MEDIDAS DESCRITIVAS

MEDIDAS DESCRITIVAS

- O objetivo das medidas descritivas é descrever um conjunto dedados quantitativos de forma compacta e organizada. Cabe ressaltar que as medidas vistas nesta etapa não podem ser levadas para fazer inferência para a população, são meramente para descrever o aspecto geral de um conjunto de observações;
- As principais medidas que serão abordadas são as medidas de:

a) Posição (Tendência Central)

b) Dispersão

c) Forma (Assimetria e Curtose)

MEDIDAS DE POSIÇÃO OU DE TENDÊNCIA CENTRAL

Uma medida de tendência central ou de posição de um conjunto de dados mostra o valor em torno do qual se agrupam as observações. As principais medidas de tendência central são a **média aritmética**, a **mediana** e a **moda**.

OBS: A utilização de medidas de posição não substitui o uso de tabelas e de gráficos.

MÉDIA ARITMÉTICA

A média assume caracterizações que veremos a seguir segundo a forma na qual os dados estão disponíveis.

Média para dados que não estão agrupados

Sejam $x_1, x_2, \dots, x_n$ os dados amostrais e *n* o tamanho da amostra. A média aritmética $\bar{X}$ dos valores $x_1, x_2, \dots, x_n$ é definida por

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

MÉDIA ARITMÉTICA

Exemplo 1: Calcular a média de peso de 10 alunos em kg

45,3	56,7	48,4	39,9	57,5
60,1	55,3	58,6	57,2	63,0

MÉDIA ARITMÉTICA

Média para dados que estão agrupados em tabelas de frequência

A média aritmética $\bar{X}$ é dada por

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{k} x_i F_i$$

onde k é o número de classes.

Exemplo 2: Para estudar o período de germinação de certa espécie de tomateiro um pesquisador montou, numa casa de vegetação, o seguinte experimento: tomou certa quantidade de terra bem homogênea e a colocou em 30 vasos iguais. Em cada vaso plantou 3 sementes dessa espécie de tomateiro. Ao final do período experimental anotou o número de sementes germinadas por vaso e os dados foram apresentados na tabela abaixo. Calcule a média de sementes germinadas.

MÉDIA ARITMÉTICA

Tabela: Distribuição de frequências no número de sementes germinadas por vaso

Nº de sementes/germinadas	Nº de vasos
0	2
1	5
2	15
3	8
Total	30

Fonte: Pesquisador

MÉDIA ARITMÉTICA

Média para dados que estão agrupados em intervalos

A média aritmética $\bar{X}$ é dada por

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{k} m_i F_i$$

onde k é o número de classes.

Exemplo 3: Considere o tempo de duração de certas lâmpadas em horas apresentados na tabela a seguir:

Tabela: Distribuição da duração em horas de um conjunto de lâmpadas

Duração/horas	Nº de lâmpadas
100 ⊢ 200	60
200 ⊢ 300	25
300 ⊢ 400	10
400 ⊢ 500	8
500 ⊢ 600	2
Total	105

Fonte: Pesquisador

MÉDIA ARITMÉTICA

Calcule a duração média desse conjunto de lâmpadas.

MÉDIA ARITMÉTICA

Observações

A média aritmética de um conjunto de dados:

- Sempre existe e é única (assume um único valor);
- Sofre muito a influência de valores extremos (muito grandes ou muito pequenos).

A soma dos desvios calculados em relação à média aritmética da distribuição é nula.

MEDIANA

Mediana para dados não agrupados

A mediana é o valor central de um rol, ou seja, é a medida que divide o rol em duas partes iguais, deixando 50% dos valores abaixo dessa medida e 50% dos valores acima dessa medida.

- Quando n é ímpar a mediana é um valor que pertence ao rol;
- Quando n é par a mediana é a média dos dois valores centrais.

- Se o rol é ímpar a mediana amostral é igual ao valor numérico do elemento dado por

$$\frac{n+1}{2}$$

- Se o rol é par a mediana amostral é igual a média dos valores numéricos dos elementos

$$\frac{n}{2} \; e \; \frac{n}{2}+1$$

MEDIANA

Exemplo 4

Calcule a mediana para os dados:

a) 1 8 4 5 4 3 2

b) 3 15 46 64 126 623

MEDIANA

Mediana para dados agrupados em tabelas de frequência

A identificação da mediana é feita através da frequência acumulada e o critério é o mesmo que usamos anteriormente.

Vamos calcular a mediana para o exemplo 2 que vimos anteriormente.

