Percentil, Quartil e Boxplot

Gilberto Pereira Sassi

Universidade Federal Fluminense
Instituto de Matemática e Estatística
Departamento de Estatística

30 de maio de 2016

Percentil

O percentil de ordem $p \cdot 100$ $(0 < p < 1)$ é o valor que ocupa a posição $p \cdot (n+1)$ dos dados ordenados.

Notação:

- $q(p)$ é o percentil de ordem $p \cdot 100$;
- se $p = 0,25$, chamamos o percentil de primeiro quartil e usamos a notação matemática Q_1;
- se $p = 0,25$, chamamos o percentil de segundo quartil ou mediana e usamos a notação matemática Q_2 ou md
- se $p = 0,75$, chamamos o percentil de terceiro quartil e usamos a notação matemática Q_3

Exemplo

Considere a população nos 20 maiores municípios brasileiros em 1996 mostrados na Tabela 1. Calcule o primeiro, o segundo, o terceiro quartil, o intervalo interquartil.

Tabela 1: Vinte maiores cidades do Brasil em 1996 em 10.000 habitantes.

Município	População	
São Paulo(SP)	988,8	$x_{(20)}$
Rio de Janeiro(RJ)	556,9	$x_{(19)}$
Salvador(BA)	224,6	$x_{(18)}$
Belo Horizonte(MG)	210,9	$x_{(17)}$
Fortaleza(CE)	201,5	$x_{(16)}$
Brasília(DF)	187,7	$x_{(15)}$
Curitiba(PR)	151,6	$x_{(14)}$
Recife(PE)	135,8	$x_{(13)}$
Porto Alegre(RS)	129,8	$x_{(12)}$
Manaus(AM)	119,4	$x_{(11)}$
Belém(PA)	116,0	$x_{(10)}$
Goiânia(GO)	102,3	$x_{(9)}$
Guarulhos(SP)	101,8	$x_{(8)}$
Campinas(SP)	92,4	$x_{(7)}$
São Gonçalo(RJ)	84,7	$x_{(6)}$
Nova Iguaçu(RJ)	83,9	$x_{(5)}$
São Luis(MA)	80,2	$x_{(4)}$
Maceió(AL)	74,7	$x_{(3)}$
Duque de Caxias(RJ)	72,7	$x_{(2)}$
São Bernardo do Campo(SP)	68,4	$x_{(1)}$

Exemplo

1° **Quartil** posição: $0,25 \cdot (20+1) = 5,25$. Então

$$Q_1 = \frac{x_{(5)} + x_{(6)}}{2} = \frac{83,9 + 84,7}{2} = 84,3$$

2° **Quartil** posição: $0,5 \cdot (20+1) = 10,5$. Então,

$$Q_2 = \frac{x_{(10)} + x_{(11)}}{2} = \frac{116 + 119,4}{2} = 117,7$$

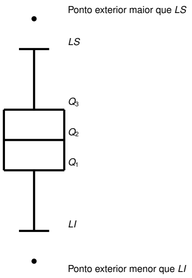
3° **Quartil** posição: $0,75 \cdot (20+1) = 15,75$. Então,

$$Q_3 = \frac{x_{(15)} + x_{(16)}}{2} = \frac{187,7 + 201,5}{2} = 194,6$$

Boxplot

A informação obtida pelo 1°, 2° e 3° quartil podem ser ser traduzidos graficamente num diagrama chamado de Boxplot.

Figura 1: Boxplot



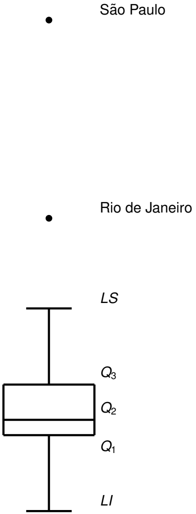
Boxplot – continuação

em que

- $LS = Q_3 + 1,5 \cdot (Q_3 - Q_1)$ é o limite superior;
- $LI = Q_1 - 1,5 \cdot (Q_3 - Q_1)$ é o limite inferior;
- Chamamos um valor entre *LI* e *LS* de valor adjacente;
- Chamamos um valor maior que *LS* ou menor que *LI* de ponto exterior, que pode possivelmente é um valor atípico.

Exemplo

Figura 2: Boxplot para as 20 maiores cidades brasileiras.



Exemplo

Considere as notas para as três turmas da Estatística Básica mostradas na Tabela. Calcule o primeiro, segundo, terceiro quartis e desenhe o Boxplot para as três turmas.

Tabela 2: Notas em ordem crescente para as três turmas de Estatística Básica.

Turma	Notas							
	$x_{(1)}$	$x_{(2)}$	$x_{(3)}$	$x_{(4)}$	$x_{(5)}$	$x_{(6)}$	$x_{(7)}$	$x_{(8)}$
A	4	5	5	6	6	7	7	8
B	1	2	4	6	6	9	10	10
C	0	6	7	7	7	7	8	

Exemplo

Turma A 1° **Quartil** posição $0,25 \cdot (8+1) = 2,25$. Então, $Q_1 = \dfrac{x_{(2)} + x_{(3)}}{2} = \dfrac{5+5}{2} = 5$;

2° **Quartil** posição $0,5 \cdot (8+1) = 4,5$. Então, $Q_2 = \dfrac{x_{(5)} + x_{(5)}}{2} = \dfrac{6+6}{2} = 6$;

3° **Quartil** posição $0,75 \cdot (8+1) = 6,75$. Então, $Q_3 = \dfrac{x_{(6)} + x_{(7)}}{2} = \dfrac{7+7}{2} = 7$.

