avaliação e gênero. Comente.
(b) Verifique se há indicação de associação entre resultado da avaliação e disciplina cursada. Comente.

Solução.

(a) Para verificar a existência de associação entre resultado da avaliação e gênero construímos o seguinte quadro:

Gênero	Resultado da avaliação		Total
	Reprovado	Aprovado	
Feminino	56 (31,82%)	120 (68,18%)	176 (100,00%)
Masculino	80 (45,45%)	96 (54,55%)	176 (100,00%)
Total	136 (38,64%)	216 (61,36%)	352 (100,00%)

Aparentemente, há indícios de associação entre o resultado da avaliação e o gênero. Pode-se observar que 31,82% dos estudantes de gênero feminino foram reprovados, porém, para os estudantes de gênero masculino foi de 45,45% (perto de 50%).

¹Realizado por: Gualberto Agamez Montalvo

(b) Para verificar a existência de associação entre resultado da avaliação e disciplina cursada construímos o seguinte quadro:

Disciplina	Resultado da avaliação		Total
	Reprovado	Aprovado	
Estatística	72 (37,50%)	120 (62,50%)	192 (100,00%)
Antropologia	64 (40,00%)	96 (60,00%)	160 (100,00%)
Total	136 (38,64%)	216 (61,36%)	352 (100,00%)

Aparentemente, não há indícios de associação entre o resultado da avaliação e a disciplina cursada. Pode-se observar que as porcentagens de cada uma das linhas, são similares à porcentagem total.

Exercício 2.

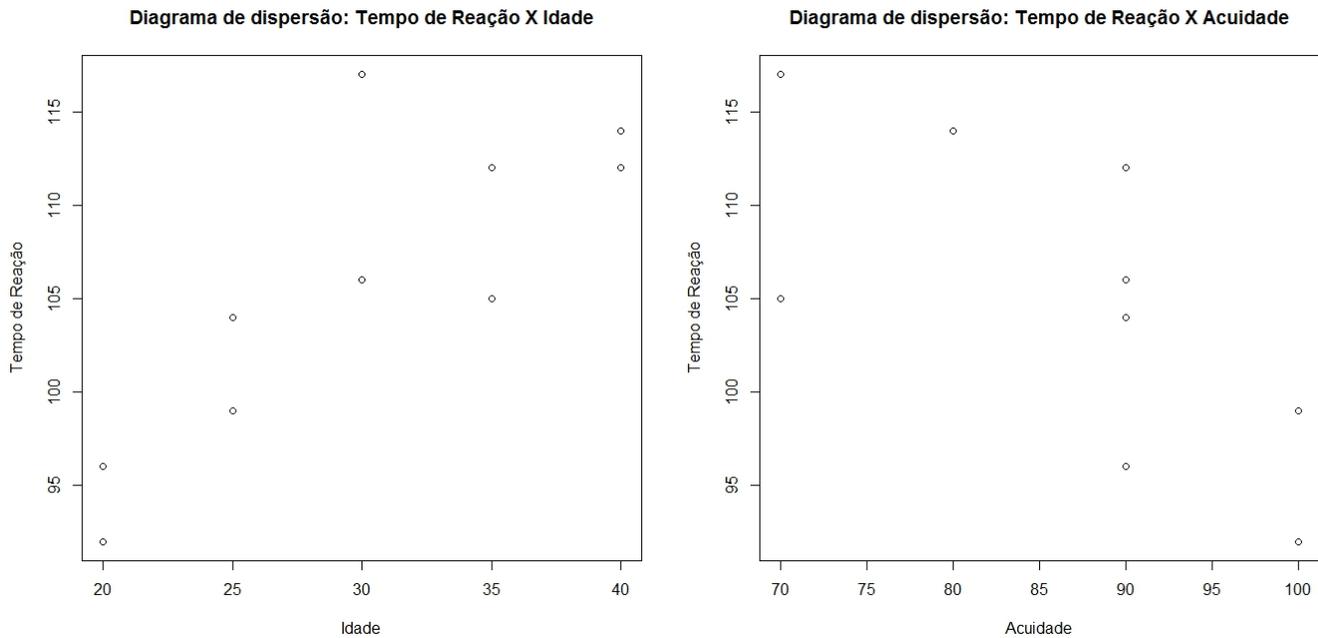
Um psicólogo está investigando a relação entre o tempo que o indivíduo leva para reagir a um certo estímulo (em segundos) e algumas de suas características tais como idade (em anos completos) e acuidade visual (medida em porcentagem). Os dados encontram-se a seguir:

