Estatística

Prof. Diego de Sousa Aguiar

Assunto da Aula de Hoje: Medidas Descritivas

INTRODUÇÃO

As vezes é necessário resumir certas características das distribuições de dados (ou mesmo de frequências dados) por meio de certas quantidades.

Tais quantidades são usualmente denominadas de Medidas, por quantificarem alguns aspectos de nosso interesse.

Nosso objetivo é apresentar algumas das chamadas **Medidas de Posição**, bem como, algumas **Medidas de Dispersão**, consideradas mais importantes no campo da aplicabilidade prática do nosso dia a dia.

Tais medidas servem para:

- Localizar uma distribuição;
- Caracterizar sua variabilidade.

MEDIDAS DE POSIÇÃO

Servem para localizar a distribuição dos dados brutos (ou das frequências) sobre o eixo de variação da variável em questão.

Veremos os três tipos principais de medidas de posição:

- **Média;**
- **Mediana;**
- **Moda.**

MEDIDAS DE POSIÇÃO

- **Média Aritmética**

Se os dados consistem de n observações $x_1, x_2, x_3, \ldots, x_n$ a média é dada pela soma das observações dividida pelo o número de observações.

➢ <u>Amostra:</u>

➢ Conjunto de Dados

$$\bar{x} = \frac{\sum_{i=1}^{n} X_i}{n}$$

➢ Distribuição de Frequência

$$\bar{x} = \frac{\sum_{i=1}^{k} X_i \, . f_i}{n}$$

EXEMPLO 01

Levando em conta as notas dos alunos na disciplina de instalações elétricas prediais das turmas de 2019.2 e 2020.1, determine qual a turma que obteve a maior média.

- **Turma de 2019.2:** 2, 3, 4, 4, 5, 6, 7, 7, 7, 7, 8
- **Turma de 2020.1:** 2, 3, 4, 4, 4, 5, 6, 7, 7, 8, 9, 10

EXEMPLO 01

- **Turma de 2019.2**

$$\bar{x} = \frac{2+3+4+4+5+6+7+7+7+7+8}{11} = 5{,}45$$

- **Turma de 2020.1**

$$\bar{x} = \frac{2+3+4+4+4+5+6+7+7+8+9+10}{12} = 5{,}75$$

EXEMPLO 02

Determine a média das notas de acordo com os dados da tabela a seguir.

Notas dos alunos em Estatística em 2019

Notas	Frequência
8 |--- 10	25
6 |--- 8	27
4 |--- 6	10
2 |--- 4	12
0 |--- 2	6
Total	80

Fonte: Relatório do Professor

EXEMPLO 02

$$\bar{x} = \frac{9\,.\,25 + 7\,.\,27 + 5\,.\,10 + 3\,.\,12 + 1\,.\,6}{80} = 6{,}325$$

MEDIDAS DE POSIÇÃO

- **Mediana**

Para calcularmos a mediana é preciso ordenarmos os dados: $x_1, x_2, x_3, \dots, x_n$. A mediana de um conjunto de dados é:

- Se *n* é ímpar $\Rightarrow$ $$Md = x_{\left(\frac{n+1}{2}\right)}$$

- Se *n* é par $\Rightarrow$ $$Md = \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n+1}{2}\right)}}{2}$$

EXEMPLO 03

Levando em conta as notas dos alunos na disciplina de instalações elétricas prediais das turmas de 2019.2 e 2020.1, determine qual a mediana de cada turma.

- **Turma de 2019.2:** 2, 3, 4, 4, 5, 6, 7, 7, 7, 7, 8
- **Turma de 2020.1:** 2, 3, 4, 4, 4, 5, 6, 7, 7, 8, 9, 10

EXEMPLO 03

- **Turma de 2019.2:** 2, 3, 4, 4, 5, 6, 7, 7, 7, 7, 8

$$Md = x_6 = 6$$

- **Turma de 2020.1:** 2, 3, 4, 4, 4, 5, 6, 7, 7, 8, 9, 10

$$Md = \frac{x_6 + x_7}{2} = \frac{5 + 6}{2} = 5{,}5$$

MEDIDAS DE POSIÇÃO

- **Moda**

Valor que ocorre com maior frequência. A moda é Obtida por inspeção da tabela de distribuição de frequências. Ao contrário do que acontece com a mediana e a média, uma amostra pode possuir mais do que uma moda.

EXEMPLO 04

Levando em conta as notas dos alunos na disciplina de instalações elétricas prediais das turmas de 2019.2 e 2020.1, determine qual a moda de cada turma.

- **Turma de 2019.2:** 2, 3, 4, 4, 5, 6, 7, 7, 7, 7, 8
- **Turma de 2020.1:** 2, 3, 4, 4, 4, 5, 6, 7, 7, 8, 9, 10

EXEMPLO 04

- **Turma de 2019.2:** 2, 3, 4, 4, 5, 6, 7, 7, 7, 7, 8

A moda da turma de 2019.2 =7

- **Turma de 2020.1:** 2, 3, 4, 4, 4, 5, 6, 7, 7, 8, 9, 10

A moda da turma de 2020.1 =4

EXEMPLO 05

Determine a moda das medições dos níveis de ruídos aferidos:

Medição de níveis de ruído em uma unidade fabril

Nível de Ruído (dB)	Frequência
79	2
81	4
82	2
84	3
85	1
86	1

Fonte: Dados Fictícios

Nesse caso a moda é 81 dB, pois sua frequência.

MEDIDAS DE DISPERSÃO

A informação fornecida pelas Medidas de Posição em geral necessitam de ser complementas pelas Medidas de Dispersão.

