MEDIDAS DE TENDÊNCIA CENTRAL

"…a tendência de dados quantitativos de se agruparem ao redor de um valor central."

As medidas de tendência central mais comuns são a média aritimética, mediana e moda

Média Aritmética: (ou simplesmente, média) - a soma de todas as medições divididas pelo número de observações no conjunto de dados.

Mediana: é uma medida de localização do centro da distribuição dos dados, definida do seguinte modo: ordenados os elementos da amostra, a mediana é o valor (pertencente ou não à amostra) que a divide ao meio, isto é, 50% dos elementos da amostra são menores ou iguais à mediana e os outros 50% são maiores ou iguais à mediana.

Medidas de Tendência Central: Média, Moda e Mediana

Mediana: é uma medida de localização do centro da distribuição dos dados, definida do seguinte modo: ordenados os elementos da amostra, a mediana é o valor (pertencente ou não à amostra) que a divide ao meio, isto é, 50% dos elementos da amostra são menores ou iguais à mediana e os outros 50% são maiores ou iguais à mediana.

Moda: O valor que aparece com maior frequência no conjunto de dados. Essa é a única medida de tendência central que pode ser usada com dados nominais, os quais tem atribuições de categoria puramente qualitativa.

Medidas de Tendência Central: Média, Moda e Mediana

AS MÉDIAS

	Média Aritmética	Média Geométrica	Média Harmónica	Média Quadrátic…
	Importa a soma das observações	Importa o produto das observações	Importa a soma dos inversos das observações	Importa a soma d… quadrados das observações
Dados não agrupados	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$	$\bar{x}_g = \sqrt[n]{\prod x_i}$	$\bar{x}_h = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$	$x_q = \sqrt{\frac{\sum x_i^2}{n}}$
Agrupados sem intervalos de classes	$\bar{x} = \frac{\sum_{i=1}^{n} wx}{\sum_{i=1}^{n} w}$	$\bar{x}_g = \sqrt[n]{\prod x_i^{f_i}}$	$\bar{x}_h = \frac{n}{\sum_{i=1}^{n} \frac{f_i}{x_i}}$	$\bar{x}_q = \sqrt{\frac{\sum f_i x_i^2}{n}}$
Agrupados com intervalos de classes	$\bar{x} = x_0 + c\frac{\sum_{i=1}^{n} d'_i f_i}{n}$	A fórmula (2) garante-lhe um cálculo breve da média para dados agrupados. No entanto, pode usar a fórmula (1) para este tipo de dados. Só que o processo usando esta última fórmula é longo e envolve, no geral, o uso de uma calculadora.		

ALGUMAS PROPRIEDADES DA MÉDIA ARITMÉTICA

- A média de um conjunto de valores é um valor que se encontra dentro desse conjunto de valores.
- Somando-se (ou subtraindo-se) uma constante (c) de todos os valores de uma variável, a média do conjunto fica aumentada (ou diminuída) dessa constante.

CÁLCULO DE MÉDIAS

Exemplo 1. As idades dos alunos de uma turma do 10.º ano são as seguintes:

15	15	17	15	14	16	15	15	14	16	17	16	15	16	16	15	14	15	14	15	16	17

Onde se localiza o centro de gravidade dos dados?

Média Aritmética Simples

Ao centro de gravidade, refere-se à média dos dados, isto é:

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + ... + x_n) = \frac{1}{22}(15 + 15 + ... + 17) = \frac{338}{22} = 15.36$$

Média Aritmética Ponderada

Se agruparmos os valores iguais, a expressão anterior também se poderia escrever:

Idade	fi
14	4
15	9
16	6
17	3
	22

$$\bar{x} = \frac{\sum_{i=1}^{22} x_i \times f_i}{n} = \frac{14 \times 4 + 15 \times 9 + 16 \times 6 + 17 \times 3}{22} = \frac{338}{22} = 15.36$$

CALCULO DA MÉDIA PONDERADA PARA DADOS COM INTERVALOS DE CLASSE

Foram colectadas as idades de de infectados pelo COVID-19 em um dia, de um determinado mês, num determinado país. Os mesmos apresentam-se na tabela abaixo (Estes dados são fectícios e servem apenas como exemplo para exercício)