Tabela: Distribuição de frequências no número de sementes germinadas por vaso

Nº de sementes/germinadas	Nº de vasos
0	2
1	5
2	15
3	8
Total	30

Fonte: Pesquisador

MEDIANA

Mediana para dados agrupados em intervalo

Para dados agrupados em intervalo o estimador assume a seguinte forma:

$$Md = l_{md} + \left(\frac{\frac{n}{2} - F_{ac}(ant)}{F_{md}} \right) . a_{md}$$

onde:

- l_{md}: é o limite inferior da classe mediana;
- $F_{ac}(ant)$: é a frequência acumulada do intervalo anterior ao intervalo mediano;
- F_{md}: é a frequência absoluta da classe mediana;
- a_{md}: é a amplitude da classe mediana.

Obs: Para o caso de dados tabelados, não se leva em consideração se é par ou impar.

MEDIANA

Exemplo 5

Vamos calcular a mediana para os dados tabelados do exemplo a seguir:

Tabela: Distribuição de frequências das notas de 200 alunos em estatística da escola X

Notas	Nº de alunos
$0 \vdash 2$	28
$2 \vdash 4$	40
$4 \vdash 6$	50
$6 \vdash 8$	65
$8 \vdash 10$	17
Total	200

Fonte: Escola X

MEDIANA

MEDIANA

Observações

- A mediana não é afetada pela grandeza dos valores extremos (aberrantes) e sim pelo número de elementos da amostra;
- Se a distribuição é assimétrica, muitas vezes, a mediana é uma medida de posição que melhor representa o conjunto de dados do que a média aritmética.

MODA

A Moda de um conjunto de dados amostrais é o valor (ou valores) mais frequente no rol.

- Pode haver mais de uma moda;
- Se todos os valores distintos têm a mesma frequência, o conjunto de dados não tem moda.

Exemplo 6

i. 10 10 12 13 18

ii. 100 100 200 200 300 600

iii. 1 2 3 4 5

iv. 1 1 3 3 5 5

MODA

Moda para dados agrupados em distribuições de frequências

Para identificar a moda quando o conjunto de dados está agrupado em uma tabela de frequências basta tomar o elemento associado a maior frequência.

Exemplo 7

Tabela: Distribuição de frequências no número de sementes germinadas por vaso

Nº de sementes/germinadas	Nº de vasos
0	2
1	5
2	15
3	8
Total	30

Fonte: Pesquisador

MODA

Moda para dados agrupados em intervalos

Para calcularmos a moda quando os dados estão agrupados em intervalos utiliza-se o **Processo de Czuber** (matemático Austríaco 1851-1925) definido por:

$$Mo = l_{mo} + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right) . a_{mo}$$

Para calcular a moda devemos identificar a classe modal, ou seja, a de maior frequência e determinar:

- l_{mo}: limite inferior do intervalo modal;
- Δ_1: diferença entre a a frequência do intervalo modal e a frequência do intervalo imediatamente anterior $(F_{mo} - F_{ant\ mo})$;
- Δ_2: diferença entre a a frequência do intervalo modal e a frequência do intervalo imediatamente posterior $(F_{mo} - F_{post\ mo})$;
- a_{mo}: Amplitude do intervalo da classe modal.

MODA

Exemplo 8

Tabela: Distribuição de frequências das notas de 200 alunos em estatística da escola X

Notas	Nº de alunos
$0 \vdash 2$	28
$2 \vdash 4$	40
$4 \vdash 6$	50
$6 \vdash 8$	65
$8 \vdash 10$	17
Total	200

Fonte: Escola X

COMO ESCOLHER

Média

- Quando é necessário um tratamento mais matemático;
- Quando os dados se distribuem, ao menos aproximadamente, simetricamente;
- Quando é necessário obter outras estimativas que dependem da média como a variância e o desvio padrão.

COMO ESCOLHER

Mediana

- Quando há valores aberrantes que podem “distorcer” a média;
- Quando desejamos conhecer o ponto central exato dos nossos dados;
- Quando a distribuição dos nossos dados é muito assimétrica.

COMO ESCOLHER

Moda

- Quando a medida de interesse é o ponto mais típico ou popular dos dados;
- Quando precisamos apenas de uma rápida ideia sobre a tendência central dos nossos dados.

COMO ESCOLHER

A média é de longe a mais usada dentre as estatísticas de tendência central, mas nem sempre é ela que melhor representa um conjunto de dados. Veja o exemplo:

Exemplo: O conjunto $1, 1, 1, 1, 1, 2, 2, 2, 2, 20$ tem $\overline{X} = 3,3$ e $Md = 1,5$ que é bem mais representativo, dado que a maior parte das observações é 1 e 2;

- O fato é que a média aritmética é muito influenciada por valores extremos, tal como 20 no exemplo dado, enquanto que a mediana não é afetada por esses valores extremos;
- Quando o tamanho da amostra é grande, esta desvantagem da média torna-se menos importante;
- A propósito de análises estatísticas, a média apresenta, sobre a mediana, a vantagem de ser mais fácil de manipular algebricamente.