As Medidas de Dispersão servem para indicar o “quanto os dados se apresentam dispersos em torno da região central”.

Portanto caracterizam o grau de variação existente em um conjunto de valores.

As Medidas de Dispersão que mais nos interessam são:

- **Amplitude;**
- **Variância;**
- **Desvio Padrão;**
- **Coeficiente de Variação.**

MEDIDAS DE DISPERSÃO

- **Amplitude**

A amplitude, já mencionada, é definida como a diferença entre o maior e o menor valores do conjunto de dados.

Salvo aplicações de Controle de Qualidade, a amplitude não é muito utilizada como Medida de Dispersão.

MEDIDAS DE DISPERSÃO

- **Variância**

A variância é definida como a "média dos quadrados das diferenças entre os valores em relação a sua própria média".

$$\sigma^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}$$

EXEMPLO 05

Levando em conta as notas dos alunos na disciplina de instalações elétricas prediais das turmas de 2019.2, determine qual a variância da turma.

- **Turma de 2019.2:** 2, 3, 4, 4, 5, 6, 7, 7, 7, 7, 8

EXEMPLO 05

- **Turma de 2019.2:** 2, 3, 4, 4, 5, 6, 7, 7, 7, 7, 8

$$\bar{x} = \frac{2+3+4+4+5+6+7+7+7+7+8}{11} = 5{,}45$$

$$\sigma^2 = \frac{\begin{array}{c}(2-5{,}45)^2+(3-5{,}45)^2+(4-5{,}45)^2+(4-5{,}45)^2 \\ +(5-5{,}45)^2+(6-5{,}45)^2+(7-5{,}45)^2 \\ +(7-5{,}45)^2+(7-5{,}45)^2+(7-5{,}45)^2+(8-5{,}45)^2\end{array}}{11}$$

$$= 3{,}52$$

MEDIDAS DE DISPERSÃO

- **Desvio Padrão**

Definimos desvio padrão como "a raiz quadrada positiva da variância". O cálculo do desvio padrão é feito por meio da variância.

O desvio padrão se expressa na mesma unidade da variável, sendo por isso, de maior interesse que a variância nas aplicações práticas.

É mais realístico para efeito de comparação de dispersões

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}}$$

EXEMPLO 06

Levando em conta as notas dos alunos na disciplina de instalações elétricas prediais das turmas de 2019.2, determine qual o desvio padrão da turma.

- **Turma de 2019.2:** 2, 3, 4, 4, 5, 6, 7, 7, 7, 7, 8

EXEMPLO 06

- **Turma de 2019.2:** 2, 3, 4, 4, 5, 6, 7, 7, 7, 7, 8

$$\bar{x} = \frac{2+3+4+4+5+6+7+7+7+7+8}{11} = 5{,}45$$

$$\sigma^2 = \frac{\begin{array}{c}(2-5{,}45)^2+(3-5{,}45)^2+(4-5{,}45)^2+(4-5{,}45)^2\\ +(5-5{,}45)^2+(6-5{,}45)^2+(7-5{,}45)^2\\ +(7-5{,}45)^2+(7-5{,}45)^2+(7-5{,}45)^2+(8-5{,}45)^2\end{array}}{11}$$
$$= 3{,}52$$

$$\sigma = \sqrt{3{,}52} = 1{,}876$$

MEDIDAS DE DISPERSÃO

- **Coeficiente de Variação**

O coeficiente de variação é definido como "o quociente entre o desvio padrão e a média", sendo frequentemente expresso em porcentagem.

A vantagem da utilização do CV é caracterizar a dispersão dos dados em termos relativos ao seu valor médio.

Pequena dispersão absoluta pode ser, na verdade considerável, quando comparada com a ordem de grandeza dos valores da variável. Quando consideramos o CV, enganos de interpretações desse tipo não ocorrem.

Além disso, por ser adimensional, o CV fornece uma maneira de se compararem as dispersões de variáveis cujas medidas são irredutíveis.

MEDIDAS DE DISPERSÃO

- **Coeficiente de Variação**

$$\boxed{CV = \frac{\sigma}{\overline{X}}}$$

EXEMPLO 07

Levando em conta as notas dos alunos na disciplina de instalações elétricas prediais das turmas de 2019.2, determine qual o desvio padrão da turma.

- **Turma de 2019.2:** 2, 3, 4, 4, 5, 6, 7, 7, 7, 7, 8

EXEMPLO 07

- **Turma de 2019.2:** 2, 3, 4, 4, 5, 6, 7, 7, 7, 7, 8

$$\bar{x} = \frac{2+3+4+4+5+6+7+7+7+7+8}{11} = 5{,}45$$

$$\sigma^2 = \frac{\begin{array}{c}(2-5{,}45)^2+(3-5{,}45)^2+(4-5{,}45)^2+(4-5{,}45)^2\\ +(5-5{,}45)^2+(6-5{,}45)^2+(7-5{,}45)^2\\ +(7-5{,}45)^2+(7-5{,}45)^2+(7-5{,}45)^2+(8-5{,}45)^2\end{array}}{11}$$
$$= 3{,}52$$

$$\sigma = \sqrt{3{,}52} = 1{,}876$$

EXEMPLO 07

- **Turma de 2019.2:** 2, 3, 4, 4, 5, 6, 7, 7, 7, 7, 8

$$CV = \frac{1,876}{5,45} = 0,3442$$

Estatística

Prof. Diego de Sousa Aguiar

Assunto da Aula de Hoje: Medidas Descritivas