

ESTATÍSTICA DESCRITIVA –notas de aula 1
Profa. Maria Lúcia dos Santos Teles

População ou universo

Um conjunto finito de objetos sobre os quais se deseja alguma informação (variável). Exemplo : censo, alunos da UNIBAN

Amostra

Qualquer subconjunto da população.

Exemplo: alunos de uma turma da UNIBAN

Natureza dos Dados :

- Dados Quantitativos: consistem em números que representam contagens ou medidas.
- Dados Qualitativos: podem ser separados em diferentes categorias que se distinguem por alguma característica não numérica.

Exemplos:

1).....

2).....

3).....

DADOS BRUTOS

Aqueles que não foram numericamente organizados.

Exemplo:

.....

ROL

É o arranjo dos dados brutos em ordem de grandeza crescente ou decrescente.

Exemplo. Ordenando os dados brutos acima na ordem crescente:

.....

VARIÁVEL DISCRETA

A variável assume valores em pontos da reta real, podendo-se enumerar todos os possíveis valores. Exemplos:

- número de alunos da classe X;
- número de acidentes da Via Anchieta;
- quantidade de livros de uma biblioteca;
- peças defeituosas num lote recebido;
-
-

VARIÁVEL CONTÍNUA

Neste caso a variável assume valores em intervalos na reta real e não é possível enumerar todos os valores. Geralmente esta variável provém de medições. Exemplos:

- peso dos alunos de uma classe;
- tempo de duração de um transistor;
- estatura dos alunos de uma classe.

-

DISTRIBUIÇÃO DE FREQUÊNCIAS

REQUISITO: SOMATÓRIO (Σ)

Quando queremos representar uma soma de n valores do tipo

$$x_1 + x_2 + x_3 + \dots + x_n = \sum_{i=1}^n x_i$$

onde:

Σ é a letra Sigma, indica instrução para somar

n indica o último elemento a ser somado

i = 1 indica que a soma vai começar pelo primeiro elemento (índice 1)

x_i é a i-ésima observação da série $x_1, x_2, x_3, \dots, x_n$

Lê-se: "Somatório de x_i , para i variando de 1 a n"

Exemplos:

$$1) x_1 + x_2 + x_3 + x_4 + x_5 = \sum_{i=1}^5 x_i$$

$$2) y_3 + y_4 + y_5 + y_6 = \sum_{i=3}^6 y_i$$

$$3) (x_1 + 10)^2 + (x_2 + 10)^2 + (x_3 + 10)^2 = \sum_{i=1}^3 (x_i + 10)^2$$

4).....

5).....

TIPOS DE FREQUÊNCIAS

Frequências simples ou absolutas (f_i) - São valores que realmente representam o número de dados de cada classe. Observe que

$$\sum_{i=1}^m f_i = n, \text{ onde } n \text{ é o número de elementos da amostra}$$

Frequências relativas (fr_i) - São os valores das razões entre as frequências simples e a frequência total. Pode ser multiplicada por 100 para que os dados sejam apresentados em porcentagem:

$$fr_i = \frac{f_i}{\sum_{j=1}^m f_j} \cdot 100$$

Frequência Acumulada Simples (F_i) - É o total das frequências de todos os valores inferiores ao limite do intervalo de uma classe:

$$F_k = f_1 + f_2 + \dots + f_k$$

ou

$$F_k = \sum_{i=1}^k f_i$$

Frequência Acumulada Relativa (Fr_i) - É a frequência acumulada da classe dividida pela frequência total da distribuição:

$$Fr_i = \frac{F_i}{\sum_{j=1}^m f_j}$$

Diagrama de Barras

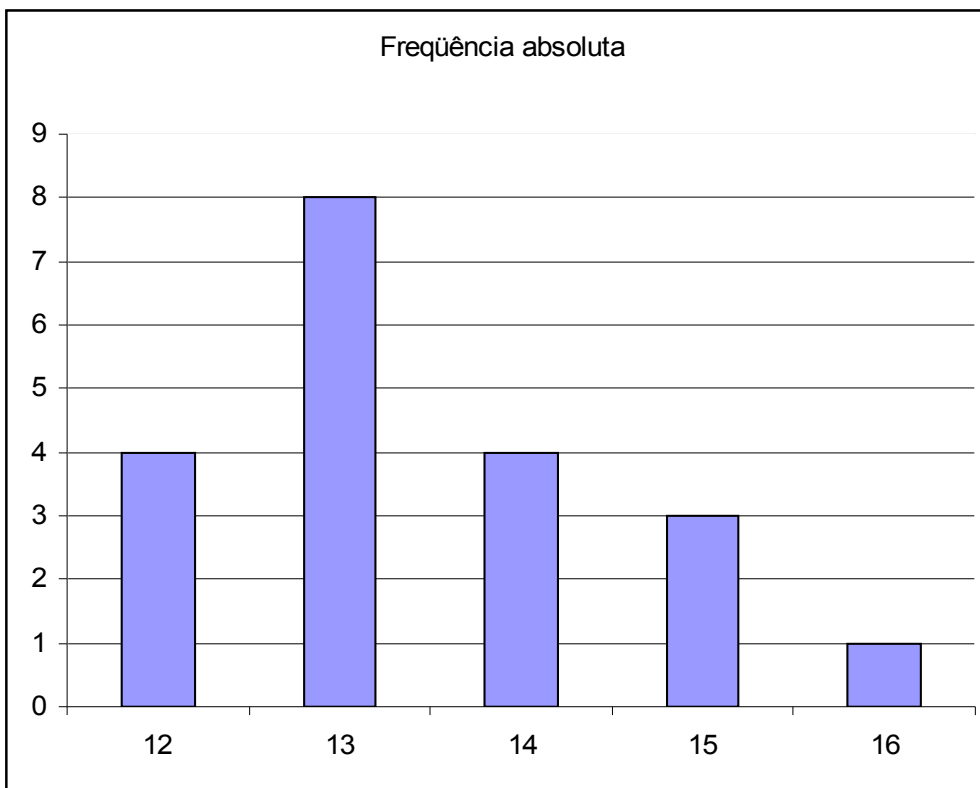
Representação gráfica que consiste em, usando um sistema de eixos coordenados xy, marcar no eixo dos x o valor das classes e nesses pontos barras verticais de altura igual à frequência absoluta ou à frequência relativa.

