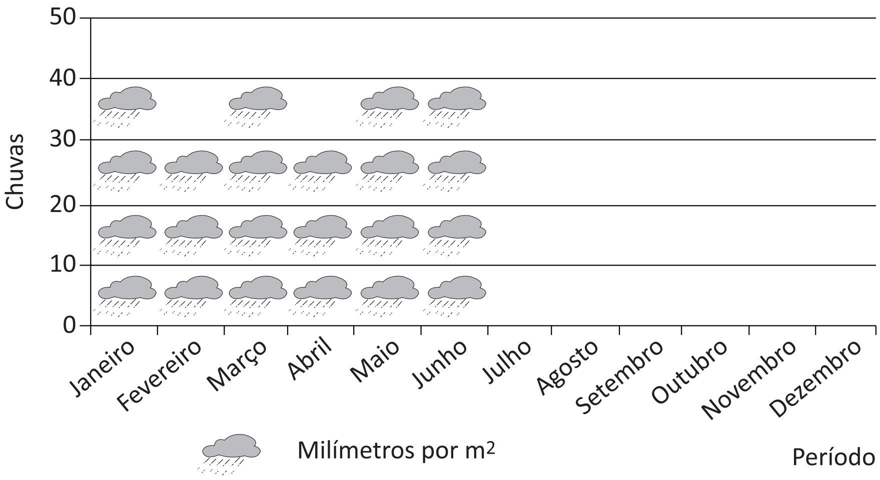
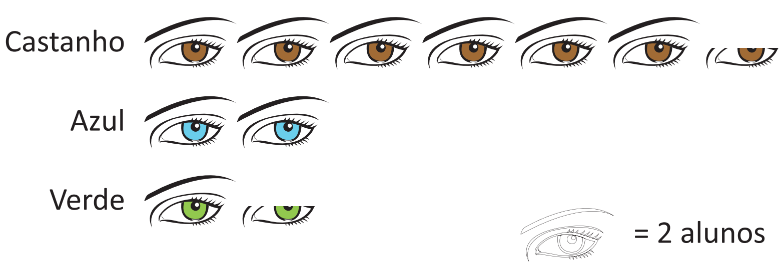
Pictogramas

É uma representação gráfica que usa símbolos sugestivos das variáveis em estudo.

1.



2.



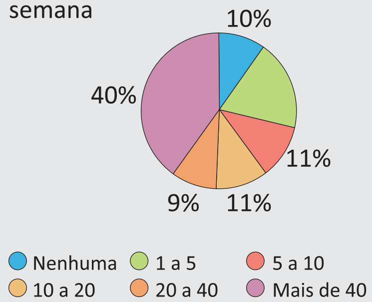
Histogramas

É uma representação gráfica de dados agrupados em classes. O gráfico é formado por uma sucessão de retângulos adjacentes, tendo cada um por base um intervalo de classe e por altura a frequência relativa ou absoluta. Deste modo, a área total coberta pelo histograma é igual a 1 caso seja construído com frequências relativas, ou igual a dimensão da amostra caso seja construído com as frequências absolutas.

Tarefa 50

Num inquérito a 410 pessoas sobre a quantidade de SMS que envia normalmente por semana, obtiveram-se os resultados registados no seguinte gráfico:

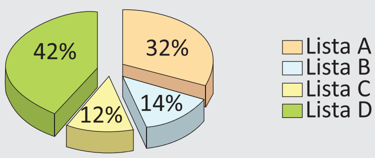
Nº de mensagens enviadas por semana



Quantas pessoas enviam cinco ou menos mensagens por semana?

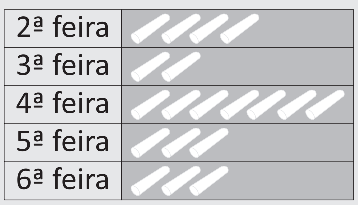
Tarefa 51

Elabora a tabela de frequências absolutas e relativas correspondentes aos resultados da sondagem registados no seguinte diagrama circular:



Tarefa 52

Numa escola registaram-se o número de paus de giz gastos por dia numa semana de aulas. Os dados obtidos estão registados no seguinte pictograma.



