10.8 Exercícios

10.8.1 Dos temas Formas de dependência, Covariância, Coeficiente de correlação linear

Exercício 172 (fonte: Bussab, Morettin, "Estatística Básica"). Recorde que se duas variáveis aleatórias X e Y são independents, então, obrigatoriamente, $\varrho(X, Y) = 0$. Recorde também que a recípoca dessa consequencia não é válida, quer dizer, existem pares de variáveis aleatórias cujo coeficiente de correlação linear é nulo mas que, a despeito disso, as variáveis são dependentes. Dois de tais pares estão apresentados nas tabelas (a) e (b) abaixo. A tarefa do presente exercício é confirmar, fazendo as contas, que em cada caso, o coeficiente de correlação linear é nulo e que as variáveis são dependentes.

(a)

Y / X	-1	0	1
-1	0	1/4	0
0	1/4	0	1/4
1	0	1/4	0

(b)

X / Y	0	1	2
1	3/20	3/20	2/20
2	1/20	1/20	2/20
3	4/20	1/20	3/20

Exercício 173. Quem é o responsável pela anulação da covâriancia entre duas variáveis aleatórias: os valores das mesmas ou as probabilidades da distribuição conjunta? A resposta é: ambos. É com o intuito de ilustrar essa resposta foram moldadas duas tarefas do presente exercício.
(a) Na distribuição (b) do Exc. 172, substitua a primeira linha das probabilidades por $5/20, 2/20, 1/20$ e mostre, fazendo as contas, que a alteração de probabilidades fez com que $\text{Cov}(X, Y)$ não se anula.
(b) Na distribuição (b) do Exc. 172, substitua os valores 0, 1, 2 da variável aleatória X por 2, 1, 3 e mostre, fazendo as contas, que a alteração de valores fez com que $\text{Cov}(X, Y)$ não se anula.

Exercício 174. Recorde que o coeficiente de correlação linear assume valor 1 ou -1 se e somente se X e Y estão em relação linear entre si (quer dizer, a relação linear sem ruído, i.e., $Y = a + bX$). Em qualquer outro caso, o coeficiente de correlação linear não pode ser nem 1 nem -1. Em particular, tais valores são impossíveis no caso quando X e Y estão em relação de dependência absoluta que não é linear. O presente exercício pede de você confirma estes fatos fazendo o cálculo para o valor de $\varrho(X, Y)$ nos dois casos que estão apresentados nas tabelas abaixo. Observe que no primeiro caso, X e Y estão em relação linear crescente, e então o resultado de seu cálculo deve dar que $\varrho(X, Y) = 1$. Se o resultado foi diferente de 1, então você errou em algum lugar; volte a ver a fórmula de cálculo para $\varrho(X, Y)$ e refaça o cálculo. Já falando do segundo caso, chamo sua atenção ao fato de que X e Y estão em relação de dependência absoluta que não é linear. Portanto, nesse caso o resultado de seu cálculo não pode ser nem 1 nem -1.

X / Y	0	1	2
1	1/3	0	0
3	0	1/3	0
5	0	0	1/3

X / Y	0	1	2
1	1/3	0	0
3	0	0	1/3
5	0	1/3	0

Exercício 175. Considere as variáveis aleatórias X e Y cuja distribuição conjunta está apresentada na tabela abaixo.

X / Y	1	2	3
1	1/3	0	
2	0	1/3	
3	0	0	

Complete a tabela da maneira tal que X e Y estejam em relação linear entre si, e confirme essa relação calculando o coeficiente de correlação linear entre X e Y.

Exercício 176. Considere as variáveis aleatórias X e Y cuja distribuição conjunta está dada na tabela abaixo.

X / Y	1	2	3
1	1/4	0	1/4
2	0	1/4	0
3	0	0	1/4

Modifique somente duas probabilidades desta tabela da maneira tal que as variáveis aleatórias X e Y estejam em relação linear, e confirme essa relação calculando o coeficiente de correlação linear entre X e Y.

