Medidas de Tendência Central

Introdução
Média Aritmética
Moda
Mediana

Introdução

- A maioria dos dados apresenta uma tendência de se concentrar em torno de um ponto central

- Portanto, é possível selecionar um valor que melhor descreva o conjunto

- Este valor é uma medida de tendência central

Introdução

- Há vários tipos de medidas utilizadas como medida de tendência central. Nós estudaremos as medidas:
  - Média aritmética
  - Moda
  - Mediana

Média Aritmética Simples

- Tipo de medida de tendência central mais utilizada

- É a soma dos valores de todas as observações dividida pelo número de observações envolvidas

- Perigo: um ou mais valores bastante discrepantes do conjunto podem distorcer a tendência apresentada pela média
  - Esta distorção pode ser amenizada aplicando-se pesos às observações (média aritmética ponderada)

Média Aritmética Simples

- A média aritmética pode ser escrita como:

$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

- Ou, de forma simplificada:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Média Aritmética Simples

- OBS: normalmente trabalha-se com a média da amostra $\overline{X}$ e não com a média da população μ devido ao custo e dificuldade de cálculo desta medida

Média Aritmética Simples

- Exercícios
  - Dada uma amostra das notas dos alunos da disciplina de estatística, calcule a média aritmética:

    {5.0, 6.5, 5.5, 8.0, 7.5, 6.0, 5.1, 7.0}

  - O que aconteceria com a média se a nota 0.1 fosse incluída na amostra?

Média Aritmética Simples

- Propriedades

  1- A soma dos desvios em relação à média é sempre igual a zero

x_i	$d_i = x_i - \bar{x}$
5	-1,325 = 5 - 6,325
6,5	0,175 = 6,5 - 6,325
5,5	-0,825 = 5,5 - 6,325
8	1,675 = 8 - 6,325
7,5	1,175 = 7,5 - 6,325
6	-0,325 = 6 - 6,325
5,1	-1,225 = 5,1 - 6,325
7	0,675 = 7 - 6,325

$$\sum d_i = 0$$

Média Aritmética Simples

- Propriedades

  2- A soma do quadrado dos desvios em relação à média é chamado desvio mínimo, valor utilizado em otimizações e regressões

x_i	$d_i = x_i - \bar{x}$	d_i^2
5	-1,325 = 5 - 6,325	1,75
6,5	0,175 = 6,5 - 6,325	0,03
5,5	-0,825 = 5,5 - 6,325	0,68
8	1,675 = 8 - 6,325	2,8
7,5	1,175 = 7,5 - 6,325	1,38
6	-0,325 = 6 - 6,325	0,11
5,5	-1,225 = 5,1 - 6,325	1,5
7	0,675 = 7 - 6,325	0,45

$$\sum d_i^2 \approx 8,7$$

Média Aritmética Simples

- Propriedades

3- Se for somada (ou subtraída) uma constante **K** a cada elemento da amostra, a média aritmética será também somada (ou subtraída) a esta constante

x_i	$x_i + 5$
5	10
6,5	11,5
5,5	10,5
8	13
7,5	12,5
6	11
5,1	10,1
7	12

$\bar{x} = 6{,}325$ $\bar{x} = 11{,}325$

Média Aritmética Simples

- Propriedades

  4- Se for multiplicada (ou dividida) uma constante **K** a cada elemento da amostra, a média aritmética será também multiplicada (ou dividida) por esta constante

➔ Exercício: demonstre esta propriedade!

Média Aritmética Ponderada

- Caso os dados se repitam, para calcular a média pode-se fazer a somatória da multiplicação de cada valor pela respectiva freqüência e dividir pelo total de valores

$$\overline{x} = \frac{\sum x_i * f_i}{\sum f_i}$$

- Esta fórmula é uma média aritmética ponderada pela freqüência
- É equivalente à média aritmética simples

Média Aritmética

- Exercícios:
  - Demonstre que a média aritmética simples e a ponderada (por freqüência) são equivalentes

  - Insira nos dados da tabela do exercício anterior um valor repetido e calcule a média aritmética simples e a ponderada

Moda

- Moda é o valor que aparece mais freqüentemente em um conjunto de dados
- Ao contrário da média aritmética, a moda não é afetada por valores extremos
- É utilizada para fins descritivos apenas, uma vez que é, dentre as medidas de tendência, a mais variável de amostra para amostra

Moda

- Moda em dados não tabulados

$$X=\{4, 2, 4, 5, 5, \mathbf{6, 6, 6}, 7, 8, 9\}$$

$$\text{Moda}=6$$

OBS: Amostras podem possuir apenas uma moda (unimodal), duas modas (bimodal), mais de duas modas (multimodal), ou nenhuma moda (amodal)

- Exercício: Dê exemplos dos casos citados acima

Moda

- Moda em dados tabulados
  - Método de Czuber (considerado o mais preciso)

$$M_0 = \ell_i + c\,\frac{f_{mo} - f_{ant}}{2f_{mo} - (f_{ant} + f_{post})}$$

onde:

- l_i é o limite inferior da classe modal
- c é o intervalo de classe
- f_{mo} é a freqüência da classe modal
- f_{ant} é a freqüência anterior à classe modal
- f_{post} é a freqüência posterior à classe modal

Moda

- Moda em dados tabulados
  - Exemplo

Idades	Freqüência
10 —19	10
20 —29	20
<u>30 —39</u>	**<u>40</u>**
40 —49	20
50 —59	10

$$M_0 = 30 + 9\frac{40 - 20}{2*40 - (20 + 20)}$$

$$M_0 = 34,5$$

Moda

- Exercício: Alterado os dados da tabela, recalcule a moda

Idades	Freqüência
10 —19	10
20 —29	50
30 —39	30
40 —49	20
50 —59	10

Mediana

- Medida de tendência central que divide uma série ordenada de dados (ROL) em duas partes iguais

- Ocupa a posição central em um ROL

- A mediana também não é afetada por valores extremos

Mediana

- Mediana em dados não tabulados

  – Amostra com número ímpar de elementos

  $X=\{1, 3, 5, 7, 8, 11, 12, 13,14\}$, onde $n=9$

Calcula-se o elemento central (E)

$$E = \frac{n+1}{2} = \frac{9+1}{2} = 5$$

Logo a mediana corresponde ao 5º elemento da amostra: $M_d = 8$

Mediana

- Mediana em dados não tabulados
  - Amostra com número par de elementos

X={1, 3, 5, 7, 9, 11}, onde n=6

Calcula-se os elementos centrais (E)

$$E=\frac{n}{2}=\frac{6}{2}=3$$

Os elementos centrais são 5 e 7. Logo a mediana é a média aritmética dos mesmos: M_d = 6

Mediana

- Mediana em dados tabulados
  - Amostra com dados discretos pares e não agrupados em classes

Custo de produção (em milhões)	Freqüência	Freqüência acumulada
2	5	5
4	10	15
6	15	30
8	12	42
10	5	47
12	3	50

$$E = \frac{n}{2} = \frac{50}{2} = 25$$

Os elementos centrais são 25 e 26 (já que se trata de uma amostra par), que estão entre 15 e 30. Logo a mediana é a média aritmética: $M_d = 6$

Mediana

- Mediana em dados tabulados
  - Amostra com dados discretos ímpares e não agrupados em classes

Custo de produção (em milhões)	Freqüência	Freqüência acumulada
2	5	5
4	10	15
6	15	30
8	12	42
10	5	47

$$E = \frac{n+1}{2} = \frac{48}{2} = 24$$

O elemento central é 24, que está entre 15 e 30. Logo a mediana é a média aritimética: $M_d = 6$

Mediana

- Mediana em dados tabulados
  - Amostra com dados contínuos agrupados em classes

$$M_d = \ell_i + c\,\frac{E - f_{ant\,ac}}{f_{md}}$$

onde:

- ℓ_i é o limite inferior da classe mediana
- c é o intervalo de classe
- f_{md} é a freqüência da classe mediana
- $f_{ant\,ac}$ é a freqüência acumulada anterior à classe mediana

Mediana

- Mediana em dados tabulados
  - Amostra com dados contínuos agrupados em classes

Custo de produção (em milhões)	Freqüência	Freqüência acumulada
10 —19	20	20
20 —29	30	50
30 —39	30	80
40 —49	20	100
50 —59	10	110

$$E = \frac{n}{2} = \frac{110}{2} = 55 \longrightarrow M_d = 30 + 9\,\frac{55 - 50}{30} = 31{,}5$$

Mediana

- Exercício
  – Calcule a mediana para o rol de dados

Custo de produção (em milhões)	Freqüência	Freqüência acumulada
10 —19	50	50
20 —29	10	60
30 —39	80	140
40 —49	90	230
50 —59	60	290

Mediana

- Exercício:

Vamos coletar a idade de 30% dos alunos desta sala, tabulá-los e dividi-los em classes. Em seguida, vamos calcular a média, moda, mediana, e comparar estes valores.

OBS: A partir do resultado obtido, vamos introduzir o conceito de simetria.

Moda, Mediana e Média

- A comparação de média, mediana e moda define a simetria dos dados

- A distribuição de dados é simétrica quando a moda, média e mediana são coincidentes

- A distribuição é assimétrica à esquerda (negativamente assimétrica) quando a média e a mediana estão à esquerda da moda

- A distribuição é assimétrica à direita (positivamente assimétrica) quando a média e a mediana estão à direita da moda

Moda, Mediana e Média

- Exercícios:
  - Ilustre graficamente as medidas de tendência central e verifique se há assimetria ou simetria
  - A distribuição de rendas anuais de um país tende a ter uma assimetria positiva ou negativa?

Obrigado!

Até a próxima aula!