

# MAE116 – Noções de Estatística

Grupo A - 1 semestre de 2012

Lista de exercícios 1 - Introdução à Estatística Descritiva – C A S A (gabarito)

---

## Exercício 1.

(4 pontos). Os dados a seguir representam medidas de Fe (%) de amostras de solos de 15 sítios obtidas da pesquisa realizada pela Expedição Central da Rússia Koala (CKE) geo-eco-química.

Sítio	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Ferro	2,29	3,84	2,87	3,11	3,43	3,62	3,84	4,28	3,01	2,81	2,3	3,65	3,93	4,58	7,1

(a) (2 pontos). Calcule a média, a moda, a mediana, o desvio padrão e os quartis.

Resposta:

Média ( $\bar{x}$ )

$$\begin{aligned}\bar{x} &= \frac{\sum_{i=1}^n x_i}{n} \\ &= \frac{2,29 + 3,84 + 2,87 + 3,11 + 3,43 + 3,62 + 3,84 + 4,28 + 3,01 + 2,81 + 2,3 + 3,65 + 3,93 + 4,58 + 7,1}{15} \\ &= 3,64\end{aligned}$$

Moda ( $M_o$ )

Dado com maior frequência.  $M_o = 3,84$

Mediana ( $M_d$ )

Primeiro ordenamos os dados: 2,29 2,3 2,81 2,87 3,01 3,11 3,43 **3,62** 3,65 3,84 3,84 3,93 4,28 4,58 7,1 e calculamos a posição da mediana como  $\frac{n+1}{2} = \frac{15+1}{2} = 8$ , e dado que  $n$  é ímpar então a mediana é o dado ordenado na posição 8, isto é,  $M_d = 3,62$ .

Desvio padrão ( $s$ )

# MAE116 – Noções de Estatística

Grupo A - 1 semestre de 2012

Lista de exercícios 1 - Introdução à Estatística Descritiva – C A S A (gabarito)

---

O desvio padrão é a raiz quadrada da variância,

$$\begin{aligned}s &= \sqrt{s^2} \\ &= \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \\ &= \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1}} \\ &= \sqrt{\frac{(2,29 - 3,64)^2 + (2,3 - 3,64)^2 + (2,81 - 3,64)^2 + (2,87 - 3,64)^2 + \dots + (7,1 - 3,64)^2}{15 - 1}} \\ &= \sqrt{\frac{19,051}{14}} \\ &= 1,17\end{aligned}$$

De outra forma,

$$\begin{aligned}s &= \sqrt{s^2} \\ &= \sqrt{\frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2}{n - 1}} \\ &= \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_{15}^2 - n\bar{x}^2}{n - 1}} \\ &= \sqrt{\frac{2,29^2 + 2,3^2 + 2,81^2 + \dots + 7,1^2 - (15)(3,64)^2}{15 - 1}} \\ &= \sqrt{\frac{218,23 - (15)(3,64)^2}{14}} \\ &= 1,17\end{aligned}$$

Quartis

Quando um conjunto de dados ordenado é dividido em 4 partes iguais, temos 3 quartis, que correspondem com os percentis 25, 50 e 75, calculados como o valor da variável que ocupa a posição  $p \times (n + 1)$ , com  $p=0,25$ ,  $p=0,50$  e  $p=0,75$ .

Dados ordenados: 2,29 2,3 2,81 2,87 3,01 3,11 3,43 3,62 3,65 3,84 3,84 3,93 4,28 4,58 7,1

- Quartil 1

$$Q_1 = \text{Percentil } 25 = \text{dado ordenado na posição } 0,25(16) = 4, \rightarrow Q_1 = 2,87$$

- Quartil 2

$$Q_2 = \text{Percentil } 50 = \text{Mediana} \rightarrow Q_2 = M_d = 3,62$$

- Quartil 3

# MAE116 – Noções de Estatística

Grupo A - 1 semestre de 2012

Lista de exercícios 1 - Introdução à Estatística Descritiva – C A S A (gabarito)

---

$$Q_3 = \text{Percentil } 75 = \text{ dado ordenado na posição } 0,75(16) = 12, \rightarrow Q_3 = 3,93$$

□

- (b) (2 pontos). Note que o sitio 15 apresenta um valor atípico. Remova esse valor e refaça o item anterior. Comente as diferenças encontradas.