Indivíduo	Tempo de Reação (T)	Idade (I)	Acuidade Visual (A)
1	96	20	90
2	92	20	100
3	99	25	100
4	104	25	90
5	117	30	70
6	106	30	90
7	112	35	90
8	105	35	70
9	114	40	80
10	112	40	90

- (a) Construa os diagramas de dispersão de: Tempo de Reação X Idade e Tempo de Reação X Acuidade Visual.
- (b) Calcule o coeficiente de correlação entre: Tempo de Reação e Idade, e Tempo de Reação e Acuidade Visual e interprete os valores obtidos.
- (c) Como seria possível prever o tempo de reação de um indivíduo através de sua idade? O aumento de um ano na idade aumenta em quanto o tempo médio de reação dos indivíduos? Faça uma previsão do tempo de reação de um indivíduo com 24 anos.
- (d) Obtenha a reta de regressão do tempo de reação em função da acuidade visual. Qual é o significado prático do coeficiente **b** encontrado.
- (e) Como seria mais razoável prever o tempo de reação de uma pessoa: pela idade ou pela acuidade visual? Por quê?

Solução.

(a) Diagramas de dispersão de: Tempo de Reação X Idade e Tempo de Reação X Acuidade Visual.



(b) O coeficiente de correlação linear de Pearson é calculado por:

$$r_{XY} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(n - 1) S_X S_Y},$$

sendo \bar{X} e \bar{Y} as médias amostrais de X e Y , respectivamente, e S_X e S_Y são os desvios padrões de X e Y , respectivamente.

T	I	A	$(T - \bar{T})$	$(I - \bar{I})$	$(A - \bar{A})$	$(T - \bar{T})(I - \bar{I})$	$(T - \bar{T})(A - \bar{A})$
96	20	90	-9,7	-10	3	97,0	-29,1
92	20	100	-13,7	-10	13	137,0	-178,1
99	25	100	-6,7	-5	13	33,5	-87,1
104	25	90	-1,7	-5	3	8,5	-5,1
117	30	70	11,3	0	-17	0,0	-192,1
106	30	90	0,3	0	3	0,0	0,9
112	35	90	6,3	5	3	31,5	18,9
105	35	70	-0,7	5	-17	-3,5	11,9
114	40	80	8,3	10	-7	83,0	-58,1
112	40	90	6,3	10	3	63,0	18,9

Assim, temos que

$$S_T^2 = \frac{(-9, 7)^2 + \dots + (6, 3)^2}{9} = 67, 3444,$$

$$S_I^2 = \frac{(-10)^2 + \dots + (10)^2}{9} = 55, 5556,$$

e

$$S_A^2 = \frac{(3)^2 + \dots + (3)^2}{9} = 112, 2222.$$

Então,

$$r_{TI} = \frac{450}{(9)\sqrt{(67, 3444)(55, 5556)}} = 0, 8174,$$

e

$$r_{TA} = \frac{-499}{(9)\sqrt{(67, 3444)(112, 2222)}} = -0, 6378.$$

Pode-se observar que existe uma relação linear positiva forte entre o Tempo de Reação e a Idade (isto é, quando aumenta a idade também aumenta o tempo de reação), e existe uma relação linear negativa moderada entre o Tempo de Reação e a Acuidade Visual (isto é, quando aumenta a acuidade visual diminui o tempo de reação).