 = 2 paus de giz

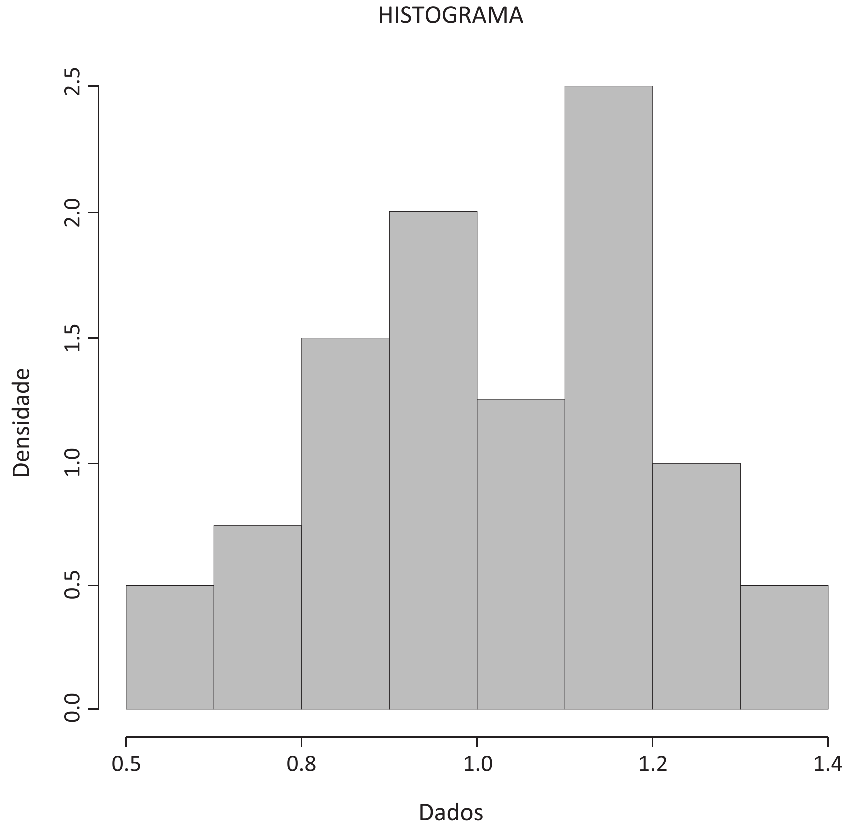
a) Quantos paus de giz são gastos numa semana?

b) Qual é o dia em que se gasta um maior número de paus de giz?

c) Indica a percentagem de paus de giz gastos até 3ª feira (inclusive).

d) Elabora uma tabela de frequências relativa ao pictograma fornecido.

Exemplo:

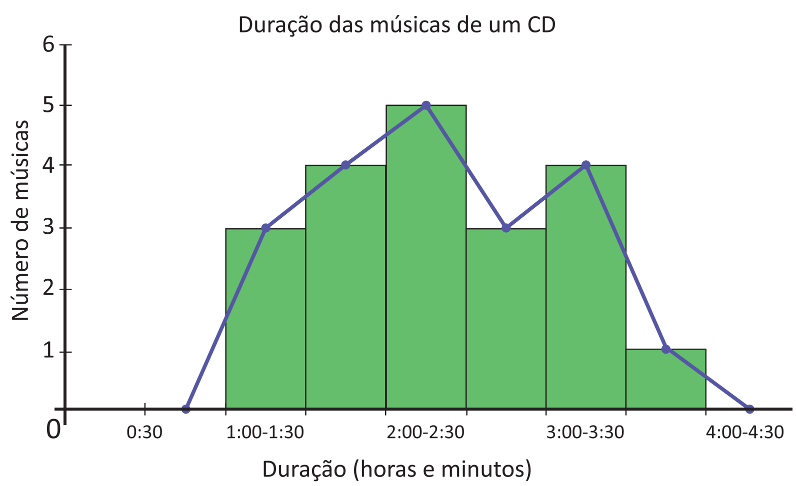


Polígono de frequências

Uma outra forma de representar graficamente dados agrupados em classes é a construção de um polígono de frequências.

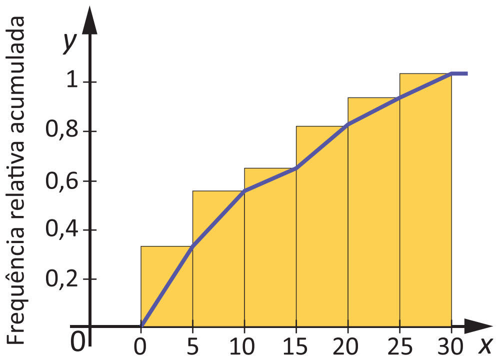
Para obter o polígono de frequências seguimos as seguintes etapas:

- Marcam-se os pontos médios os extremos superiores da barra correspondente a cada classe.
- Unem-se os pontos médios determinados na etapa anterior.
- Unem-se os pontos médios da primeira e da última classe a dois pontos do eixo horizontal, pontos médios de duas classes fictícias.



No caso das variáveis contínuas, para representar a função cumulativa podemos proceder da seguinte forma:

Antes do limite inferior da 1ª classe a frequência acumulada é nula. Admitindo que em cada classe a frequência se distribui uniformemente e tendo em conta que a frequência acumulada no limite inferior de cada classe é a frequência acumulada da classe anterior, unimos o ponto correspondente à frequência acumulada no início de cada classe com o ponto correspondente à frequência acumulada no início da classe seguinte.



