

Capítulo 4- Dados bivariados(242)

Introdução (pág. 243)

Aqui pretendemos analisar a relação entre duas variáveis.

Por exemplo, a relação entre o peso e a altura de uma pessoa.

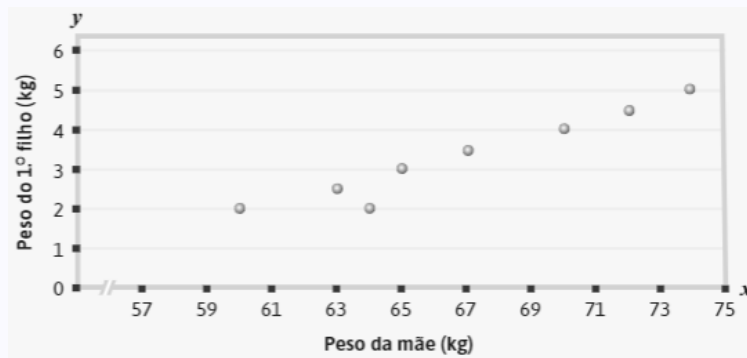
Aqui recolhemos pares de dados, que designamos por **dados bivariados**.

Sugestão: consulte o vídeo da página 243.

4.1-Diagrama de dispersão.(pág. 243)

Aqui usamos diagramas de dispersão ou gráficos de correlação.

Exemplo:



A este tipo de gráfico chamamos **gráfico de correlação** ou **diagrama de dispersão**.

O conjunto dos pontos num gráfico de correlação designa-se por **nuvem de pontos**.

Exemplo 1(pág. 243)

Nota: no exemplo acima, existe correlação entre as variáveis.

Diagrama de dispersão ou **gráfico de correlação** é um gráfico de pontos em que as coordenadas de cada ponto são os valores das duas variáveis em estudo.

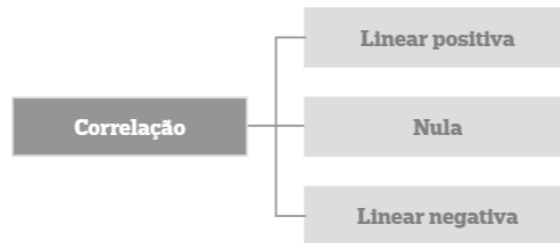
Check (pág. 244).

Exemplo 2(pág. 245)

A **correlação** diz-se **linear** se a nuvem de pontos se distribuir ao longo de uma linha reta, a reta de regressão.(245)

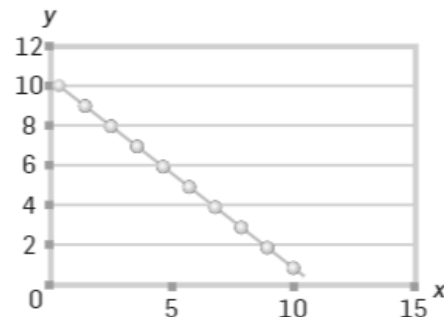
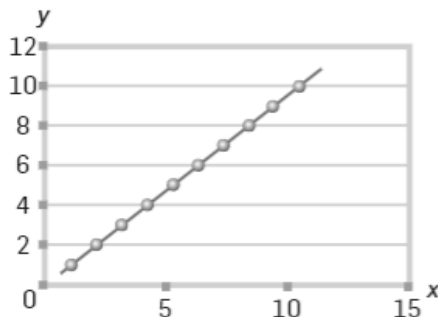
Correlação **linear positiva**- à medida que os valores de uma variável aumentam, os valores correspondentes da outra variável também aumentam.

Correlação **linear negativa**- à medida que os valores de uma variável aumentam, os valores correspondentes da outra variável diminuem.



(pág. 246)

Relação linear perfeita(246)



São casos em que os pontos da nuvem se encontram todos sobre a reta de regressão.

A correlação diz-se, nestes casos:

- ▶ **linear total (ou perfeita) positiva**, se se trata do gráfico da esquerda;
- ▶ **linear total (ou perfeita) negativa**, se se trata do gráfico da direita.

Exemplo 3(pág.246)

4.2-Reta de regressão. (pág. 247)

Sugestão: consulte os 3 vídeos da página 247.

Reta de regressão é a reta que melhor se ajusta aos pontos de um diagrama de dispersão. (247)

Nota: Esta reta serve sobretudo para descrever a relação entre as variáveis ou para fazer estimativas para valores desconhecidos.

Importante: devemos evitar fazer previsões para valores de x que estejam fora do intervalo dos valores observados. Tais previsões podem nem fazer sentido!...

Exemplo 4 (pág. 247). (Observe na calculadora gráfica...) Instruções em:

<https://auladigital.leya.com/resources/qrcodes/multiresource/QR2980/index.html>

Centro de gravidade da nuvem de pontos é o ponto cujas coordenadas são as médias das distribuições em análise. Calculadora: 2var...

Exemplo 5 (pág. 249).

Nota importante: Na equação $Y=ax+b$, muitas vezes dão-nos um valor de x e pedem uma estimativa para o valor de y .