(c) Para prever o tempo de reação de um indivíduo através de sua idade ajustamos a reta:

$$\hat{Y} = a + bX,$$

sendo \hat{Y} o valor predito para o tempo de reação, X a idade,

$$b = \frac{\sum_{i=1}^n Y_i X_i - n\bar{Y}\bar{X}}{(n-1)S_X^2} \quad \text{e} \quad a = \bar{Y} - b\bar{X}.$$

Assim, temos que

$$b = \frac{450}{(9)(55, 5556)} = 0, 90 \quad \text{e} \quad a = 105, 7 - (0, 9)(30) = 78, 7$$

Portanto

$$\hat{Y} = 78, 7 + 0, 9X$$

Da equação anterior, podemos inferir que, o aumento de um ano na idade aumenta em 0,9 segundos o tempo médio de reação dos indivíduos.

Finalmente, a previsão do tempo de reação de um indivíduo com 24 anos é $\hat{Y} = 78,7 + 0,9(24) = 100,3$ (100,3 segundos).

(d) Para prever o tempo de reação de um indivíduo em função da acuidade visual, denotamos \hat{Y} como o valor predito para o tempo de reação e X como a acuidade visual. Então

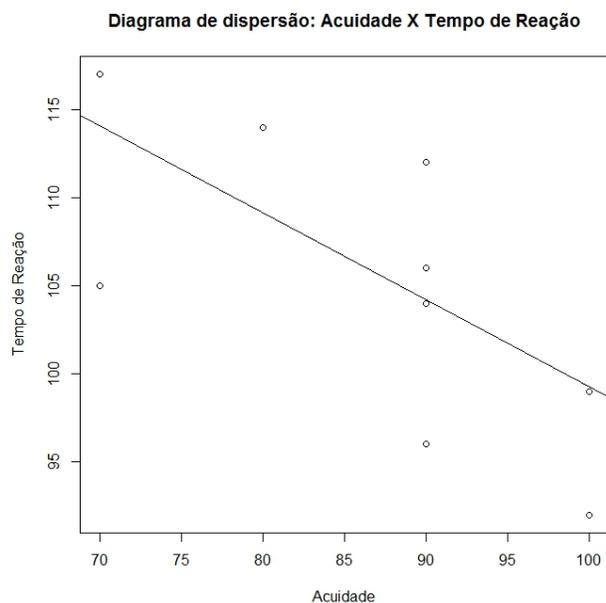
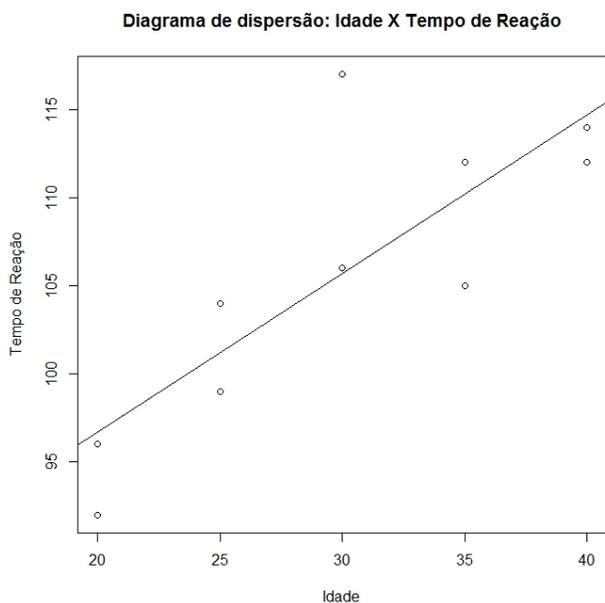
$$b = \frac{-499}{(9)(112,2222)} = -0,4941 \quad \text{e} \quad a = 105,7 - (-0,4941)(87) = 148,6832.$$

Portanto

$$\hat{Y} = 148,6832 - 0,4941X$$

Interpretação de b: para um aumento de uma unidade da acuidade visual, o tempo médio de reação dos indivíduos diminui em 0,4941 segundos.

(e) Seria mais razoável prever o tempo de reação de uma pessoa pela idade, pois sua correlação é maior em comparação com a acuidade visual, e aparentemente a ajuste dos dados é melhor (observar as seguintes figuras).



Exercício 3.

Os dados a seguir referem-se a porcentagem de crianças imunizadas pela vacina tríplice (difteria, tétano e coqueluche) e a taxa de mortalidade infantil em 20 países, em 1992.

