REVISTA DE CIÊNCIA ELEMENTAR

Quartis

Maria Eugénia Graça Martins
Faculdade de Ciências da Universidade de Lisboa

CITAÇÃO
Martins, M. E. G.(2014)
Quartis,
Rev. Ciência Elem., V2(04):268
doi.org/10.24927/rce2014.268

EDITOR
José Ferreira Gomes,
Universidade do Porto

RECEBIDO EM
21 de julho de 2011

ACEITE EM
03 de outubro de 2011

PUBLICADO EM
31 de dezembro de 2014

rce.casadasciencias.org

Os quartis são medidas de localização que dividem a amostra (ou coleção) de dados de tipo quantitativo (ou qualitativo ordinal) ordenada, em quatro partes, cada uma com uma percentagem de dados aproximadamente igual. O 1º quartil ou quartil inferior, representado por Q_1 (ou $Q_{0,25}$) e o 3º quartil ou quartil superior, representado por Q_3 (ou $Q_{0,75}$) são medidas que localizam alguns pontos da distribuição dos dados de tal forma que aproximadamente 25% dos dados são inferiores ou iguais a Q_1, aproximadamente 25% dos dados são superiores ou iguais a Q_3 e os restantes dados, aproximadamente 50%, situam-se entre Q_1 e Q_3. De um modo geral, quando nos referimos aos quartis, estamos a referir-nos ao 1º e 3º quartis, uma vez que o 2º quartil é designado por mediana.

Dada uma amostra, considera-se que o 1º quartil é o valor tal que pelo menos 25% dos dados são não maiores do que ele e pelo menos 75% dos dados são não menores do que ele e o 3º quartil é o valor tal que pelo menos 75% dos dados são não maiores do que ele e pelo menos 25% dos dados são não menores do que ele.

Há vários processos para calcular os quartis, que não conduzem necessariamente aos mesmos valores, mas a valores próximos, desde que a amostra tenha uma dimensão razoável, que é a situação de interesse em Estatística, em que se procura resumir a informação contida nos dados, através de algumas medidas.

Uma metodologia que pode ser seguida para obter os quartis é a seguinte:

- Ordenar os dados por ordem crescente e calcular a mediana;
- O 1.º quartil, Q_1, é a mediana dos dados que ficam para a esquerda da mediana;
- O 3.º quartil, Q_3, é a mediana dos dados que ficam para a direita da mediana.

Quando o número de dados é ímpar, o processo seguido, por exemplo, em MANN (1995) ou MOORE (1996), considera cada uma das partes em que fica dividida a amostra pela mediana, sem a incluir. Por exemplo,

10 12 13 16 **17** 19 20 23 25

$Q_1 = \frac{12+13}{2}$ Mediana $Q_3 = \frac{20+23}{2}$

Esta metodologia não é seguida por outros autores, como por exemplo DE VEAUX et al. (2004) que inclui a mediana nas duas partes. Assim, o 1º e 3º quartis viriam, no caso do exemplo anterior

10	12	13	16	**17**	19	20	23	25
		Q_1=13		Mediana		Q_3=20		

No caso do número de dados ser par, 2n, com n ÐN, não há ambiguidade na definição dos quartis que serão sempre as medianas de conjuntos de dados com n elementos.

Do mesmo modo que a mediana, também se podem obter os quartis a partir das tabelas de frequências com os dados agrupados.

À diferença entre o 3º quartil e o 1º quartil dá-se o nome de amplitude interquartil, que é uma medida da variabilidade ou dispersão (medidas de dispersão) entre os dados.

Os quartis são utilizados na construção do diagrama de extremos e quartis e do diagrama caixa de bigodes.

A característica populacional que corresponde à característica amostral quartil, também se denomina quartil e determina-se da mesma maneira (para uma população finita). Depende do contexto saber se nos estamos a referir à estatística - quartil amostral, ou ao parâmetro - quartil populacional.