Turma B 1° **Quartil** posição $0,25 \cdot (8+1) = 2,25$. Então, $Q_1 = \dfrac{x_{(2)} + x_{(3)}}{2} = \dfrac{2+4}{2} = 3$;

2° **Quartil** posição $0,5 \cdot (8+1) = 4,5$. Então, $Q_2 = \dfrac{x_{(5)} + x_{(5)}}{2} = \dfrac{6+6}{2} = 6$;

3° **Quartil** posição $0,75 \cdot (8+1) = 6,75$. Então, $Q_3 = \dfrac{x_{(6)} + x_{(7)}}{2} = \dfrac{9+10}{2} = 9,5$.

Turma C 1° **Quartil** posição $0,25 \cdot (8+1) = 2,25$. Então, $Q_1 = \dfrac{x_{(2)} + x_{(3)}}{2} = \dfrac{2+4}{2} = 3$;

2° **Quartil** posição $0,5 \cdot (8+1) = 4,5$. Então, $Q_2 = \dfrac{x_{(5)} + x_{(5)}}{2} = \dfrac{6+6}{2} = 6$;

3° **Quartil** posição $0,75 \cdot (8+1) = 6,75$. Então, $Q_3 = \dfrac{x_{(6)} + x_{(7)}}{2} = \dfrac{9+10}{2} = 9,5$.

Exemplo

Figura 3: Turma A

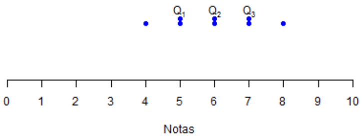


Figura 4: Turma B

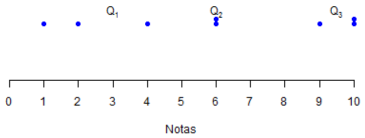
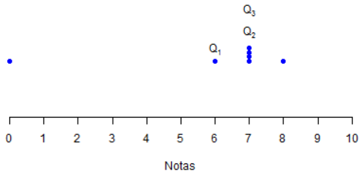
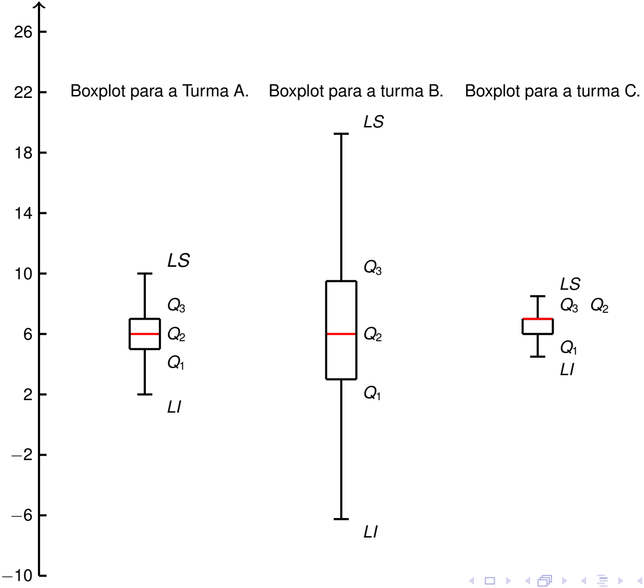


Figura 5: Turma C



Boxplot para as três turmas.



Exemplo

Considere as notas finais de uma estatística com 12 alunos.

Tabela 3: Notas finais dos 12 alunos.

$x_{(1)}$	$x_{(2)}$	$x_{(3)}$	$x_{(4)}$	$x_{(5)}$	$x_{(6)}$	$x_{(7)}$	$x_{(8)}$	$x_{(9)}$	$x_{(10)}$	$x_{(11)}$	$x_{(12)}$
0	3.35	3.41	3.83	3.98	4.08	4.19	4.55	5	5.55	5.84	10

1° **Quartil** posição $0,25 \cdot (12+1) = 3,25$. Então, $Q_1 = \frac{x_{(3)}+x_{(4)}}{2} = \frac{3,41+3,83}{2} = 3,62$;

2° **Quartil** posição $0,5 \cdot (12+1) = 7,5$. Então, $Q_2 = \frac{x_{(7)}+x_{(8)}}{2} = \frac{4,19+4,55}{2} = 4,37$;

3° **Quartil** posição $0,75 \cdot (12+1) = 9,75$. Então, $Q_2 = \frac{x_{(9)}+x_{(10)}}{2} = \frac{5+5,55}{2} = 5,275$;

Limite Superior $LS = Q_3 + 1,5(Q_3 - Q_1) = 5,275 + 1,5 \cdot 1,655 = 7,7575$;

Limite Inferior $LI = Q_1 - 1,5(Q_3 - Q_1) = 3,62 - 1,5 \cdot 1,655 = 1,1375$.

Exemplo

Figura 6: Boxplot para a turma de Estatística Básica.