País	X	Y
Bolívia	77	118
Brasil	69	65
Cambodja	32	184
Canadá	85	8
China	94	43
República Tcheca	99	12
Egito	89	55
Etiópia	13	208
Finlândia	95	7
França	95	9
Grécia	54	9
Índia	89	124
Itália	95	10
Japão	87	6
México	91	33
Polônia	98	16
Rússia	73	32
Senegal	47	145
Turquia	76	87
Reino Unido	90	9

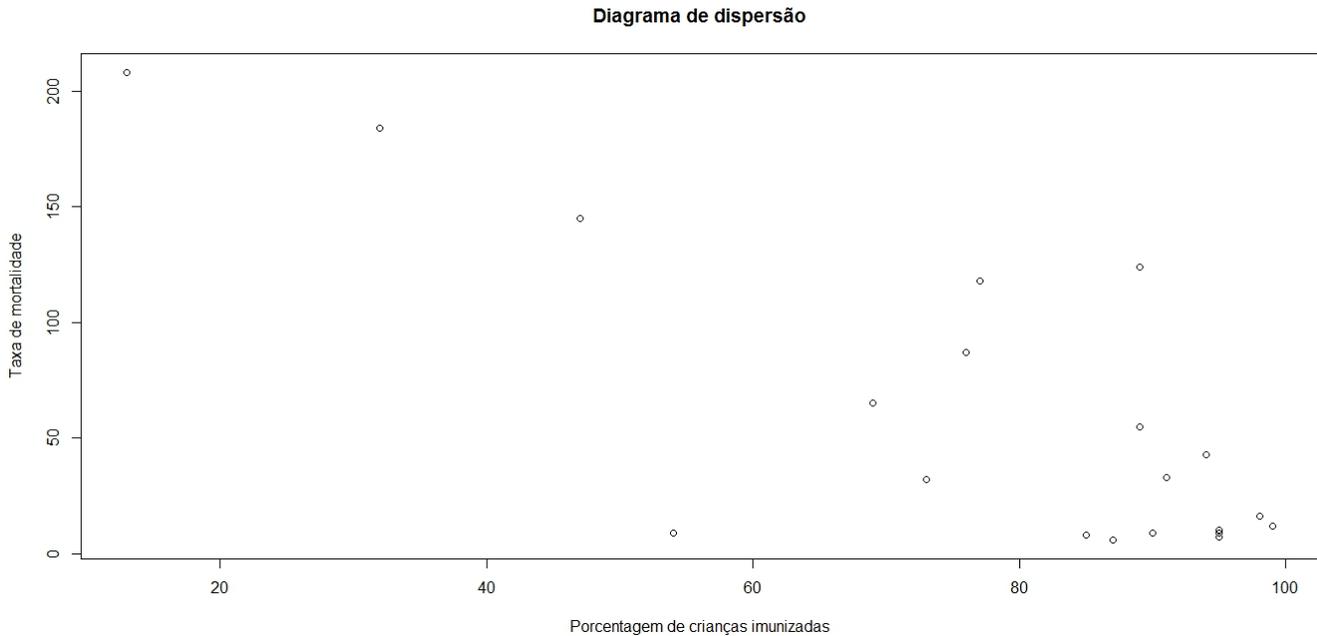
sendo; X como a porcentagem de crianças imunizadas e Y como a taxa de mortalidade por 1000 nascidos.

- (a) Construa o gráfico de dispersão para Y versus X . Comente sobre a relação entre as variáveis.
- (b) Obtenha a reta de regressão da taxa de mortalidade em função da porcentagem de crianças imunizadas. Qual é a interpretação do coeficiente b encontrado?

(c) Qual é a taxa de mortalidade prevista em um país que tenha 70% das crianças imunizadas?

Solução.

(a) Diagrama de dispersão da taxa de mortalidade por 1000 nascidos versus a porcentagem de crianças imunizadas.



Pode-se notar que aparentemente, conforme aumenta a porcentagem de crianças imunizadas (X), a taxa de mortalidade (Y) tende a diminuir. Nota-se também que a tendência é linear.