Medidas de localização

Num estudo estatístico podemos estabelecer as medidas de **tendência central ou da localização: Média, Moda e Mediana.**

Média

Esta medida de localização (média aritmética) é a mais utilizada.

Representa-se por $\bar{x}$ e determina-se da seguinte forma :

- Somam-se todos os dados;
- Divide-se a soma anterior pelo total dos dados observados.

Exemplo:

Registaram-se as classificações de um aluno:

11; 11; 12; 12; 12; 14

Para determinar a média fazemos: $\frac{11+11+12+12+12+14}{6}=12$

Mas também podemos fazer de uma forma mais cómoda:

$$\frac{2\times11+3\times12+14}{6}=12$$

Tarefa 53

Num teste de 79 perguntas aplicado a 600 pessoas, o número de respostas certas está representado na tabela seguinte:

Nº de respostas corretas	Nº de pessoas
$[0,10[$	40
$[10,20[$	60
$[20,30[$	75
$[30,40[$	90
$[40,50[$	105
$[50,60[$	85
$[60,70[$	80
$[70,80[$	85

a) Constrói um histograma e um polígono de frequências absolutas da distribuição.

b) Constrói um histograma de frequências relativas acumuladas e o respetivo polígono.

Formalizando

Para uma amostra formada por "n" elementos chama-se média amostral ou simplesmente média ao quociente entre a soma de todos os dados e a dimensão da amostra.

Se $x_1, x_2, x_3, \cdots x_n$ são os "n" dados que constituem amostra

tem-se $\bar{x} = \dfrac{x_1 + x_2 + \cdots + x_n}{n} = \dfrac{\sum_{i=1}^{n} x_i}{n}$

Em geral, dada uma amostra ou população de dimensão "n", onde cada variável x_i toma "k" valores diferentes, sendo f_i fr_i a frequência absoluta e relativa respetivamente ao valor x_i, tem-se:

$$\bar{x} = \frac{f_1 \times x_i + f_2 \times x_2 + \cdots + f_k \times x_k}{n} = \frac{\sum_{i=1}^{k} f_i \times x_i}{n}$$

$$\bar{x} = fr_1 \times x_1 + fr_2 \times x_2 + \cdots fr_k \times x_k = \sum_{i=1}^{k} fr_i \times x_k$$

Tarefa 54

Um serviço telefónico de apoio ao cliente, recolheu informação acerca do tempo de espera, em minutos, até a chamada ser encaminhada para um assistente. A seguinte tabela regista os resultados obtidos.

Tempo de atendimento (min)	Nº de pessoas
$[0,5[$	54
$[5,10[$	80
$[10,15[$	32
$[15,20[$	16
$[20,25[$	12
$[25,30[$	6

a) Representa a informação usando um histograma de frequências relativas.

b) Representa o polígono de frequências absolutas.

Exemplo:

Para efetuar um teste de qualidade, pesaram-se 100 pacotes de arroz de 250 gramas. Os resultados obtidos estão registados na seguinte tabela:

Classes	Nº de pacotes
$[247,248[$	11
$[248,249[$	25
$[249,250[$	32
$[250,251[$	19
$[251,252[$	13

Pretendemos determinar o peso médio dos pacotes observados.

Começamos por encontrar a marca da classe.

Classes	Marca da classe	Nº de pacotes
$[247,248[$	247,5	11
$[248,249[$	248,5	25
$[249,250[$	249,5	32
$[250,251[$	250,5	19
$[251,252[$	251,5	13

Seguidamente determina-se o produto entre a marca da classe e a frequência absoluta.

Classes	Marca da classe	Nº de pacotes	$f_i \times x_i$
$[247,248[$	247,5	11	2722,5
$[248,249[$	248,5	25	6212,5
$[249,250[$	249,5	32	7984
$[250,251[$	250,5	19	4759,5
$[251,252[$	251,5	13	3269,5
			$\sum_{i=1}^{5} f_i \times x_i = 24948$

Vamos agora determinar a média:

$$\overline{x} = \frac{\sum_{i=1}^{5} f_i \times x_i}{100} = 249,48$$

Propriedades da média

1. Dada uma distribuição estatística cuja média seja $\overline{x}$, se adicionarmos uma constante k a todos os dados observados , obtém-se uma nova distribuição com média:

$$\overline{y} = \overline{x} + k \ , \ k \in IR$$

Exemplo:

Num restaurante o valor pago por um cliente, em dólares, pelas refeições de 8 dias foi:

9 10 10,5 9,5 12 8,5 8 8,5

A média do preço pago pelo cliente foi 9,5 dólares.