Resposta:

Média ( $\bar{x}$ )

$$\begin{aligned}\bar{x} &= \frac{\sum_{i=1}^n x_i}{n} \\ &= \frac{2,29 + 3,84 + 2,87 + 3,11 + 3,43 + 3,62 + 3,84 + 4,28 + 3,01 + 2,81 + 2,3 + 3,65 + 3,93 + 4,58}{14} \\ &= 3,40\end{aligned}$$

Moda ( $M_o$ )

Dado com maior frequência.  $M_o = 3,84$

Mediana ( $M_d$ )

Primeiro ordenamos os dados: 2,29 2,3 2,81 2,87 3,01 3,11 **3,43 3,62** 3,65 3,84 3,84 3,93 4,28 4,58 e calculamos a posição da mediana como  $\frac{n+1}{2} = \frac{14+1}{2} = 7,5$ . Neste caso como  $n$  é par então a mediana é a média dos dados nas duas posições centrais, ou seja, nas posições 7 e 8, deste modo, temos:  $M_d = \frac{3,43+3,62}{2} = 3,52$

Desvio padrão ( $s$ )

# MAE116 – Noções de Estatística

Grupo A - 1 semestre de 2012

Lista de exercícios 1 - Introdução à Estatística Descritiva – C A S A (gabarito)

---

O desvio padrão é a raiz quadrada da variância,

$$\begin{aligned}s &= \sqrt{s^2} \\ &= \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \\ &= \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n - 1}} \\ &= \sqrt{\frac{(2,29 - 3,40)^2 + (2,3 - 3,40)^2 + (2,81 - 3,40)^2 + (2,87 - 3,40)^2 + \cdots + (4,58 - 3,40)^2}{14 - 1}} \\ &= \sqrt{\frac{6,2539}{13}} \\ &= 0,69\end{aligned}$$

De outra forma,

$$\begin{aligned}s &= \sqrt{s^2} \\ &= \sqrt{\frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2}{n - 1}} \\ &= \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2 - n\bar{x}^2}{n - 1}} \\ &= \sqrt{\frac{2,29^2 + 2,3^2 + 2,81^2 + \cdots + 4,58^2 - (14)(3,4)^2}{14 - 1}} \\ &= \sqrt{\frac{167,82 - (14)(3,4)^2}{13}} \\ &= 0,69\end{aligned}$$

Quartis

Dados ordenados: 2,29 2,3 2,81 2,87 3,01 3,11 3,43 3,62 3,65 3,84 3,84 3,93 4,28 4,58

- Quartil 1

$Q_1 = \text{Percentil } 25 = \text{dado ordenado na posição } 0,25(15) = 3,75$ , neste caso, média dos dados nas posições 3 e 4  $\rightarrow Q_1 = \frac{2,81+2,87}{2} = 2,84$

- Quartil 2

$Q_2 = \text{Percentil } 50 = \text{Mediana} \rightarrow Q_2 = M_d = 3,52$

- Quartil 3

$Q_3 = \text{Percentil } 75 = \text{dado ordenado na posição } 0,75(15) = 11,25$ , neste caso, média dos dados nas posições 11 e 12  $\rightarrow Q_3 = \frac{3,84+3,93}{2} = 3,89$

# MAE116 – Noções de Estatística

Grupo A - 1 semestre de 2012

Lista de exercícios 1 - Introdução à Estatística Descritiva – C A S A (gabarito)

---

## Observações:

Na seguinte tabela são comparados os resultados obtidos da amostra completa e da amostra sem o dado atípico 7,1. Note-se que as medidas mais sensíveis à presença deste dado são a média e o desvio padrão, elas diminuem em menor medida que a mediana e os quantis, no entanto, a moda é a mesma.