Por exemplo, consideremos uma equação que resultou de uma recolha de dados onde o x variou entre 3 e 18, tendo obtido a equação da reta de regressão:

$$Y= 0.58x+8.08.$$

Para $x=14$, qual seria a estimativa para Y ?

Resposta: basta substituir: $Y= 0.58 \times 14 + 8.08 \approx 16.2$

Check (pág. 249). (Evitamos calcular o x e função do y . Não resolver o .6)

Sugestão: consulte os vídeos tutoriais da calculadora gráfica da página 249.

Sugestões para a calculadora gráfica:

Nota: Para obtermos a equação da reta de regressão usando a **calculadora gráfica**.

Sugestão1: consulte os vídeos tutoriais da página 249.

Sugestão2: Pode usar os tutoriais indicados no código QR apresentado no verso da capa do manual.

Sugestão 3:

<https://auladigital.leya.com/resources/qrcodes/multiresource/QR2980/index.html>

ou:

Casio e Texas:

https://pedronoia.net/_private/Calculadoras/calc10EstatBiv2009.pdf

TI-Nspire:

<https://pedronoia.net/nspire.pdf>

Numworks:

<https://www.numworks.com/pt/professores/tutoriais/regressao/>

Exemplo 6 (pág. 250).

Nota:(Pág. 250) O exemplo anterior mostra que nunca devemos considerar a reta de regressão independente da nuvem de pontos, pois a existência de **pontos anómalos** pode induzir em erro o cálculo de estatísticas e de estimativas. A eliminação destes pontos conduz a estatísticas e estimativas mais reais.

4.3 Coeficiente de correlação linear. (251).

Sugestão: vídeo da página 251.

O coeficiente de correlação linear mede o grau de associação linear entre duas variáveis. Representa-se por r e varia entre -1 e 1 .

- ▶ $r = 1$, a correlação diz-se **total** (ou **perfeita**) **positiva**;
- ▶ $r = 0$, a correlação diz-se **nula**: não há correlação linear;
- ▶ $r = -1$, a correlação diz-se **total** (ou **perfeita**) **negativa**.

Para os valores intermédios, a correlação é tanto mais forte quanto mais próximo o valor de r se encontrar de 1 ou de -1 , enfraquecendo à medida que se aproxima de zero.

Nota: o facto de o coeficiente de correlação ser zero, ou próximo de zero, não significa que as duas variáveis não tenham qualquer relação. Podem ter uma relação não linear, como podemos observar no exemplo 7:

Exemplo 7(pág. 252).

Nota: o facto de o coeficiente de correlação ser elevado não significa que as duas variáveis tenham uma relação de causa-efeito, como podemos ver no exemplo 8:

Exemplo 8(pág. 252).

Exemplo 9(pág. 253)Utilização da calculadora gráfica.

Utilização da calculadora gráfica.

Os procedimentos são os mesmos que vimos para a reta de regressão. O coeficiente de correlação é " r " e surge juntamente com a reta de regressão.

Apenas nas calculadoras **Texas**, pode ser necessário pedir explicitamente o valor de " r ", fazendo: //Catalog//Diagnostic On// Enter//. Ao pedir a correlação linear, já aparecerá o valor de " r ".

Check (pág. 253). (Em .1, .3, e .4, pode usar a Calculadora gráfica. 5. Tente resolver o.6 indicando os cálculos fazendo y -nº de netos e x - idade da avó)

Sugestão: Consulte o vídeo tutorial da sua calculadora indicado na página 253.

Exemplo 10 (pág. 254)

Nota:(Pág. 254) O exemplo 10 a existência de **pontos anómalos(outliers)** pode condicionar o valor do coeficiente de correlação linear. Por vezes, a eliminação destes pontos conduz a uma interpretação mais realista dos dados.

4.4 Tabelas de contingência. (255).

Quando pelo menos uma das variáveis estatísticas em estudo é do tipo qualitativo, recorre-se à representação dos dados em **tabelas de contingência**.

Exemplo 11(pág. 255).

Prova dos 9 (pág. 256 a 259). (Resolva na aula tudo exceto 2.1 e 6.4, que pode deixar para treinar.)

Sugestão: pode consultar o vídeo tutorial da calculadora correspondente ao exercício 2.3 da página 256

Pensamento computacional-Python (pág. 260)

Programa para determinar as coordenadas do centro de gravidade.

Nota: No programa da página 60, definimos uma lista de pontos (“pontos”) e usamos a função “**sum**(x for x, _ in pontos)” para somarmos todas as coordenadas x. Fazemos o mesmo para o y. Depois é só dividir pelo comprimento (“**len**”) da lista.

Obtemos as coordenadas do centro de gravidade.

Sugestão 1: Consulte os vídeos tutoriais de Python.

Sugestão 2: aprenda um pouco sobre listas: <https://pedronoia.pt/python/pyl4.htm>

Exercícios de Aplicação (pág. 262) 1 a 15.

Síntese (pág. 261)- Dados Bivariados.

Preparado?(pág. 267)

A Pensar no Exame(pág. 269)