Se em cada refeição tivesse consumido cerveja e sobremesa teria que ter pago mais 10 dólares. Teríamos então:

19 20 20,5 22 18,5 18 18,5

A média será de 19,5 dólares.

Repara então que se verifica a propriedade referida, a média final é a primeira adicionada de 10 dólares.

Tarefa 55

Na tabela seguinte estão registadas as temperaturas observadas na cidade de Díli, durante 24 dias de um determinado mês:

26	27	27	26	29	29	30
31	27	31	29	30	31	26
29	27	30	31	26	27	29
29	30	30				

Determina a média da temperatura durante esses 24 dias em Díli.

Tarefa 56

Os pesos, em quilogramas, de 24 pessoas são:

68,5	34,2	47,5	39,2	47,3
79,2	46,5	58,3	62,5	58,7
80	63,4	58,6	50,2	60,5
70,8	30,5	42,7	59,4	39,3
48,6	56,8	72	60	

a) Agrupa os dados em intervalos de amplitude 10.

b) Determina a média do peso das pessoas inquiridas.

Tarefa 57

Uma equipa de futebol participou num torneio internacional, realizando 8 jogos. Relativamente a cada jogador, registou-se o número de jogos do torneio em que participou, obtiveram-se os seguintes resultados:

8 0 1 2 2 8 6 0 5
8 4 8 4 8 2 2 8 1
0 4 6 0 5 3

a) Quantos jogadores fazem parte da equipa?

b) Determina em quantos jogos participou em média cada jogador?

2. Dada uma distribuição estatística cuja média é $\overline{x}$, ao multiplicar todos os dados por uma constante k, obtém-se uma nova distribuição em que a média é igual a $\overline{y} = \overline{k}\,\overline{x}\,, k \in IR\{0\}$.

Exemplo:

Numa empresa ao fim de cada ano se distribuem prémios de produtividade pelos funcionários, os resultados estão registados na seguinte tabela:

Prémio (em dólares)	Nº Funcionários
50	1
80	6
110	3
140	2

A média dos valores obtidos é 95 dólares.

No ano seguinte os prémios duplicaram, mantendo-se as mesmas frequências absolutas.

Prémio (em dólares)	Nº Funcionários
100	1
160	6
220	3
280	2

Neste caso a média é 190 dólares, observa que a média é dada por 95 X 2 = 190.

Moda

Quando nos referimos a moda não faz sentido falar no seu cálculo, não há nada a calcular, procuramos o valor da variável quantitativa ou qualitativa que ocorre mais vezes , isto é que tem maior frequência absoluta . A esse valor damos o nome de moda.

Formalizando

Sendo $x_1, x_2, \cdots, x_k$ os k valores distintos de uma variável estatística, chama-se moda ao valor que tem maior frequência absoluta e representa-se M_0.

Exemplos:

1. Na seguinte distribuição: 3; 4; 3; 6; 2; 3; 6; 5; 3. A moda é 3.

2. Determinação da moda nas seguintes distribuições:

   3 6 3 6 2 1 — A moda é 6.

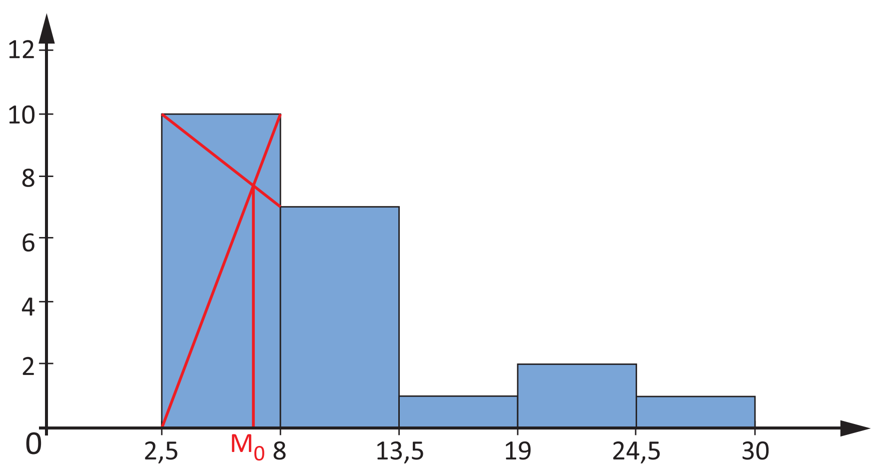
   4 3 5 7 1 2 — Distribuição amodal (não existe moda).

No caso de os dados estarem agrupados em classes, á classe com maior frequência absoluta dá-se o nome de **Classe Modal**.

Se representamos os dados agrupados em classes por um histograma é possível fazer uma **estimativa geométrica para a Moda**.