$n = 15$	$n = 14$ (sem dado 7,1)
$\bar{x} = 3,64$	$\bar{x} = 3,40$
$M_o = 3,84$	$M_o = 3,84$
$M_d = 3,62$	$M_d = 3,52$
$s = 1,17$	$s = 0,69$
$Q_1 = 2,87$	$Q_1 = 2,84$
$Q_2 = 3,62$	$Q_2 = 3,52$
$Q_3 = 3,93$	$Q_3 = 3,89$

□

# MAE116 – Noções de Estatística

Grupo A - 1 semestre de 2012

Lista de exercícios 1 - Introdução à Estatística Descritiva – C A S A (gabarito)

---

## Exercício 2.

(2 pontos). Um experimento para verificar o efeito de insulina na produção de glicogênio muscular em peixes foi realizado com duas espécies de peixes. Para cada espécie, dezesseis peixes similares foram distribuídos ao acaso em dois grupos sendo que ao primeiro grupo (7 peixes) foi dada a solução salina (placebo) e ao segundo (9 peixes) uma solução com insulina. Após 24 horas os animais foram sacrificados e mediu-se a quantidade de glicogênio muscular (mg/g). Analise descritivamente os resultados, utilizando a média, a mediana, o desvio padrão e o coeficiente de variação.

Espécie	Solução	Quantidade de glicogênio muscular (mg/g)								
		1,05	0,58	0,62	0,50	1,01	1,18	0,57		
A	salina	1,05	0,58	0,62	0,50	1,01	1,18	0,57		
	insulina	0,7	1,5	1,30	1,30	1,23	0,63	1,30	1,45	1,44
B	salina	1,37	0,35	0,56	0,30	1,27	1,30	0,35		
	insulina	0,40	2,01	1,47	1,60	1,25	0,15	1,40	1,50	1,49

*Resposta:*

Para as duas espécies, A e B, em os dois tipos de soluções foram obtidas as medidas descritivas, registradas na seguinte tabela.

Espécie	Solução	Medidas descritivas			
		Média $\bar{x}$	Mediana $M_d$	Desvio padrão $s$	Coef. de variação $CV = (s/\bar{x}) * 100$
A	salina	0,79	0,62	0,28	35,69
	insulina	1,21	1,30	0,32	26,49
B	salina	0,79	0,56	0,50	63,78
	insulina	1,25	1,47	0,59	47,44

Note-se que para ambas as espécies A e B, o valores da média, a mediana e o desvio padrão são menores na solução do placebo (salina) que na solução insulina, porém, segundo o coeficiente de variação, o efeito da insulina na produção de glicogênio muscular em relação a média é menos disperso que a solução salina, sendo que a espécie B apresenta quase o dobro da variabilidade da espécie A.

Em relação as espécies, as médias e as medianas do efeito das soluções salina e insulina na produção de glicogênio muscular são aproximadamente iguais, para a solução salina uma média aproximada de 0,8 e mediana de 0,6, para a solução insulina, a média e a mediana são aproximadamente 1,2 e 1,4. No entanto, os efeitos de ambas soluções na especie A são

# MAE116 – Noções de Estatística

Grupo A - 1 semestre de 2012

Lista de exercícios 1 - Introdução à Estatística Descritiva – C A S A (gabarito)

---

*menos variáveis que na espécie B, sendo que o efeito do placebo (solução salina) é mais disperso.*

*Deste modo, pode ser concluído que para ambas as espécies, a solução insulina é melhor do que o placebo para produzir glicogênio muscular.*

□

# MAE116 – Noções de Estatística

Grupo A - 1 semestre de 2012

Lista de exercícios 1 - Introdução à Estatística Descritiva – C A S A (gabarito)

---

## Exercício 3.

(4 pontos). Na tabela abaixo estão os dados referentes a uma amostra de 21 trabalhadores. As variáveis são:

S: Renda (Milhares de reais);

T: Trabalhador da indústria moderna (M) ou tradicional (T);

Z: O período em que está trabalhando: manhã (M), tarde (T), noite (N).

S	T	Z	S	T	Z	S	T	Z
3,5	M	N	1,7	T	M	3,2	M	M
4	T	M	2,5	T	T	2,4	M	N
3,2	M	M	2,2	T	N	3,4	M	T
2,7	M	M	3,7	M	N	2,7	M	T
2,9	T	T	4,5	M	T	1,8	T	M
3,1	T	N	3,8	M	T	1,5	T	M
1,9	T	M	2,4	T	M	1,9	T	T

(a) (1 pontos). Classifique cada uma das variáveis.

