UNIVERSIDADE DE SÃO PAULO
Faculdade de Zootecnia e Engenharia de Alimentos

ZAB1111 – Estatística Básica

Prof. César Gonçalves de Lima cegdlima@usp.br

Aula 3

Medidas de posição ou de tendência central

3. MEDIDAS ASSOCIADAS A VARIÁVEIS QUANTITATIVAS

3.1. MEDIDAS DE POSIÇÃO OU DE TENDÊNCIA CENTRAL

Visam determinar o **centro** da distribuição dos dados de uma variável **quantitativa**, em torno do qual os dados se distribuem.

Para os dados brutos de uma série:

- **Média aritmética:** $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$

- **Mediana ($Md = P_{50}$):** é o valor central da <u>série ordenada</u> de valores (ordem crescente ou decrescente).

- **Moda (Mo):** é o valor mais frequente da série.

Uma série de dados pode não ter moda, pode ter uma ou duas modas.

O cálculo da mediana depende do número de elementos da amostra:

- Se n é ímpar $\Rightarrow$ a mediana é igual ao valor central que ocupa a posição $\left(\frac{n+1}{2}\right)$ da série ordenada.

  Exemplo: Se $n = 11 \Rightarrow$ mediana é o valor da série que ocupa a 6ª posição

- Se n é par $\Rightarrow$ a mediana é igual à média dos valores centrais da série e que ocupam as posições $\left(\frac{n}{2}\right)$ e $\left(\frac{n}{2} + 1\right)$

  Exemplo: Se $n = 10 \Rightarrow$ mediana é a média dos valores da série que ocupam a 5ª e a 6ª posição.

Exemplo 1. Calcular a média, a moda e a mediana da amostra com as idades de 10 alunos da disciplina de Estatística Básica:

18 20 18 19 19 18 21 17 20 18

Média: $\bar{x} = \frac{188}{10} = 18{,}8$ anos (18 anos, 9 meses e 18 dias...)

Moda: $Mo = 18$ anos

Dados ordenados: 17 18 18 18 **18** **19** 19 20 20 21

Mediana: $Md = \frac{(18+19)}{2} = 18{,}5$ anos

Exemplo 2: Para a série de pesos de coelhos ao desmame, não classificados numa distribuição de frequências:

492	552	560	583	657	657	666	697	699	716
727	731	737	750	770	798	808	817	823	823
830	842	842	860	873	878	880	883	900	910
940	960	960	963	992	1000	1000	1000	1020	1040

Temos: $n = 40$

$\bar{x} = 815{,}9\text{g}$ $\qquad Md = \frac{823+830}{2} = 826{,}5\text{g}$ $\qquad Mo = 1000\text{g}$

Sobre as medidas de posição:

- A média é a medida mais usada para representar uma série.
- A média é bastante influenciada por valores atípicos, aberrantes ou discrepantes (muito grandes ou muito pequenos). A moda e a mediana são pouco influenciadas por esses valores.
- Quando os dados estiverem agrupados em classes de frequências, os valores da média, moda e mediana serão afetados pela maneira como as classes de frequências são escolhidas e pela quantidade delas.
- Para distribuições de frequências unimodais e pouco assimétricas vale a relação empírica: $\text{Média} - \text{Moda} \cong 3(\text{Média} - \text{Mediana})$

Se os dados da <u>variável discreta</u> já estiverem classificados em uma distribuição de frequências:

- $\bar{x} = \frac{\sum_{i=1}^{k} x_i \, f_i}{n}$, onde $k =$ número de classes e $n = \sum_{i=1}^{k} f_i$
- $Md(X)$ é o valor que ocupa o valor central da série ordenada
- $Mo(X)$ é o valor mais frequente da série

Exemplo 3: Tamanho de ninhadas de coelhos desmamados no 1º trimestre de 1989 no setor de cunicultura do Campus:

5	4	4	4	3	3	3	8	7	1	6	8	4	9	3	5	5	5	7	2
7	6	2	4	6	5	3	3	4	7	5	6	7	4	5	6	3	6	4	5

Dados ordenados:

1	2	2	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	5	5
5	5	5	5	5	5	6	6	6	6	6	6	7	7	7	7	7	8	8	9

A distribuição de frequências do tamanho de ninhadas de coelhos desmamados auxilia nos cálculos das medidas de posição.