O procedimento a seguir é o seguinte:

- Representam-se os dados num histograma.
- Determina-se a classe modal.
- Unem-se os vértices superiores do retângulo da classe modal com os vértices das classes contíguas.
- A perpendicular traçada desde o ponto de interseção das linhas obtidas anteriormente até ao eixo das abcissas, determina a localização gráfica da moda.



Mediana

A mediana é o dado que ocupa a posição central da distribuição, quando ordenada, bem seja ordem crescente ou decrescente.

Tarefa 58

Uma empresa exportadora registou as encomendas (em toneladas) dos seus clientes durante um semestre.

4 3 5 6 7 8 5 3 7 6 5 8

a) Qual é a média das encomendas?

b) No semestre seguinte, as encomendas, duplicaram, tendo registado os seguintes valores:

   8 6 10 12 14 16 10 6 14 12 10 16

   b1)Qual é a média das encomendas?

   b2)Qual é a relação existente entre esta média e a das encomendas do 1º semestre?

c) Determina a moda das encomendas.

Exemplo:

Consideramos as classificações, arredondadas as unidades, dos testes de um aluno ao longo do ano letivo.

11 10 9 13 10 11 12

Ordenamos os dados :

9 10 10 11 11 12 13

O centro da distribuição é 11.

A partir de um conjunto de dados ordenados dá-se o nome de Mediana ao valor (pertencente ou não a esse conjunto de dados) que o divide ao meio, ou seja 50% dos elementos da amostra são inferiores ou iguais à mediana e 50% são superiores ou iguais à mediana. A mediana representa-se por:

$$M_e \quad ou \quad \tilde{x}$$

Para determinar a mediana de uma amostra de "n" elementos primeiro ordenam-se os elementos da amostra.

Se n é impar:

A mediana é o elemento que ocupa a posição $\frac{n+1}{2}$.

Ou seja a mediana é: $\tilde{x} = x_{\frac{n+1}{2}}$.

Se n é par:

A mediana é a semissoma dos elementos que ocupam as posições

$\frac{n}{2}$ e $\frac{n}{2}+1$.

Ou seja a mediana é: $\tilde{x} = \dfrac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$.

Tarefa 59

Considera a seguinte distribuição de idades:

x_i	f_i
26	**6**
28	**7**
30	**4**
32	**3**

a) Determina a média da distribuição dada.

b) Calcula a média se multiplicarmos cada idade da tabela por 3.

c) Indica a média depois de dividir todos os valores da distribuição por 2.

Exemplo:

Registou-se o número de crias nascidas por ninhada de um casal de coelhos:

Nº de crias	Frequência Absoluta
2	3
3	24
4	30
5	13
5	7
Total	77

Tarefa 60

A partir de uma amostra de 75 pilhas, obtiveram-se os seguintes dados referentes ao tempo de duração em horas.

Duração (horas)	Nº de pilhas
$[25,30[$	3
$[30,35[$	5
$[35,40[$	21
$[40,45[$	28
$[45,50[$	12
$[50,55[$	6

a) Completa a tabela com a frequência relativa.

b) Indica a classe modal.

c) Localiza graficamente a moda.

O número total de crias é 77 (ímpar), a mediana é o valor que ocupa a posição $\frac{77+1}{2}=39$, depois de ter os dados ordenados, acrescenta-se à tabela anterior as frequências absolutas acumuladas:

Nº de crias	f_i	F_i
2	3	3
3	24	27
4	30	57
5	13	70
5	7	77
Total	77	

Observando a coluna das frequências absolutas podemos concluir:

Os elementos x_1 x_2 x_3 são todos iguais a 2.

Do x_4 até x_{27} são todos iguais a 3.

E assim sucessivamente, a mediana é $\tilde{x}=x_{39}=4$.

Em 50% das ninhadas nasceram quatro crias ou menos.

Exemplo:

Registou-se o número de crias nascidas por ninhada de um outro casal de coelhos.

Tarefa 61

Num concerto musical esteve presente um conjunto de pessoas com a seguinte distribuição de idades:

Idade (anos)	Nº pessoas
[14,21[	10
[21,28[	8
[28,35[	6
[35,42[	3
[42,49[	3
[49,56[	7
[56,63[	10
[63,70[	3

a) Constrói a tabela de frequências desta distribuição.

b) Quantas pessoas com menos de 49 anos assistiram ao concerto?

c) Calcula a média desta distribuição.

d) Determina, caso exista, a classe modal.