60	85	33	52	65	77	84	65	74	57	71	35	81	50	35	64	74	47	54
68	80	61	41	91	55	73	59	53	77	45	41	55	78	48	69	85	67	39
60	76	94	98	66	66	73	42	65	94	88	89							

De acordo com os dados, calcule a média das idades dos 50 estados positivo para o COVID-19

$$k = \frac{\ln n}{\ln 2} + 1 = \frac{\ln 50}{\ln 2} + 1 = 5,644 + 1 = 6,644,\ então,\ k = 6$$

$$A_t = 98 - 33 = 65$$

$$c = \frac{65}{6} = 10,8 \approx 11$$

Classes	fi	m. classe
[33 - 44[	7	38,5
[44 - 55[	7	49,5
[55 - 66[	11	60,5
[66 - 77[	11	71,5
[77 - 88[	8	82,5
[88 - 99[	6	73,5
	50	

$$\overline{X} = \frac{\sum_{i=1}^{50} f_i \times x_0}{n} = \frac{(38,5 \times 7 + 49,5 \times 7 + \ldots + 93,5 \times 6)}{50} = 65,78$$

Medidas de Tendência Central: Média, Moda e Mediana

Classes	fi	m. classe	d'	d'xfi	d'	d'xfi
[33 - 44[	7	38,5	-2	-14	-3	-21
[44 - 55[	7	49,5	-1	-7	-2	-14
[55 - 66[	11	60,5	0	0	-1	-11
[66 - 77[	11	71,5	1	11	0	0
[77 - 88[	8	82,5	2	16	1	8
[88 - 99[	6	73,5	3	18	2	12
	50			24		-26

$$d_i^{'} = \frac{x_i - x_0}{c}, onde\, i = 1\, à\, k$$

$$d_1^{'} = \frac{x_1 - x_0}{c} = \frac{38,5 - 60,5}{11} = -2$$

$$d_2^{'} = \frac{x_2 - x_0}{c} = \frac{49,5 - 60,5}{11} = -1$$

...

$$d_6^{'} = \frac{x_6 - x_0}{c} = \frac{93,5 - 60,5}{11} = 3$$

$$\overline{X} = x_0 + c\frac{\sum_{i=1}^{n} d_i^{'} \times f_i}{c} = 60,5 + 11 \times \frac{24}{50} = 65,78$$

$$\overline{X} = x_0 + c\frac{\sum_{i=1}^{n} d_i^{'} \times f_i}{c} = 71,5 - 11 \times \frac{26}{50} = 65,78$$

No caso do número de classes ser ímpar, considera-se “zero” a classe central:

Classes	fi	m. classe	d'	d'xfi
[33 - 44[	7	38,5	-2	-14
[44 - 55[	7	49,5	-1	-7
[55 - 66[	11	60,5	0	0
[66 - 77[	11	71,5	1	11
[77 - 88[	8	82,5	2	16
	44			6

$$\overline{X} = x_0 + c\frac{\sum_{i=1}^{n} d_i^{'} \times f_i}{c} = 60,5 + 11 \times \frac{6}{44} = 62$$

$$\overline{X} = \frac{\sum_{i=1}^{n} f_i \times x_0}{n} = \frac{(38,5 \times 7 + 49,5 \times 7 + \ldots + 82,5 \times 8)}{44} = 62$$

A MODA

É o valor que ocorre com maior frequência em uma serie de valores. Por exemplo, o salário modal dos empregados de uma indústria é o salário mais comum, isto é, o salário recebido pelo maior número de empregados dessa indústria.

Moda para dados não agrupados

Quando se lida com valores não agrupados, a moda é facilmente reconhecida: basta, de acordo com a definição, procurar o valor que mais se repete. Mas antes, para facilitar, recomenda-se organizar os dados em ordem – crescente ou decrescente.

Medidas de Tendência Central: Média, Moda e Mediana

Considere a série de dados: 9, 8, 7, 10, 12, 11, 10, 13, 10, 15, 10

Rol: 7, 8, 9, 10, 10, 10, 11, 12, 13, 15

TIPO DE SÉRIE DE MODA

Amodal – serie de dados sem valores repetidos.

7, 8, 9, 10, 11, 13, 15, 18, 20, 21

Unimodal – com apenas uma moda. Um único valor repete-se mais vezes.

7, 8, 9, 10, 10, 10, 11, 12, 13, 15

Bimodal – com pelo menos três modas.