Ninhada	f_i	F_i	Posição
1	1	1	1º
2	2	3	2º até 3º
3	7	10	4º até 10º
4	8	18	11º até 18º
5	**8**	**26**	19º até 26º
6	6	32	27º até 32º
7	5	37	33º até 37º
8	2	39	38º até 39º
9	1	40	40º

$Md = 5{,}0$ coelhos

$\bar{x} = \frac{194}{40} = 4{,}85$ coelhos $Mo = 4$ e 5 coelhos (distribuição bimodal)

Se os dados da variável contínua estiverem classificados em uma distribuição de frequências:

- $\bar{x} = \frac{\sum_{i=1}^{k} Pm_i \, f_i}{n}$, onde $k =$ número de classes e $n = \sum_{i=1}^{k} f_i$

- $Md(X) = L_{Md} + \frac{\left(\frac{n}{2} - F_a\right)}{f_{Md}} h$

  Em que L_{Md} é o limite inferior da classe que contem a mediana, F_a é a frequência acumulada da classe anterior à classe mediana, f_{Md} é a frequência absoluta da classe que contem a mediana e h é a amplitude da classe que contem a mediana.

- $Mo(X) = L_{Mo} + \frac{(f_{Mo} - f_a)}{(f_{Mo} - f_a) + (f_{Mo} - f_p)}\ h$

  Em que L_{Mo} é o limite inferior da classe que contem a moda, f_{Mo} é a frequência absoluta da classe modal, f_a e f_p são as frequências absolutas das classes anterior e posterior da classe modal, respectivamente.

Exemplo 4. Calcular a média, a moda e a mediana dos pesos médios (gramas) de ninhadas de coelhos desmamados agrupados na seguinte distribuição de frequências.

Peso (g)	P_{mi}	f_i	F_i
490 ⊦ 590	540	4	4
590 ⊦ 690	640	3	7
690 ⊦ 790	740	8	15
790 ⊦ 890	840	13	28
890 ⊦ 990	940	6	34
990 ⊦ 1090	1040	6	40
Total		40	

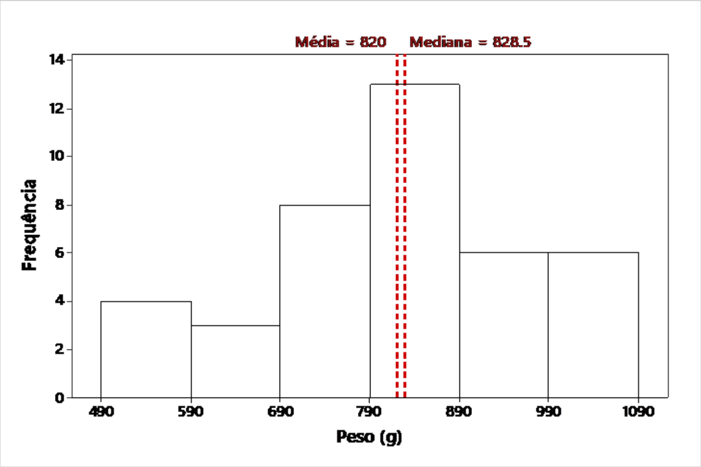
Posição

1º até 4º
5º até 7º
8º até 15º
16 até 28º
29º até 34º
35º até 40º

$\bar{x} = 820$ g

$$Md = 790 + \frac{\left(\frac{40}{2} - 1\ 5\right)}{13} 100 = 828{,}5 \text{ g}$$

$$Mo = 790 + \frac{(13-8)}{(13-8) + (13-6)} 100 = 831{,}7 \text{ g}$$



Histograma dos pesos médios (gramas) de ninhadas de coelhos desmamados

Observações finais:

- É melhor usar os dados brutos nos cálculos das medidas de tendência central, sempre que estiverem disponíveis.
- Se tivermos acesso somente à distribuição de frequências dos dados de uma variável contínua, vale lembrar que as medidas calculadas serão sempre aproximadas.

Exemplo: Peso ao desmame dos coelhos

Dados	Média	Moda	Mediana
Brutos	815,9g	1000,0g	826,5g
Distr. frequências	820,0g	831,7g	828,5g