Nº de crias	Frequência Absoluta
2	3
3	25
4	20
5	5
6	3
Total	56

Agora o número de dados é par então a mediana é o valor que se encontra na posição obtida pela semissoma dos valores.

$$x_{\frac{n}{2}} = x_{28} \text{ e } x_{\frac{n}{2}+1} = x_{29}$$

Primeiro temos que ordenar os dados:

Nº de crias	f_i	F_i
2	3	3
3	25	28
4	20	48
5	5	53
6	3	56
Total	56	

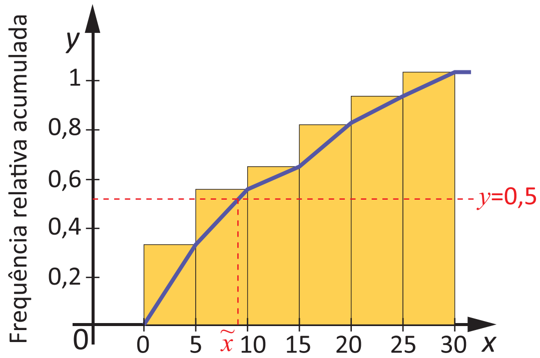
A mediana é: $\tilde{x} = \dfrac{x_{28} + x_{29}}{2} = \dfrac{3+4}{2} = 3{,}5$.

Em 50% das ninhadas nasceram três crias ou menos.

No caso dos dados estarem **agrupados em classes**, identificamos a **classe mediana** e só localizaremos graficamente a mediana.

A obtenção geométrica de um valor aproximado da mediana, pode realizar-se construindo o gráfico da função cumulativa e determinando a interseção do mesmo com a reta horizontal y=0,5.

A abcissa do ponto de interseção é o valor aproximado da mediana.



Quartis

Em muitos estudos estatísticos, com populações ou amostras de grande dimensão, é interessante conhecer a maior ou menor concentração de dados ao longo do intervalo de variação dos seus valores. Ou seja, entre o maior valor que a distribuição toma e o seu menor valor.

As estatísticas que dividem o conjunto de dados, depois de ordenados, em quatro partes, chamam-se quartis.

O **1º Quartil** representa-se por Q_1 e é o valor da variável, tal que 25% dos dados são menores ou iguais e 75% são maiores ou iguais.

O **2º quartil** Q_2 é o valor da variável que coincide com a mediana.

O **3º Quartil** representa-se por Q_3 e é o valor da variável tal que 75% dos dados são menores ou iguais e 25% são maiores ou iguais.

Exemplo:

O número de mensagens SMS recebidas num telemóvel durante 11 dias está registado a seguir:

12 12 6 6 5 8 7 10 9 10 11

Primeiro ordenamos os elementos da distribuição:

5 6 **6** 7 8 **9** 10 10 **11** 12 12

Q_1 Q_2 Q_3

Como no caso da mediana existem fórmulas para determinar os quartis. Essas fórmulas são diferentes caso a distribuição tenha um número de elementos par ou ímpar.

Tarefa 62

Os dados seguintes referem-se ao número de passageiros transportados pelos autocarros A e B em cada um dos percursos efetuados durante um dia.

Autocarro A:

30 34 78 46 38 50 44
61 65 84 44 28

Autocarro B:

41 39 83 48 60 72 54
90 33 47

Determina a mediana do número de passageiros transportados por cada um dos autocarros.

Tarefa 63

Determina a mediana da seguinte distribuição.

x_i	f_i
0	10
1	15
2	25
3	20
4	15
5	15

Tarefa 64

Registou-se o número de horas que uma pessoa gastou a ver televisão durante 24 semanas.

7 9 9 10 11 11 3 4 4 5

5 6 6 11 11 12 12 14 13

14 16 15 15

a) Justifica que a mediana é 11.

b) A grupa os dados em classes e identifica a classe mediana.

c) Localiza geometricamente o valor da mediana.

Tarefa 65

O número de mensagens SMS recebidas em 17 dias consecutivos foram as indicadas a seguir.

13 14 58 10 21 19 16

15 25 20 19 37 36 34

32 32 25

Indica os quartis da distribuição.

Para n par:

Localização	Quartil
$k=\frac{n+2}{4}$	$Q_1 = x_k$
$k=\frac{n}{2}$	$Q_2 = \tilde{x}$
$k=\frac{3n+2}{4}$	$Q_3 = x_k$

Para n ímpar:

Localização	Quartil
$k=\frac{n+1}{4}$	$Q_1 = x_k$
$k=\frac{n+1}{2}$	$Q_2 = \tilde{x}$
$k=3\times\frac{n+1}{4}$	$Q_3 = x_k$

Nota

Quando estas fórmulas não conduzem a números inteiros, calcula--se a média aritmética dos valores que ocupam as posições dadas pelos números inteiros mais próximos do valor obtido.