Resposta:

*Las variáveis são classificadas como:*

*S: Renda (Milhares de reais). Variável Quantitativa contínua*

*T: Trabalhador da indústria moderna (M) ou tradicional (T). Variável Qualitativa Nominal*

*Z: O período em que está trabalhando: manhã (M), tarde (T), noite (N). Variável Qualitativa Nominal, porém, esta variável também pode ser classificada como Qualitativa Ordinal.*

□

(b) (1,5 pontos). Divida os trabalhadores em dois grupos a partir da variável T. Calcule a média, a mediana e o coeficiente de variação da renda para os dois grupos. Compare essas medidas

Resposta:

*A variável T classifica os trabalhadores como empregados de dois tipos de indústria, moderna e tradicional. Na seguinte tabela são selecionados los valores da renda destes trabalhadores e obtidas as medidas descritivas.*



# MAE116 – Noções de Estatística

Grupo A - 1 semestre de 2012

Lista de exercícios 1 - Introdução à Estatística Descritiva – C A S A (gabarito)

---

Medidas descritivas	Renda	
	Indústria moderna (M)	Indústria tradicional (T)
	2,4	1,5
	2,7	1,7
	2,7	1,8
	3,2	1,9
	3,2	1,9
	3,4	2,2
	3,5	2,4
	3,7	2,5
	3,8	2,9
	4,5	3,1
		4,0
Media ( $\bar{x}$ )	3,3	2,4
Mediana ( $M_d$ )	3,3	2,2
Desvio padrão ( $s$ )	0,6	0,7
Coef. variação $CV$	18,7	31,5

Note-se que a renda média dos trabalhadores da indústria tradicional (2,35) é menor do que a renda dos trabalhadores da indústria moderna (3,10). Observando os valores da mediana podemos concluir que a metade dos trabalhadores da indústria tradicional ganha R\$ 2.200 ou menos, no entanto, a metade dos trabalhadores da indústria moderna ganha R\$3.200 ou menos. Considerando que a variação da renda, segundo o coeficiente de variação, é maior na indústria tradicional, pode concluir-se que o tipo de indústria influi na renda.

□

- (c) (1,5 pontos). Divida agora os trabalhadores em 3 grupos, de acordo com o período em que está trabalhando. Repita os cálculos feitos em (b) para estes grupos e responda, você diria que o período em que trabalhador exerce sua atividade afeta sua renda?

Resposta:

A variável  $Z$  classifica os trabalhadores no período de trabalho. Na seguinte tabela são selecionados los valores da renda destes trabalhadores e obtidas algumas medidas descritivas.

Da tabela anterior note-se que em relação a renda média, os trabalhadores da tarde tem uma renda melhor, porém, a metade dos empregados da noite tem uma renda de R\$ 3.100 ou mas e sua variabilidade é menor, comparada, com a variabilidade da renda dos trabalhadores das outras jornadas.

# MAE116 – Noções de Estatística

Grupo A - 1 semestre de 2012

Lista de exercícios 1 - Introdução à Estatística Descritiva – C A S A (gabarito)

---

Medidas descritivas	Renda		
	Manhã	Tarde	Noite
	3,2	4,5	3,5
	2,7	3,8	3,7
	3,2	3,4	2,4
	4,0	2,7	3,1
	1,9	2,9	2,2
	1,7	2,5	
	2,4	1,9	
	1,8		
	1,5		
Media ( $\bar{x}$ )	2,49	3,10	2,98
Mediana ( $M_d$ )	2,40	2,900	3,10
Desvio padrão ( $s$ )	0,85	0,870	0,66
Coef. variação $CV$	34,12	28,06	22,18

*Ainda tendo maior quantidade de empregados o período da manhã, a renda é a menor do que a renda dos trabalhadores da jornada da tarde e noite. Assim, pode-se concluir que o período de trabalho aparenta influir na renda dos trabalhadores.*

□