Exercício 177. Na tabela abaixo há a distribuição conjunta de duas variáveis aleatórias X e Y. Conforme você pode ver, a tabela não definiu dois valores. Coloque nas celulas vazios os valores a seu gosto. Note, entretanto, que seus valores devem obedecer à seguinte restrição: a soma de todas as probabilidades da tabela deve dar 1. Calcule $\text{Cov}(X, Y)$ para a tabela completada. Agora escolhe outros dois valores e calcule $\text{Cov}(X, Y)$ novamente. As duas covariâncias são iguais? Elas têm obrigação a dar valores iguais?

Y / X	-1	0	1
-1	1/20	3/20	
0	4/20		2/20
1	1/20	3/20	2/20

Exercício 178. Considere o experimento aleatório "três laçamentos de uma moeda honesta". No conjunto de seus resultados, foi definida variável aleatória X da maneira expressa pela tabela abaixo. Sua tarefa é definir no mesmo conjunto variável aleatória Y da maneira tal que X e Y estejam em relação linear, e tal que os valores que Y assume nos resultados (hht) e (hth) sejam 3 e 6 respectivamente (isso já está imposto pela tabela). Faça a tabela da distribuição conjunta de X e Y para confirmar que, de fato, há relação linear entre X e Y.

resultado	valor de X	valor de Y
hhh	3	
hht	1	3
hth	2	6
thh	2	
tth	3	
tht	1	
htt	1	
ttt	2	

Exercício 179. Considere o experimento aleatório "lançamento de um dado equilibrado". Neste experimento foram definidos todos os valores da variável aleatória X e dois valores da variável aleatória Y; os valores defindos estão na tabela abaixo.

resultado	valor de X	valor de Y
1	2	4
2	3	6
3	4	
4	4	
5	3	
6	2	

(a) Pede-se completar a definição da variável aleatória Y de maneira tal que as variáveis aleatórias X e Y estejam em relação linear.
(b) Pede-se construir a tabela da distribuição conjunta das variáveis aleatórias X e Y.

Exercício 180. Lançamos um dado equilibrado e uma moeda honesta com os números -1 e 1 nas suas faces "cara" e "coroa". Designamos por X o número obtido no dado. Designamos por Y o dobro do número obtido no dado acrescido pelo número obtido na moeda. Pede-se desenhar a tabela da distribuição conjunta de X e Y. Responda:
(a) X e Y são variáveis aleatórias independentes?
(b) X e Y têm dependência absoluta?
(c) Calcule a covariância e o coeficiente de correlação linear entre X e Y.
(d) Verifique a validade da relação genérica $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y)$ fazendo os cálculos para as variáveis aleatórias do presente exercício.

Exercício 181. Lançamos um tetraedro equilibrado com os números 1,2,3 e 4 em suas faces, e uma moeda honesta com os números 1 e 2 nas suas faces "cara" e "coroa". Definimos duas variáveis aleatórias U e T da seguinte maneira: U é a soma dos valores obtidos no tetráedro e na moeda; T é a diferença entre estes valores. Pede-se desenhar a tabela da distribuição conjunta de U e T. Responda:
(a) U e T são variáveis aleatórias independentes?
(b) U e T têm dependência absoluta?
(c) Calcule a covariância e o coeficiente de correlação linear entre U e T.
(d) Verifique a validade da relação genérica $\text{Var}(U+T) = \text{Var}(U) + \text{Var}(T) + 2\text{Cov}(U,T)$ fazendo os cálculos para as variáveis aleatórias do presente exercício.

