

Medidas Descritivas

VARIÁVEL QUANTITATIVA

Medidas Descritivas

❑ **Ideia:** Resumir os dados usando um ou alguns valores. Esses valores se dividem em dois casos:

❑ **Medidas de Posição:** Valores representativos ou valores típicos.

❑ Média, Mediana e Moda

❑ **Medidas de Dispersão:** Valor que representa como os dados se distribuem em torno de um valor representativo.

❑ Variância, Desvio Padrão e Desvio Médio

Média e Moda

- Média é uma medida de centro, uma medida de posição dos dados.

Seja x_1, \dots, x_n valores de uma variável aleatória quantitativa (discreta ou contínua), então a média é dada por

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}.$$

- Moda é mais indicada para variável para variável quantitativa discreta e pode ser usada também para variável qualitativa. A moda é o valor mais frequente dos dados.

Mediana

A Mediana é uma medida de posição.

Seja x_1, \dots, x_n valores de uma variável aleatória quantitativa, ordene esses valores em ordem crescente

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}.$$

A mediana é o “ponto do meio” de $x_{(1)}, x_{(2)}, \dots, x_{(n)}$. Mais precisamente,

$$md(x) = \begin{cases} x_{\left(\frac{n+1}{2}\right)}, & \text{se } n \text{ é ímpar} \\ \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2}, & \text{se } n \text{ é par} \end{cases}$$

Exemplo

Considere a variável nota para 3 turmas de estatística descrito na tabela abaixo:

A	B	C
4	1	0
5	2	6
5	4	7
6	6	7
6	6	7
7	9	7
7	10	8
8	10	

Exemplo

❑ Média, Mediana e Moda para a turma A:

❑ Dados ordenados: 4, 5, 5, 6, 6, 7, 7, 8

❑ Média: $\bar{x} = \frac{4+5+5+6+6+7+7+8}{8} = 6$

❑ Mediana: $md(x) = \frac{6+6}{2} = 6$

❑ Moda: $mo(x) = 5, 6, 7$

Exemplo

□ Média, Mediana e Moda para a turma B

□ Média: $\bar{x} = \frac{1+2+4+6+6+9+10+10}{8} = 6$

□ Mediana: $md(x) = \frac{6+6}{2} = 6$

□ Moda: $mo(x) = 6$

Exemplo

□ Média, Mediana e Moda para a turma C

□ Média: $\bar{x} = \frac{0+6+7+7+7+7+8}{7} = 6$

□ Mediana: $md(x) = 7$

□ Moda: $mo(x) = 7$

Dispersão unidimensional para as turmas

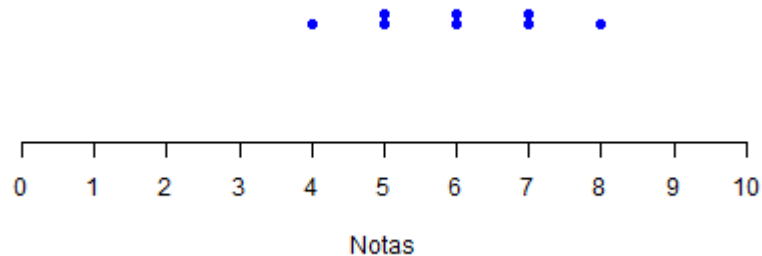
Primeiro construímos as tabelas de distribuição de frequência para as três turmas.

Turma A			
Notas	Frequência	Proporção	Porcentagem
4	1	0.1250	12.50
5	2	0.2500	25.00
6	2	0.2500	25.00
7	2	0.2500	25.00
8	1	0.1250	12.50
total	8	1.0000	100.00

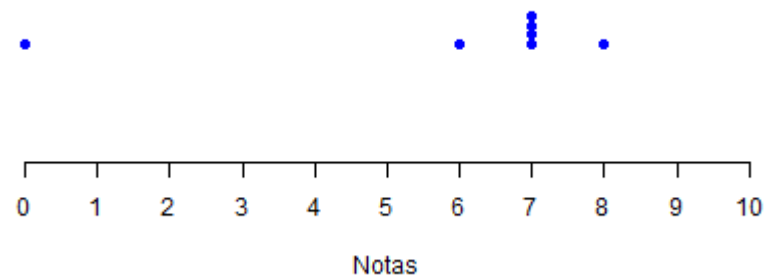
Turma B			
Notas	Frequência	Proporção	Porcentagem
1	1	0.1250	12.50
2	1	0.1250	12.50
3	0	0.0000	0.00
4	1	0.1250	12.50
5	0	0.0000	0.00
6	2	0.2500	25.00
7	0	0.0000	0.00
8	0	0.0000	0.00
9	1	0.1250	12.50
10	2	0.2500	25.00
total	8	1.0000	100.00