9, 9, 9, 10, 10, 10, 11, 11, 11, 15

Bimodal – com apenas duas modas. Dois valores repetem-se mais vezes.

7, 8, 9, 10, 10, 10, 11, 11, 11, 15

Moda para dados agrupados sem intervalos de classe.

Uma vez agrupados os dados, é possível determinar imediatamente a moda. Para tal, basta fixar o valor da variável de maior frequência.

Tabela – Peças produzidas na Petropeças em 2000

Número de peças com defeito xi	Número de meses f_i
0	2
1	4
2	6
3	8
4	4
5	2
6	1

Neste caso o valor de maior frequência é 3 (a frequência para este valor é 8).

Logo, a moda é 3 peças com defeito.

Moda para dados agrupados com intervalos de classe.

Existem métodos diferentes para determinação da moda. Mas para o nosso estudo, iremos determinar a moda usando o método de King.

$$M_o = l + c\frac{f_{post}}{f_{ant} + f_{post}}$$

Em que

l... *limite inferior da classe modal;*

c... *amplitude do intervalo de classe;*

$f_{ant.}$... *frequência simples da classe adjacente anterior à classe modal;*

f_{post}... *frequência simples da classe posterior à classe modal*

Classes	fi
[33 - 44[	7
[44 - 55[	7
[55 - 66[	11
[66 - 77[	11
[77 - 88[	8
[88 - 99[	6
	50

Temos duas classes modais: [55 – 66[e [66 – 77[. Logo, a série é bimodal.

$$M_o = l_{\inf} + c\frac{f_{post}}{f_{post} + f_{ant}} = 55 + 11\frac{11}{11+7} = 61.72$$

$$M_o = l_{\inf} + c\frac{f_{post}}{f_{post} + f_{ant}} = 66 + 11\frac{8}{8+11} = 70,88$$

A MEDIANA

É a terceira grandeza de tendência central e pode ser defendia como valor que divide uma serie ordenada de tal forma que pelo menos a metade ou cinquenta por cento dos itens sejam iguais ou maiores do que ela, e que haja pelo menos outra metade ou cinquenta por cento itens maiores do que ela.

Mediana para dados agrupados em classes.

$$M_d = l + c\frac{\frac{N}{2} - F_{ant}}{f_{Md}}$$

N... número de elementos

$F_{ant...}$ Frequência acumulada até a classe anterior à classe médiana

f_{Md}... frequência simples da classe médiana

Passos para cálculo da mediana para dados agrupados em classes

classes	fi	F
[33 - 44[	7	7
[44 - 55[	7	14
[55 - 66[	11	25
[66 - 77[	11	36
[77 - 88[	8	44
[88 - 99[	6	50
	50	

Passo 1: Cálculo da frequência acumulada

Passo 2: Cálculo do quociente $\frac{N}{2} = \frac{50}{2} = 25$

Passo 2: Encontrar a classe em que $\frac{N}{2} \leq F$

$$M_d = l_{\inf} + c \times \frac{\frac{N}{2} - F_{ant}}{f_{M_d}} = 66 + 11 \times \frac{\frac{50}{2} - 14}{11} = 66$$

Exercício prático: Considere uma amostra de 40 plantas, cujas alturas encotram-se agrupadas em classes.

altura	[10 - 14[	[14 - 18[	[18 - 22[	[22 - 26[	[26 - 30[	[30 - 34[	[34 - 38[	[38 - 42[
nº de plantas	15	28	40	30	20	15	10	5

Calcule as três medidas de tendência central

Classes	fi			
[10 - 14[	15			
[14 - 18[	28			
[18 - 22[	40			
[22 - 26[	30			
[26 - 30[	20			
[30 - 34[	15			
[34 - 38[	10			
[38 - 42[	5			
	163			

Medidas de Tendência Central: Média, Moda e Mediana

A seguir, propõe-se um quadro resumo de aplicação de medidas de tendência central para diferentes tipos de dados:

	Média	**Moda**	**Médiana**
Discretas	Não recomendável	Pode-se calcular	Pode-se calcular
Contínuas	Pode-se calcular	Pode-se calcular	Pode-se calcular
Nominais	Não se calcula	Pode-se calcular	Não se calcula
Ordinais	Não se calcula	Pode-se calcular	Pode-se calcular

SUMÁRIO:

Medidas de tendência central:

- Média
- Moda
- Mediana