Exercício 182. Sejam o tetráedro e a moeda como no Exercício 181. Definimos V como a soma dos valores obtidos no tetráedro e na moeda e definimos W como o dobro desta soma. Pede-se desenhar a tabela da distribuição conjunta de V e W. Responda:
(a) V e W são variáveis aleatórias independentes?
(b) V e W têm dependência absoluta?
(c) Calcule a covariância e o coeficiente de correlação linear entre V e W.
(d) Verifique a validade da relação genérica $\text{Var}(V+W) = \text{Var}(V) + \text{Var}(W) + 2\text{Cov}(V,W)$ fazendo os cálculos para as variáveis aleatórias do presente exercício.

Exercício 183. Sejam o tetráedro e a moeda como no Exercício 181. Designamos por A o número obtido no tetráedro e por B o número obtido na moeda. Pede-se desenhar a tabela da distribuição conjunta de A e B. Responda:
(a) A e B são variáveis aleatórias independentes?
(b) A e B têm dependência perfeita?

(c) Calcule a covariância e o coeficiente de correlação linear entre A e B.
(d) Verifique, fazendo as contas numéricas, a validade da relação genérica $\mathrm{Var}(A+B) = \mathrm{Var}(A) + \mathrm{Var}(B) + 2\mathrm{Cov}(A,B)$ fazendo os cálculos para as variáveis aleatórias do presente exercício.

Exercício 184. Considere o experimento aleatório "lançamento de um dado equlibrado". Define neste experimento
(a) variáveis aleatórias X e Y da maneira tal que elas estejam em relaçao linear entre si;
(b) variáveis aleatórias Z e U da maneira tal que elas estejam em relação absoluta mas não linear;
(c) variáveis aleatórias V e W da maneira tal que não haja relação absoluta entre elas.

Exercício 185. Calcule o coeficiente de correlação linear entre
(a) X e Y, (b) Z e U, (c) V e W,
construidas no Exercício 184.

Exercício 186 *(nada parecido será cobrado de você nas provas do curso).* Construa a distribuição conjunta de duas variáveis aleatórias A e B da maneira tal que elas sejam independentes, mas que a distribuição de A coincida com a de X (do Exercício 184) e a distribuição de B coincida com a de Y (do Exercício 184). Responda: as variáveis aleatórias A e B construidas podem ser definidas no mesmo experimento aleatório que foi usado para a construção das variáveis aleatórias X e Y no Exercício 184?

Exercício 187. Seja X uma variável aleatória que assume valores 1, 2 e 3 com as probabilidades 1/4, 1/2 e 1/4, e seja E uma oura variável aleatória, que é independente de X, e que assume valores $-1/10$ e $+1/10$ com as probabilidades 1/2 e 1/2. Definimos $Y = 3X + 1 + E$. Calcule a covariância e o coeficiente de correlação linear entre X e Y.

Exercício 188. Seja X uma variável aleatória que assume valores 1, 2 e 3 com as probabilidades 1/4, 1/2 e 1/4, e seja E uma oura variável aleatória, que é independente de X, e que assume valores -1 e 1 com as probabilidades 1/2 e 1/2. Definimos $Y = 3X + 1 + E$. Calcule a covariância e o coeficiente de correlação linear entre X e Y. Compare os resultados com os do Exercício 187.

comentar

Exercício 189. Seja X uma variável aleatória que assume valores 1, 2 e 3 com as probabilidades 1/3, 1/3 e 1/3, e seja E uma oura variável aleatória, que é independente de X, e que assume valores -10 e 10 com as probabilidades 1/2 e 1/2. Definimos $Y = 3X + 1 + E$. Calcule a covariância e o coeficiente de correlação linear entre X e Y. Compare os resultados com os dos Exercício 187 e 188.

comentar

Exercício 190. Seja X uma variável aleatória que assume valores 1, 2 e 3 com as probabilidades 1/4, 1/2 e 1/4. Definimos $Y = X^2$. Calcule a covariância e o coeficiente de correlação linear entre X e Y. Compare o valor do coeficiente de correlação linear obtido agora com os obtidos nos Exercícios 187, 188 e 189.

comentar

Exercício 191 *(nada parecido será cobrado de você nas provas do curso).*
(a) Construa um experimento aleatório e define nele três variáveis aleatórias X, E, Y da maneira tal que relação entre Y e X seja via função linear desregrada por um ruido homogêneo e não viesado, mas que o ruido não tenha distribuição simêtrica.
(b) Construa um experimento aleatório e define nele três variáveis aleatórias X, E, Y da maneira tal que relação entre Y e X seja via função linear desregrada por um ruido não viesado e simêtrico, mas que este ruido não seja homogêno.