Turma C			
Notas	Frequência	Proporção	Porcentagem
0	1	0.1429	14.29
1	0	0.0000	0.00
2	0	0.0000	0.00
3	0	0.0000	0.00
4	0	0.0000	0.00
5	0	0.0000	0.00
6	1	0.1429	14.29
7	4	0.5714	57.14
8	1	0.1429	14.29
total	7	1.0000	100.00

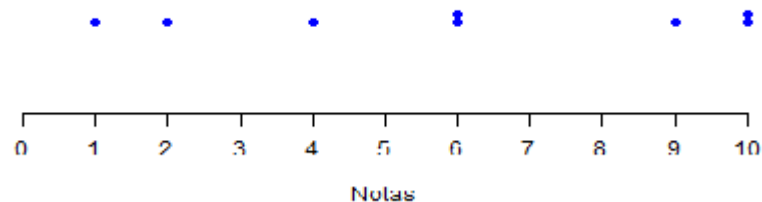
Gráfico de distribuição de unidimensional



Turma A



Turma C



Turma B

Observações

1. A média e a mediana indicam o centro das observações.
2. A média, mediana e a moda resumem um conjunto de dados em um valor típico.
3. A moda é mais indicada para variável quantitativa discreta e pode ser calculada para variável qualitativa.
4. Notamos que as três turmas em nosso exemplo têm a mesma média, mas os dados são distribuídos de forma diferente. A turma A tem valores menos dispersos que a turma B e C, a turma tem o valor 0 isolado, provavelmente um valor atípico. Para medir essa distribuição dos valores em torno da média, usamos as medidas de dispersão.

Medidas de dispersão

Avalia como os valores são distribuídos em torno da média. O princípio básico é analisar os desvios das observações em relação a média das observações, em que o desvio é a diferença entre um valor e a média: $x_i - \bar{x}$.

Podemos calcular três tipos de medidas de dispersão: variância, desvio padrão e desvio médio.

Variância

Média dos desvios ao quadrado:

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n-1},$$

em que x_1, \dots, x_n são os valores observados de uma variável quantitativa, \bar{x} é a média e n é o número de elementos na amostra.

Como calcular:

- ❑ Calculo os desvios ao quadrado: $(x_1 - \bar{x})^2, (x_2 - \bar{x})^2, \dots, (x_n - \bar{x})^2$
- ❑ Somos os desvios ao quadrado: $(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2$
- ❑ Divido a soma por $n - 1$: $\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n-1}$

Observação sobre a variância

A variância não tem a mesma unidade dos dados. Por exemplo, se x_1, \dots, x_n são medidos em cm , a unidade da variância é cm^2 . Uma solução é tirar a raiz quadrada da variância que é exatamente o desvio padrão ou usar o desvio médio.

Desvio Padrão e Desvio Médio

□ Desvio Padrão:

$$s = \sqrt{s^2}$$

□ Desvio Médio:

$$dm(x) = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \cdots + |x_n - \bar{x}|}{n}$$

Como calcular o desvio médio

1. Calcule os desvios absolutos: $|x_1 - \bar{x}|, |x_2 - \bar{x}|, \dots, |x_n - \bar{x}|$
2. Faça a média dos desvios absolutos:

$$dm(x) = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \dots + |x_n - \bar{x}|}{n}$$

Exemplo

Considere as três notas da turma de estatística.

Turma	Notas								Média
A	4	5	5	6	6	7	7	8	6
B	1	2	4	6	6	9	10	10	6
C	0	6	7	7	7	7	8		6

Exemplo

□ Turma A:

$$□ s^2 = \frac{(4-6)^2 + (5-6)^2 + (5-6)^2 + (6-6)^2 + (6-6)^2 + (7-6)^2 + (7-6)^2 + (8-6)^2}{8-1} = 1,71$$

$$□ s = \sqrt[2]{s^2} = 1,31$$

$$□ dm = \frac{|4-6| + |5-6| + |5-6| + |6-6| + |6-6| + |7-6| + |7-6| + |8-6|}{8} = 1,00$$