Nota:

Sendo os quartis um caso especial dos quantis, podem-se utilizar as fórmulas aí consideradas para a sua obtenção. Um método mais elaborado para o cálculo dos quartis (Murteira, 2002), será descrito a seguir, por ser também o método utilizado na folha de Excel. Neste processo, começa por se considerar para a determinação do 1º e 3º quartis, respetivamente as "posições" $\frac{n+3}{4}$ e $\frac{3n+1}{4}$ (1). No que respeita o 1º quartil, Q_1, se $\frac{n+3}{4}$ for inteiro, então Q_1 será a observação de ordem $\frac{n+3}{4} = k$, na amostra ordenada. Se $\frac{n+3}{4}$ não for inteiro, seja $\frac{n+3}{4} = k + \varepsilon$, em que representamos por ε a parte decimal (igual a 0,25, 0,5 ou 0,75). Para obter Q_1 considera-se uma interpolação linear, fazendo uma média ponderada entre a observação de ordem k e a observação de ordem $(k+1)$:

Q1=observação de ordem k + ε (observação de ordem $(k+1)$ – observação de ordem k)

O raciocínio utilizado para obter o 3º quartil, Q_3, é idêntico ao descrito para a determinação do 1º quartil, considerando agora a posição $\frac{3n+1}{4}$ em vez de $\frac{3n+1}{4}$.

Adiantamos que quando a dimensão da amostra é ímpar, este processo conduz aos mesmos resultados que o considerado por DE VEAUX et al. (2004), em que se considera os quartis como as medianas de cada uma das partes em que fica dividida a amostra, incluindo a mediana em cada uma dessas partes.

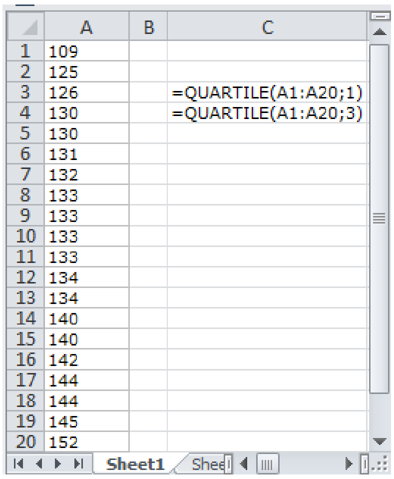
Considere-se, por exemplo, os seguintes dados correspondentes às alturas em cm, dos 20 alunos de uma turma (a amostra já se encontra ordenada):

109	125	126	130	130	131	132	133	133	133	133	134	134	140	140	142	144	144	145	152

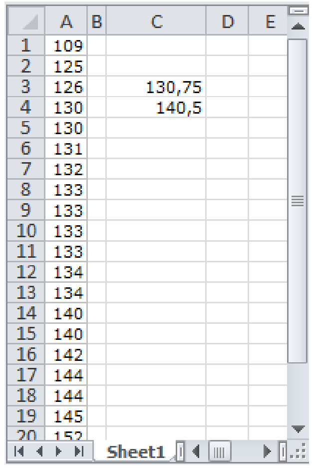
De acordo com a metodologia descrita anteriormente, o 1º quartil e o 3º quartil estão, respetivamente, nas "posições" 5,75 e 15,25 pelo que realizando as interpolações lineares, tem-se

Q1=130+0,75×(131−130)	e	=140+0,25×(142−140)
=130,75		=140,5

Na folha de Excel obtém-se:



	A	B	C
1	109		
2	125		
3	126		=QUARTILE(A1:A20;1)
4	130		=QUARTILE(A1:A20;3)
5	130		
6	131		
7	132		
8	133		
9	133		
10	133		
11	133		
12	134		
13	134		
14	140		
15	140		
16	142		
17	144		
18	144		
19	145		
20	152		



	A	B	C	D	E
1	109				
2	125				
3	126		130,75		
4	130		140,5		
5	130				
6	131				
7	132				
8	133				
9	133				
10	133				
11	133				
12	134				
13	134				
14	140				
15	140				
16	142				
17	144				
18	144				
19	145				
20	152				

(1) Dada uma amostra de dimensão n, ordenada, considera-se que a "posição" da mediana é definida por $\frac{n+1}{2}$. Quando n é ímpar, a mediana corresponde à observação de ordem $\frac{n+1}{2}$. Quando n é par é a semissoma (interpolação linear) das observações de ordem $\frac{n}{2}$ e $\frac{n}{2}+1$. Para definir o 1º quartil considera-se a "posição" definida por $\frac{posic\tilde{a}o\ da\ mediana+1}{2}=\frac{\frac{n+1}{2}+1}{2}=\frac{n+3}{4}$.

Para obter a "posição" do 3º quartil basta considerar $n-\frac{n+3}{4}+1=\frac{3n+1}{4}$. Da mesma forma que se calcula a mediana, se $\frac{n+3}{4}$ e $\frac{3n+1}{4}$ são inteiros, os quartis são as observações que correspondem a essas ordens. Caso contrário faz-se uma interpolação linear entre as observações correspondentes às ordens adjacentes.

Repare-se que estamos aqui a generalizar o conceito de "posição" para valores não necessariamente inteiros.