(c) Construa um experimento aleatório e define nele três variáveis aleatórias X, E, Y da maneira tal que relação entre Y e X seja via função linear desregrada por um ruido homogêneo, cuja distribuição seja simêtrica em torno do valor de $\mathbb{E}[E]$, mas que este valor seja diferente de zero, quer dizer, que o ruido esteja viezado.

10.8.2 Do tema Método de mínimos quadrados

Exercício 192. **(a)** Ajuste a reta de mínimos quadrados $y = \alpha + \beta x$ para para cada um dos quatro conjuntos abaixo:

$$\begin{array}{ll} (a) & (0\,;\,4),\ (1\,;\,4),\ (2\,;\,3),\ (3\,;\,1) \\ (b) & (0\,;\,4),\ (1\,;\,3),\ (3\,;\,2),\ (4\,;\,0) \\ (c) & (1\,;\,4),\ (2\,;\,3),\ (3\,;\,1),\ (4\,;\,1) \\ (d) & (0\,;\,5),\ (1\,;\,3),\ (2\,;\,2),\ (4\,;\,1) \end{array}$$

Desenhe os ponots do conjunto e a reta ajustada em cada caso (a posição da reta pode ser desenhada aproximadamente). Em cada caso, marque os resíduos em relação à reta ajustada e calcule e apresente a soma dos quadrados desses resíduos.
(b) Marque os resíduos em relação à reta
$y = 3$ no caso (a) (observe que $y = 3$ não é a reta de mínimos quadrados para *(a)*),
$y = x$ no caso (b) (observe que $y = x$ não é a reta de mínimos quadrados para *(b)*),
$y = -x + 3$ no caso (c) (observe que $y = -x + 3$ não é a reta de mínimos quadrados para *(c)*),
$y = -x + 5$ no caso (d) (observe que $y = -x + 5$ não é a reta de mínimos quadrados para *(d)*),
e calcule a soma dos quadrados desses resíduos; em cada caso, compare a soma calculada com a soma dos quadrados dos resíduos em relação à respectiva reta de mínimos quadrados.

Exercício 193. Dois pontos são dados

$$(0\,;\,0) \quad \text{e} \quad (3\,;\,0)$$

(a) Pede-se acrescentar terceiro ponto de sorte tal que a reta ajustada para os três pontos pelo método de mínimos quadrados tenha inclinação negativa.
(b) Pede-se acrescentar terceiro ponto de sorte tal que a reta ajustada para os três pontos pelo método de mínimos quadrados tenha inclinação positiva.

Exercício 194. Dos pontos são dados:

$$(0\,;\,1) \quad \text{e} \quad (3\,;\,2)$$

(a) Pede-se acrescentar terceiro ponto de sorte tal que a reta ajustada para os três pontos pelo método de mínimos quadrados tenha inclinação negativa.
(b) Pede-se acrescentar terceiro ponto de sorte tal que a reta ajustada para os três pontos pelo método de mínimos quadrados tenha inclinação positiva.

Exercício 195. Dado o conjunto

$$(0\,;\,0),\ (0,5\,;\,0,7),\ (1\,;\,1)$$

verifique se a reta ajustada a esta pelo método de mínimos quadrados, tem inclinação positiva ou negativa. Acrescente um ponto à amostra de sorte tal que a reta ajustada para os quatro pontos da nova amostra pelo método de mínimos quadrados tenha inclinação oposta à da reta ajustada para três pontos.