Exemplo:

Se n = 17

Para de terminar o 1º Quartil: $k=\frac{17+1}{4}=4,5$

Então $Q_1=\frac{x_4+x_5}{4}$

Para de terminar o 3º Quartil: $k=3\times\frac{17+1}{4}=13,5$

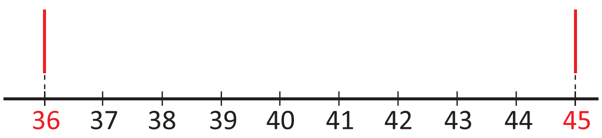
Então $Q_3=\frac{x_{13}+x_{14}}{4}$

Diagrama de extremos e quartis

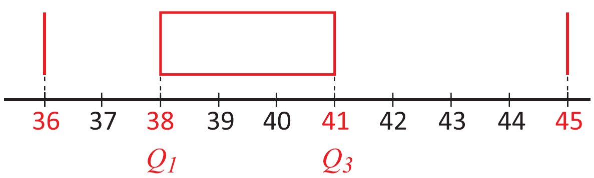
O diagrama de extremos e quartis é uma forma esquemática de representar os valores máximo e mínimo de uma distribuição assim como a mediana e os quartis.

Para construir o **diagrama de extremos e quartis** procedemos da seguinte forma:

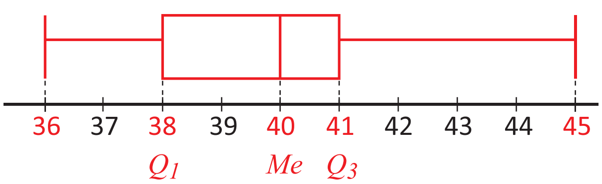
- Determinar na amostra os extremos da distribuição, a mediana e os quartis;
- Traçar um eixo graduado para assinalar os valores determinados anteriormente;
- Traçar dois segmentos de reta correspondentes aos extremos da distribuição;



- Construir um retângulo em que dois dos lados correspondam ao 1º e 3º quartil;



- Dividir o retângulo anterior em duas partes usando um segmento de reta que corresponde ao valor da mediana. Finalmente unir os valores extremos do retângulo como se indica na figura seguinte.



Observa que a forma como o diagrama se constrói permite retirar imediatamente informação importante sobre os dados, nomeadamente sobre o centro da amostra (mediana) e variabilidade (maior ou menor concentração dos dados).

Tarefa 66

Registou-se nas tabelas seguintes o número de horas estudo diário dos alunos de duas turmas.

Turma 1

0	**3**
1	8
2	7
3	6
4	3
5	3

Turma 2

0	**3**
1	8
2	7
3	6
4	3
5	3

a) Determina o tempo diário médio dos alunos da primeira turma.

b) Determina os quartis da distribuição de valores da primeira turma.

c) Indica um intervalo de tempo de estudo a que correspondam 25% destes alunos.

d) Indica qual das turmas tem uma média mais elevada em relação as horas de estudo.

e) Junta os dados das duas turmas e determina os quartis desta nova distribuição.

Medidas de Dispersão

Observa o seguinte estudo estatístico sobre os resultados obtidos por duas turmas num exame de português:

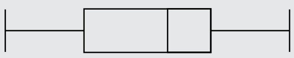
	Média	Moda	Mediana
Turma C	9,8	12	9
Turma D	9,8	12	9

Nota

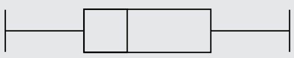
A utilização de um diagrama de extremos e quartis facilita a identificação de simetria e assimetrias de uma amostra.

Dados Simétricos:

Dados Enviesados à esquerda:



Dados enviesados à direita:



Tarefa 67

Constrói um diagrama de extremos e quartis referente aos dados seguintes.

x_i	f_i
14	5
15	14
16	6
17	2
18	1
19	1

Nota

A letra grega σ :

lê-se sigma minúsculo.

Observemos agora as classificações de cada turma:

Turma C	20	19	18	12	12	12	12	9	8	8	7	4	3	2	1
Turma D	13	13	12	12	12	12	10	9	9	8	8	8	7	7	7

Podemos observar que os resultados da turma D são mais homogéneos que os da turma C. O desempenho das duas turmas é diferente.

É importante analisar, não só as medidas de tendência central mas também a forma de dispersão dos dados. Os dois tipos de medidas em conjunto fornecem uma melhor visualização da distribuição em estudo.

Amplitude Total e amplitude Interquartis

Num conjunto de dados chama-se **amplitude total** à diferença entre o valor máximo e o valor mínimo da variável.

Se os dados estão **agrupados em classes**, a amplitude total é a diferença entre o limite superior da ultima classe e o limite inferior da primeira classe

A **amplitude Interquartis** é a diferença entre o terceiro quartil e o primeiro.

A amplitude interquartis evidencia a variabilidade ou dispersão dos dados em relação à mediana.

Variância e Desvio Padrão

Como a medida de localização central mais usada é a media, é importante estabelecer uma medida de dispersão relativa a ela, o desvio-padrão.