Exemplo

Tabela de frequências

x_i	freq. abs. (f_i)	freq. rel. (fr_i)
12	4	0,2
13	8	0,4
14	4	0,2
15	3	0,15
16	1	0,05
Total	20	1,0

Diagrama de barras



MEDIDAS DE TENDÊNCIA CENTRAL

1. MÉDIA

Média aritmética simples

Para uma sequência numérica $X: x_1, x_2, \dots, x_n$, a média aritmética simples, que designamos por $\bar{X}$ é definida por:

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

Ex.: Se $X: 2, 0, 5, 3$; então:

$$\bar{X} = \frac{2 + 0 + 5 + 3}{4} = 2,5$$

Média aritmética ponderada

Para uma sequência numérica $X: x_1, x_2, \dots, x_n$, afetados de pesos $p_1, p_2, \dots, p_m$, respectivamente, então:

$$\bar{X} = \frac{\sum_{i=1}^m x_i p_i}{\sum_{j=1}^m p_j}$$

Ex.: Seja $X: [2, 4, 5]$ com pesos 1, 3 e 2 respectivamente, então:

$$\bar{X} = \frac{\sum_{i=1}^3 x_i p_i}{\sum_{j=1}^3 p_j} = \frac{2 \times 1 + 4 \times 3 + 5 \times 2}{1 + 3 + 2} = \frac{2 + 12 + 10}{6} = 4$$

Cálculo da média aritmética ponderada usando uma tabela de frequências

Se os dados estão apresentados na forma de uma variável discreta faremos a média aritmética ponderada considerando as frequências simples de f_i como sendo as ponderações dos elementos x_i correspondentes:

Assim a fórmula para o cálculo da média é:

$$\bar{X} = \frac{\sum_{i=1}^m x_i f_i}{\sum_{j=1}^m f_j}$$

Exemplo: Determinar a média da seguinte distribuição:

x_i	f_i
2	1
5	4
6	3
8	2
Σ	10

$$\bar{X} = \frac{\sum x_i f_i}{\sum f_i} = \frac{2 \times 1 + 5 \times 4 + 6 \times 3 + 8 \times 2}{1 + 4 + 3 + 2} = \frac{2 + 20 + 18 + 16}{10} = \frac{56}{10} = 5,6$$

Observação: Embora a variável assuma somente valores inteiros, o valor da média pode ser um número real.

2. MODA

Denominamos **moda** o valor que ocorre com maior frequência em uma série de valores.

Quando dois valores ocorrem com a mesma frequência máxima, cada um deles é uma moda, e o conjunto se diz **bimodal**. Se mais de dois valores ocorrem com a mesma frequência máxima, cada um deles é uma moda, e o conjunto é **multimodal**. Quando nenhum valor é repetido, o conjunto não tem moda, podendo ser chamado de **amodal**.

Exemplos:

a) Na série X: [1; 2; 1; 3; 2; 5; 3; 1; 2; 4; 1] a moda é 1, pois é o valor que aparece com maior frequência.

b) A série Y: [1; 2; 3; 2; 4; 4; 5] apresenta como moda os valores 2 e 4, pois esses valores aparecem igualmente com a máxima frequência na série, classificada como *bimodal*.

c) A série Z: [1; 4; 2; 5] é considerada *amodal*, ou seja, não tem moda, pois nenhum valor se repete na série.

3. MEDIANA

Mediana é o valor localizado na posição central dos valores ordenados da variável, tal que 50% dos valores da variável são menores do que a mediana e os 50% restantes são maiores.

Dados não agrupados

Estando ordenados os valores da série e sendo n o número de elementos da série, o valor mediano será:

- o termo de ordem $(n+1)/2$ se n for ímpar;
- a média aritmética dos termos de ordem $n/2$ e $n/2 + 1$, se n for par.