Exercício 196. Uma reta $y = \beta x + \alpha$ foi ajustada pelo método de mínimos quadrados, a um certo conjuto

$$(x_1\,;\,y_1),\ (x_2\,;\,y_2), \ldots, (x_n\,;\,y_n)$$

Acerca desse conjunto e da reta ajustada, dou-lhe as seguintes informações parciais:
(i) que o valor de β é negativo,
(ii) que o conjunto contém somente um ponto com coordenadas $(0\,,\,0)$ e que o quadrado do resíduo do ponto em relação da reta ajustada é 4,
(iii) que o conjunto contem somente um ponto com coordenadas $(4\,,\,1)$ e que o quadrado do resíduo do ponto em relação da reta ajustada é 9.
Peço que, com base nestas informações, você ache os valores α e β da reta ajustada.

10.8.3 Do tema Regressão linear simples

Os exercícios dessa seção são parecidos entre si. Cada um fornece uma amostra, pede calcular o coeficiente amostral de correlação linear, interpretar seu valor em termos da indicação da presença do Modelo de Regressão Linear Simples entre as variáveis aleatórias que geraram a amostra, estimar os valores de a e b desse modelo usando a reta de mínimos quadrados, e, por fim, fazer uma previsão usando a reta ajustada. Entretanto, de um exercício para outro, há uma ligeira variação de linguagem usada para colocar a tarefa. Eu propositalmente não unifiquei as linguagens, pois assim que consigo mostrar a diversificação de das formas de expressão usadas na prática.

Dois avisos:
(1) Os termos "o coeficiente amostral de correlacão linear" e "a estimativa de coeficiente de correlacão linear" são sinónimos, e a notação usada para esses é r ou r_{xy}.
(2) A expressão "reta de regressão" no âmbito da área Regressão Linear Simples significa "a reta de mínimos quadrados".

Exercício 197. (Exercício 1 do Capítulo 10 do livro "Noções de Estatística".)
(a) Calcule, sem o uso de computador, r_{xy}, a estimativa de coeficiente de correlação linear de Pearson entre duas variáveis aleatórias X e Y, na base de 6 pares das suas realizações

x_i	19	24	26	31	31	32
y_i	17	35	11	29	14	31

O resultado indica a existência de relação linear entre X e Y? Se "sim", porque os pontos $(x_i, y_i), i = 1, 2, \ldots, 6$, não pertencem a uma reta? Faça o diagrama de dispersão das observações de X e Y para se certificar que os pontos do diagrama realmente não estão alinhados.
(b) Quais são os valores de α e β que minimizam a expressão

$$\sum_{i=1}^{6} (y_i - (\alpha + \beta x_i))^2 \tag{10.9}$$

em que $(x_1, y_1), \ldots, (x_6, y_6)$ são as pares da tabela acima? Para os valores encontrados de α e β, trace uma reta $y = \alpha + \beta x$ no diagrama de dispersão que foi construida na sua resposta ao item (a). Qual a interpretação desta reta? Qual a sua interpretação em termos das variáveis aleatórias X e Y?

Exercício 198. Com o objetivo de estudar a relação entre nivel educacional de pais e filhos, tomou-se a seguinte amostra (obs.: o nível educacional esta mensurado em anos completos de frequência a escola).

Família	Pais	Criança
A	12	12
B	10	8
C	6	6
D	16	11
E	8	10
F	9	8
G	12	11

(a) Encontre o coeficiente de correlacão linear amostral e faça a conclusão sobre a existência de relação linear entre os fatores "nivel educacional de pais" e "nivel educacional de filho".
(b) Faça um esboço do gráfico da reta de mínimos quadrados junto com os pontos da amostra.
(c) Comente a relação entre as duas variáveis usando os resultados obtidos. (Pode desconsiderar a tarefa (c) caso esta o incomoda por razão qualquer que seja.)