□ Turma B:

$$□ s^2 = \frac{(1-6)^2 + (2-6)^2 + (4-6)^2 + (6-6)^2 + (6-6)^2 + (9-6)^2 + (10-6)^2 + (10-6)^2}{7} = 12,29$$

$$□ s = \sqrt[2]{s^2} = \sqrt[2]{12,29} = 3,51$$

$$□ dm = \frac{|1-6| + |2-6| + |4-6| + |6-6| + |6-6| + |9-6| + |10-6| + |10-6|}{8} = 2,75$$

□ Turma C:

$$□ s^2 = \frac{(0-6)^2 + (6-6)^2 + (7-6)^2 + (7-6)^2 + (7-6)^2 + (7-6)^2 + (8-6)^2}{6} = 7,33$$

$$□ s = \sqrt[2]{s^2} = 2,71$$

$$□ dm = \frac{|0-6| + |6-6| + |7-6| + |7-6| + |7-6| + |7-6| + |8-6|}{7} = 1,71$$

Uso de distribuição de frequência

Caso discreto

Para calcular a média, moda, variância, desvio padrão e desvio médio podemos usar a tabela de distribuição de frequência. Por exemplo, a média, a variância e o desvio médio podem par as três turmas de estatística como abaixo:

□ Turma A:

$$\square \bar{x} = \frac{2+2\cdot5+2\cdot6+2\cdot7+8}{8} = 6$$

$$\square s^2 = \frac{(4-6)^2+2\cdot(5-6)^2+2\cdot(6-6)^2+2\cdot(7-6)^2+(8-6)^2}{8} = 1,71$$

$$\square dm = \frac{|4-6|+2\cdot|5-6|+2\cdot|6-6|+2\cdot|7-6|}{8} = 1,00$$

□ Turma B:

$$\square \bar{x} = \frac{1+2+4+2\cdot6+9+2\cdot10}{8} = 6$$

$$\square s^2 = \frac{(6-1)^2+(6-2)^2+(6-4)^2+2\cdot(6-6)^2+(9-6)^2+(10-6)^2}{8} = 12,29$$

$$\square dm = \frac{|6-1|+|6-2|+|6-4|+2\cdot|6-6|+|9-6|+|10-6|}{8}$$

□ Turma C:

$$\square \bar{x} = \frac{0+6+4\cdot7+8}{7} = 6$$

$$\square s^2 = \frac{(6-0)^2+(6-6)^2+4\cdot(7-6)^2+(8-6)^2}{7} = 7,33$$

$$\square dm = \frac{|6-0|+|6-6|+4\cdot|6-7|+|8-6|}{7} = 1,71$$

Uso de distribuição de frequência

Caso contínuo

Considere as taxas de mortalidade em 1982 em 34 cidades do oeste catarinense conforme tabela abaixo.

Taxa de mortalidade infantil							
9,9	10,3	11,9	13,1	13,9	15,7	17	18
18,3	18,4	19,6	20	20,3	21,7	22	22,6
22,7	23,5	23,7	23,8	25,4	27,2	27,3	28,9
29,7	29,9	32,3	32,9	33,0	36,3	36,4	38,3
39,2	62,2						

Exemplo – Distribuição de Frequência

Taxa de Mortalidade Infantil	Frequência	Proporção	Porcentagem	ponto média
0 ----- 10	1	0.0294	2.94	5
10 ----- 20	10	0.2941	29.41	15
20 ----- 30	15	0.4412	44.12	25
30 ----- 40	7	0.2059	20.59	35
40 ----- 50	0	0.0000	0.00	45
50 ----- 60	0	0.0000	0.00	55
60 ----- 70	1	0.0294	2.94	65
Total	34	1.0000	100.00	--

Exemplo

□ Usando a distribuição de frequência:

$$\square \bar{x} = \frac{5 \cdot 1 + 15 \cdot 10 + 25 \cdot 15 + 35 \cdot 7 + 55}{34} = 24,71$$

$$\square s^2 = \frac{1 \cdot (5 - 24,71)^2 + 10 \cdot (5 - 24,71)^2 + 15 \cdot (25 - 24,71)^2 + 7 \cdot (35 - 24,71)^2 + (55 - 24,71)^2}{33} = 112,03$$

$$\square s = \sqrt[2]{112,03} = 10,58$$

$$\square dm = \frac{|5 - 24,71| + 10 \cdot |15 - 24,71| + 15 \cdot |25 - 24,71| + 7 \cdot |35 - 24,71| + |55 - 24,71|}{34} = 6,87$$

□ Usando os dados reais:

$$\square \bar{x} = 24,86$$

$$\square s^2 = 107,63$$

$$\square s = 10,37$$

$$\square dm = 7,70$$