Exemplos:

.....

.....

Dados agrupados

Vamos utilizar a frequência acumulada (= total das frequências de todos os valores inferiores ao limite superior de uma dada classe):

$$F_k = f_1 + f_2 + f_3 + \dots + f_k$$

A mediana é primeiro valor cuja frequência acumulada supera a metade da soma total das frequências $\sum f_i / 2$

Exemplo 1. Distribuição do número de filhos do sexo masculino em uma amostra de 34 famílias.

no.de meninos	f_i	F_i
0	2	2
1	6	8
2	10	18
3	12	30
4	4	34
soma	34	

Temos $\sum f_i / 2 = 17$ e como a menor frequência acumulada que supera este valor é 18, que corresponde ao valor 2 da variável, este o valor mediano.

Mediana = 2 meninos

Exemplo 2

x_i	f_i	F_i
12	1	1
14	2	3
15	1	4
16	2	6
17	1	7
20	1	8
soma	8	

$\sum f_i / 2 = 4$, que corresponde ao valor de F_3 .

Neste caso, a mediana é o valor médio entre x_3 e x_4 :

Mediana = 15,5

Exercícios

1) (somatórios) Desenvolver por extenso cada uma das somas indicadas:

a) $\sum_{i=2}^7 x_i =$

b) $\sum_{i=1}^4 (5x_i + 2) =$

c) $\sum_{i=2}^6 (x_i - a)^2 \cdot f_i =$

d) $\sum_{j=1}^6 y_j =$

e) $\sum_{i=-3}^3 i^2 =$

2) (somatórios) Escrever em notação de somatório:

- a) $(x_1 + 2) + (x_2 + 2) + (x_3 + 2)$
- b) $(x_1 - 15)^2 f_1 + (x_2 - 15)^2 f_2 + (x_3 - 15)^2 f_3$
- c) $x_7 + x_8 + x_9 + x_{10} + x_{11}$
- d) $(x_1 - y_1) + (x_2 - y_2) + \dots + (x_8 - y_8)$
- e) $x^2_1 + x^2_2 + x^2_3 + \dots + x^2_n$

3) (distribuição de freqüências) Lançou-se um dado 100 vezes, obtendo-se os seguintes resultados:

6-1-6-3-1-4-5-2-5-6-1-5-3-4-1-4-6-1-2-1
 3-1-4-6-2-5-4-3-1-5-2-2-3-6-3-5-2-4-1-6
 6-2-4-5-3-6-4-6-3-6-6-4-4-2-2-5-1-6-3-1
 2-5-3-1-1-4-3-5-1-5-3-3-1-6-5-2-6-1-4-5
 5-1-4-6-3-2-5-6-5-2-2-4-1-3-3-4-6-4-3-2

Apresentar os resultados numa distribuição de freqüência discreta.

4) (distribuição de freqüências) A tabela abaixo apresenta as vendas de um determinado aparelho elétrico, durante um mês, por empresa

14	12	11	13	14	13
12	14	13	14	11	12
12	14	10	13	15	11
15	13	16	17	14	14

Apresente os resultados numa distribuição de frequência discreta.

5) (distribuição de frequências, média e moda) As notas dos alunos de uma turma foram

2 3 3 4 4 4 4 5 5 6 7 7 7 7 7 7 8 8 9 10

- Obtenha a distribuição de frequência.
- Obtenha as frequências relativas e acumuladas.
- Calcule a média.
- Calcule a moda.
- Represente graficamente a distribuição.

6) (distribuição de frequências) A distribuição abaixo mostra o número de acidentes ocorridos com 70 motoristas de uma empresa de ônibus.

no. de acidentes	0	1	2	3	4	5	6	7
no. de motoristas	20	10	16	9	6	5	3	1

Determine:

- o número de motoristas que não sofreram nenhum acidente;
- o número de motoristas que sofreram pelo menos 4 acidentes;
- o número de motoristas que sofreram menos de 3 acidentes;
- o número de motoristas que sofreram no mínimo 3 e no máximo 5 acidentes;
- a percentagem de motoristas que sofreram no máximo 2 acidentes.

7) Calcule a média, a moda e a mediana da seguinte amostra:

Unidades vendidas de um produto em um mês:

5 - 12 - 10 - 15 - 18 - 20 - 15 - 12 - 13 - 19 - 14 - 7 - 10 - 12 - 15 - 10 - 12 - 7 - 6 - 9 - 13 - 19 - 12 - 7 - 9 - 13 - 6 - 5 - 19 - 20

8) Dadas as tabelas abaixo, calcule a média, a moda e a mediana em cada caso.

a)

x_i	f_i
12	1
14	2
15	1
16	2
17	1
20	1

b)

x_i	f_i
2	3
4	7
6	12
8	8
10	4

c)

x_i	f_i
0	2
1	5
2	9
3	7

4	6
5	3

9) Considere a distribuição de frequências abaixo.

i	x_i	f_i
1	0	5
2	1	5
3	2	5
4	3	7
5	4	4

Calcule a média, a moda e a mediana.