Exercício 199.
(a) Calcule a estimativa, r_{xy}, do coeficiente de correlação linear entre duas variáveis aleatórias X e Y na base de 6 pares das suas realizações

x_i	19	24	26	31	31	32
y_i	0,8	1,2	1,5	2,0	2,3	2,5

O resultado indica a existência de relação linear entre X e Y? Assumindo que X e Y satisfazem o modelo de regressão linear $Y = a + bX + E$, estime os valores de a e b.
(b) Na tabela abaixo, cada valor da segunda linha é 10 vezes maior que o correspondente valor na segunda linha da tabela do item (a) acima (para evitar a confusão, a segunda linha da tabela abaixo foi denotada por "z", e toda a tabela interpreta-se como amostra do par de variáveis aleatórias (X, Z)).

x_i	19	24	26	31	31	32
z_i	8	12	15	20	23	25

Com base no resultado do item (a) e na relação entre as segundas linhas das duas tabelas (pelo fator 10, conforme ilusinado acima), dé o valor de r_{xz}. Assuma agora que X e Z satisfazem o modelo $Z = a' + b'X + \mathcal{E}$. Quais são as estimativas de a' e b' fornecidas pela amostra da tabela via o método de Regressão Linear? Responda nesta pergunta com base no resultado do item (a) e na relação entre as segundas linhas das duas tabelas (pelo fator 10, conforme ilusinado acima). Agora faça contas diretas para confirmar suas respostas acerca dos valores de r_{xz} e das estimativas de a' e b'.

Exercício 200 (banco de exercícios de MAE0116 ministrada na USP no 1^o sem. de 2013).
Uma substância utilizada em pesquisas nas áreas médica e biológica é enviada pelo fabricante aos seus usuários por via aérea em caixas contendo 1000 ampolas. Os dados abaixo, envolvendo 10 remessas, correspondem ao número de vezes que a caixa foi transferida de uma aeronave para a outra durante o percurso (X) e o número de ampolas que foram encontradas quebradas no momento em que a caixa foi aberta pelo usuário (Y).

nr. de vezes que a caixa foi transferida(X)	1	0	2	0	3	1	0	1	2	0
número de ampolas quebradas (Y)	16	9	17	12	25	13	7	15	19	11

(a) Faça à mão o diagrama de dispersão entre X e Y. O que este gráfico sugere?
(b) Calcule à mão o coeficiente amostral de correlação linear de Pearson entre X e Y.

(c) Ajuste uma reta de regressão para a relação entre as variáveis Y (dependente) e X (independente).
(d) Interprete o valor do coeficiente angular "b"da reta ajustada.

Exercício 201 (banco de exercícios de MAE0116 ministrada na USP no 1^o sem. de 2013). Uma empresa que presta serviço de manutenção de computadores coletou dados referentes a tempos de reparo (em minutos) e número componentes reparados ou substituídos em 30 computadores. Os dados são apresentados na tabela abaixo.

Número de componentes reparados	Tempo de reparo	Número de componentes reparados	Tempo de reparo	Número de componentes reparados	Tempo de reparo
1	19	9	125	13	171
1	24	9	147	13	184
2	23	9	140	14	183
2	35	10	148	15	180
3	49	10	157	15	191
5	75	11	155	16	194
5	87	11	168	17	195
6	96	12	170	18	198
7	99	12	185	20	197
8	120	12	174	20	208

(a) Faça o diagrama de dispersão entre "Número de componentes reparados" e "Tempo de reparo". O que este gráfico sugere?
(b) Calcule o coeficiente amostral de correlação linear entre os fatores "Número de componentes reparados" e "Tempo de reparo".
(c) Ajuste uma reta de regressão para os dados da tabela.
(d) Interprete o valor do coeficiente angular β da reta ajustada. Considerando a reta ajustada, estime o tempo esperado para reparo de computador com 8 componentes a serem reparados. Observe que entre os dados coletados há o par $(8, 120)$, quer dizer, para um caso de substituição de 8 componentes, foi preciso 120 minutos de reparo. Sua estimação do tempo esperado para a substituição de 8 componentes é 120? É obrigado a ser exatamente 120?