A variância que se representa por $s^2 \quad ou \quad \sigma^2$, é a medida que se obtém dividindo a soma dos quadrados dos desvios das observações relativamente à média, pelo número de observações:

$$\sigma^2 = \frac{\sum_{i=1}^{n} \left(x_i - \bar{x}\right)^2}{n}$$

Quando os dados são discretos e organizados numa tabela, a determinação da variância toma a seguinte forma:

$$\sigma^2 = \frac{\sum_{i=1}^{n} f_i \left(x_i - \bar{x}\right)^2}{n}$$

Exemplo:

O número de golos obtidos nas primeiras 20 jornadas de um clube foi registado na seguinte tabela.

Nº de golos	0	1	2	3	4	5
Nº de jogos	2	5	4	5	3	1

Queremos determinar a variância.

Primeiro determinamos a média: $\overline{x} = 2,25$

Agora construímos uma tabela com os valores envolvidos no calculo da variância o que facilitará a organização dos dados.

x_i	f_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$	$f_i(x_i - \overline{x})^2$
0	**2**	**-2,25**	**5,0625**	**10,125**
1	**5**	**-1,25**	**1,5625**	**7,8125**
2	**4**	**-0,25**	**0,0625**	**0,25**
3	**5**	**0,75**	**0,5625**	**2,8125**
4	**3**	**1,75**	**3,0625**	**9,1875**
5	**1**	**2,75**	**7,5625**	**7,5625**
Total	**20**			$\sum_{i=1}^{5} f_i(x_i - \overline{x})^2 = 37,75$

Temos então que a variância é dada por:

$$\sigma^2 = \sum_{i=1}^{5} \frac{f_i\left(x_i - \overline{x}\right)^2}{20} = 1,89 \quad (2c.d.)$$

Desvio Padrão

A variância pode ser utilizada como medida comparativa entre diferentes distribuições, quanto maior forem os quadrados dos desvios em relação à média, maior será a variância. Não é, no entanto, reveladora em relação aos dados da distribuição, pois a ordem de grandeza é diferente da dos desvios.

Por este motivo, em lugar de usar a variância como medida de dispersão relativa à média, usamos a sua raiz quadrada a que chamamos desvio padrão.

Tarefa 68

Num controlo de qualidade a um fio elétrico produzido por uma empresa, realizou-se a medida da longitude do fio e registaram-se os resultados:

10,4 10,3 9,8 10,2 10
10,2 10,1 9,8 9,9 10
10,2 9,7

a) Qual foi a amplitude das medições?

b) Determina a variância da distribuição.

Nota

Quando os dados são contínuos, para determinar a variância procede-se da mesma forma só que em lugar de usar os valores da variável, deveremos usar a marca da classe.

Tarefa 69

Os dados que se seguem correspondem ao comprimento, em cm, de uma amostra de peixes.

29,9	40,2	37,8	19,7	30,0
29,7	19,4	39,2	24,7	20,4
19,1	34,7	33,5	18,3	19,4
27,3	38,2	16,2	36,8	33,1
41,4	13,6	32,2	24,3	19,1
37,4	23,6	33,3	31,6	20,1
17,2	13,3	37,7	12,6	39,6
24,6	18,6	18,0	33,7	38,2

a) Agrupa os dados em classes.

b) Determina a variância.

O desvio padrão de uma distribuição é a medida que se obtém calculando a raiz quadrada da variância.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} f_i\left(x_i - \bar{x}\right)^2}{n}}$$

Exemplo:

A seguinte distribuição refere-se as alturas de 100 árvores de uma amostra aleatória.

Altura (m)	$[0; 0,5[$	$[0,5; 1[$	$[1; 1,5[$	$[1; 2[$
Frequência	18	30	40	12

Queremos determinar o desvio padrão.

Primeiro determinamos a média: $\bar{x} = 0,98$

Classes	x_i	f_i	$x_i - \bar{x}$	$\left(x_i - \bar{x}\right)^2$	$f_i\left(x_i - \bar{x}\right)^2$
$[0; 0,5[$	0,25	18	-0,73	0,5329	9,5922
$[0,5; 1[$	0,75	30	-0,23	0,0529	1,587
$[1; 1,5[$	1,25	40	0,27	0,0729	2,916
$[1,5; 2[$	1,75	12	0,77	0,5929	7,1148
Total		100			$\sum_{i=1}^{4} f_i\left(x_i - \bar{x}\right)^2 = 21,21$

O desvio padrão é:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{4} f_i\left(x_i - \bar{x}\right)^2}{100}} = 0,21 \quad (2c.d.)$$

Propriedades

1. Se a todos os dados de uma distribuição com desvio padrão σ adicionarmos o mesmo valor k, o desvio padrão não se altera.
2. Se multiplicarmos todos os dados de uma distribuição com desvio padrão σ, por uma mesma constante K, obtemos uma nova distribuição com desvio padrão kσ.