(b) Para obter a reta de regressão da taxa de mortalidade em função da porcentagem de crianças imunizadas, precisamos dos valores de a e b da reta:

$$\hat{Y} = a + bX,$$

sendo \hat{Y} o valor predito para a taxa de mortalidade, e X a porcentagem de crianças imunizadas.

$$b = \frac{\sum_{i=1}^n Y_i X_i - n\bar{Y}\bar{X}}{(n-1)S_X^2} \quad \text{e} \quad a = \bar{Y} - b\bar{X}.$$

Assim, temos que

$$b = \frac{68626 - (20)(77,4)(59)}{(20)(559,5158)} = -2,136 \quad \text{e} \quad a = 59 - (-2,136)(77,4) = 224,326$$

Portanto

$$\hat{Y} = 224,326 - 2,136X$$

Então, podemos inferir que, o aumento de uma unidade da porcentagem de crianças imunizadas diminui em 2,136 a taxa média de mortalidade (isto é, deixam de morrer 2136 nascidos).

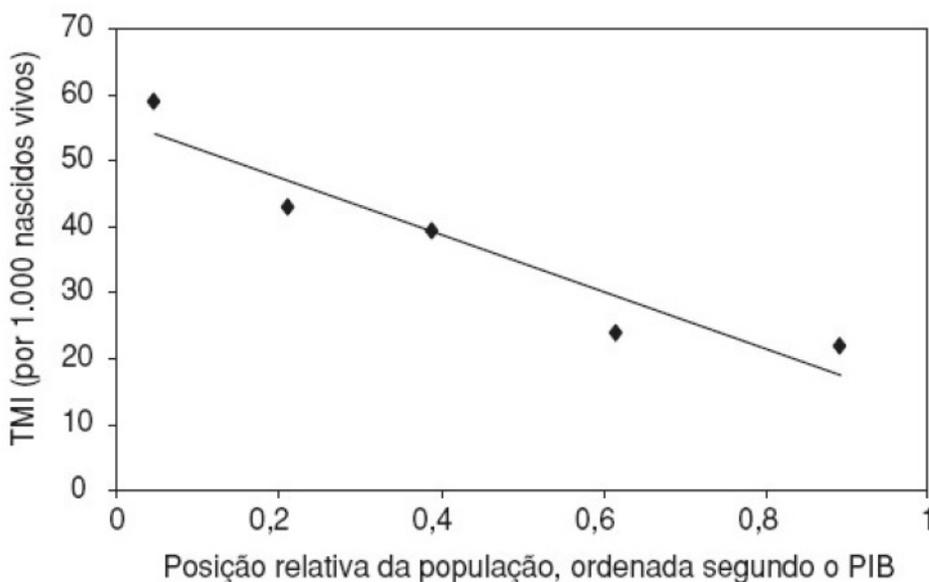
(c) A taxa de mortalidade prevista em um país que tenha 70% das crianças imunizadas é $\hat{Y} = 224,326 - (2,136)(70) = 74,806$ (74806 crianças mortas).

Exercício 4.

No artigo intitulado “Métodos de mensuração das desigualdades em saúde” publicado em Rev. Panam. Salud Publica 12(6), 2002, determinou-se que a relação entre a Taxa de mortalidade infantil e uma posição relativa de um país, obtida segundo o PIB, é linear e dada pela expressão:

$$\hat{y} = 53,38 - 40,46x,$$

em que \hat{y} é a Taxa de mortalidade infantil estimada (por 1000 nascidos vivos) e x a posição relativa do país segundo o PIB. Também é apresentado o gráfico a seguir.



- (a) Interprete o coeficiente angular estimado.
- (b) Qual será a mortalidade esperada se a posição do país for 0,67?

Solução.

(a) Ao aumentar 1 unidade na posição relativa do país segundo o PIB tem-se em média 40,46 unidades a menos na taxa de mortalidade infantil (isto é, deixam de morrer 40460 crianças nascidas vivas).

(b) A mortalidade esperada se a posição do país for 0,67 é aproximadamente de 26272 crianças nascidas vivas, pois, $\hat{y} = 53,38 - (40,46)(0,67) = 26,2718$.