Exercício 202. Uma indústria submete seus novos operários a um teste de aptidão (fator x) e três meses depois mede a produtividade destes operários (fator y). Os resultados estão na tabela a seguir:

Operário	1	2	3	4	5	6
Aptidão	22	25	15	19	22	18
Produtividade	45	37	25	40	33	30

Faça o diagrama de dispersão da amostra e calcule o coeficiente amostral de correlação linear. Ajuste a reta $y = \alpha + \beta x$ aos dados pelo método de mínimos quadrados. Dê a interpretação (em termos da aptidão dos operários e suas produtividades) dos valores a e b. Calcule a estimativa da produtividade média de operários com aptidão igual a 20.

Exercício 203 (banco de exercícios de MAE0116 ministrada na USP no 1^o sem. de 2013). São apresentados abaixo o comprimento e a altura (em cm) de 10 tartarugas do sexo masculino de uma determinada espécie.

Tartaruga	1	2	3	4	5	6	7	8	9	10
Comprimento (X)	93	94	96	101	102	103	104	106	107	112
Altura (Y)	36	35	36	38	38	37	39	39	38	40

(a) Faça o diagrama de dispersão entre comprimento (X) e altura (Y). O que este gráfico sugere?
(b) Calcule o coeficiente amostral de correlação linear de Pearson entre as variáveis X e Y.
(c) Ajuste uma reta de regressão para a relação entre as variáveis Y (dependente) e X (independente). Interprete o valor do coeficiente angular "b"da reta ajustada.
(d) Considerando a reta ajustada em (c), estime a altura esperada de tartarugas com comprimento de 95 cm e 105 cm.

Exercício 204 (banco de exercícios de MAE0116 ministrada na USP no 1^o sem. de 2013). Deseja-se provar a eficiência de certo fungicida para o controle de pragas em plantações de trigo. Diferentes doses (em gr/ha) de um fungicida foram aplicadas em 10 canteiros de produção, contendo 100 plantas cada um. Após 15 dias da aplicação do fungicida foi feita a contagem do número de plantas doentes. A tabela a seguir apresenta os dados do número de plantas doentes (Y) de acordo com a dose (X):

X (dose)	100	125	200	250	275	300	325	350	375	400
Y (plantas doentes)	50	48	39	35	30	24	20	12	10	5

(Os valores que podem ajudar em seus cálculos: $\sum_{i=1}^{10} x_i = 2700$, $\sum_{i=1}^{10} y_i = 273$, $\sum_{i=1}^{10} x_i^2 = 822500$, $\sum_{i=1}^{10} y_i^2 = 9695$, $\sum_{i=1}^{10} x_i y_i = 59450$.)
(a) Construa o diagrama de dispersão e interprete o relacionamento entre as variáveis Y e X.
(b) Calcule o coeficiente amostral de correlação linear de Pearson entre X e Y. (c) Ajuste uma reta de regressão para a relação entre as variáveis Y: número de plantas doentes (dependente) e X: dose (independente). Interprete o valor do coeficiente angular "b"da reta ajustada.
(d) Considerando a reta ajustada em (c), estime o número esperado de plantas doentes para cada 100 plantas tratadas com o fungicida na dose 260 gr/ha.

Exercício 205. Os dados a seguir referem-se a meses de experiência de 10 datilógrafos e o número de erros cometidos na datilografia de determinado texto.

Meses de experiência	1	2	2	3	4	6	7	8	9	10	10
Número de erros	30	28	10	24	20	18	14	13	10	7	6

(a) Construa (à mão) o diagrama de dispersão.
(b) Calcule (à mão) o coeficiente amsotral de correlação linear entre o fator "Meses de experiência" (X) e o fator "Número de erros" (Y) e interprete seu valor.
(c) Obtenha (à mão) a reta de regressão de Y em função de X e desenhe esta reta no diagrama de dispersão do item (a).
(d) Qual é o significado do coeficiente b obtido?
(e) Existe algum datilógrafo com comportamento diferente dos demais? Se existir elimine o ponto correspondente e calcule o coeficiente de correlação e ajuste a reta de regressão ao conjunto de dados sem essa observação.
(f) Compare os valores de b e do coeficiente de correlação para os dois conjuntos de dados.

Exercício 206. Os dados a seguir correspondem às variáveis X (peso médio do corpo, em kg) e Y (peso médio do cérebro, em g) para 13 espécies de animais. EXPLICAR PORQUE TOMARAM A MÉDIA POR CADA ESPCIE.

Animal	Peso do corpo	Peso do cérebro
Vaca	465	423
Gato	3,3	25,6
Canguru	35	56
Camundongo	0,023	0,4
Coelho	2,5	12,1
Jaguar	100	157

Animal	Peso do corpo	Peso do cérebro
Rato	0,28	1,9
Porco	192	180
Chimpanzé	52,16	440
Hamster	0,12	1
Girafa	529	680
Gorilla	207	406
Cavalo	521	655

(a) Construa o diagrama de dispersão e obtenha o coeficiente amostral de correlação linear entre as variáveis X e Y. Comente.
(b) Obtenha a equação de regressão do peso médio do cérebro em função do peso médio do corpo do animal.
(c) Faça uma previsão do peso médio do cérebro para animais com peso médio corporal igual a 62 kg.
(d) Estudos apontaram que o peso médio do homem adulto é 62 kg e o peso médio do cérebro é 1320g. Com base no item (b), o que você pode concluir sobre a relação entre peso do cérebro e peso do corpo para o homem em comparação com a dos animais listados?

Exercício 207. Os dados a seguir representam 26 pacientes que foram contaminados pelo veneno de certo tipo de inseto e submetidos a um de três tipos de tratamento. As variáveis são:
Idade: idade do paciente no momento de admissão, em anos;
Diag: tempo, em horas, gasto entre o contato com o inseto e administração do tratamento;
Recup: tempo, em horas, entre a administração do tratamento e recuperação;
Tratam: tipo do tratamento administrado;
Coag: presença de coágulos no momento de admissão.

Nr de Paciente	*Idade*	*Diag*	*Recup*	*Tratam*	*Coag*
1	28	7	3	II	Não
2	15	52	45	I	Não
3	76	30	23	III	Sim
4	15	53	46	I	Sim
5	21	3	2	II	Não
6	11	46	42	I	Não
7	16	55	47	I	Não
8	16	54	47	I	Sim
9	47	13	12	III	Sim
10	18	59	51	II	Não
11	40	20	11	III	Sim
12	24	3	1	II	Não
13	32	9	3	II	Não
14	31	9	3	II	Não
15	10	44	40	I	Sim
16	31	9	3	II	Sim
17	31	10	4	II	Sim
18	46	13	11	III	Sim
19	21	1	2	II	Sim
20	39	17	8	III	Sim
21	15	53	46	I	Sim
22	9	42	39	I	Não
23	75	30	22	III	Sim
24	54	18	16	III	Não
25	35	12	5	II	Sim
26	18	58	50	II	Sim

(a) Construa o gráfico de dispersão entre as variáveis *Diag* e *Recup*. Identifique no mesmo gráfico os pacientes quanto à presença ou não de coágulos.
(b) Calcule o coeficiente amostral de correlação e a reta de regressão entre *Diag* e *Recup* para os dois grupos de pacientes, com coágulos presentes e ausentes. O relacionamento entre as variáveis depende da ocorrência de coágulos?