

A) Medidas de tendência central

Considere os seguintes dados brutos: 1, 2, 3, 1, 1, 6, 5, 4, 1, 4, 4, 3.

i) Qual é o valor que ocorre com maior frequência?

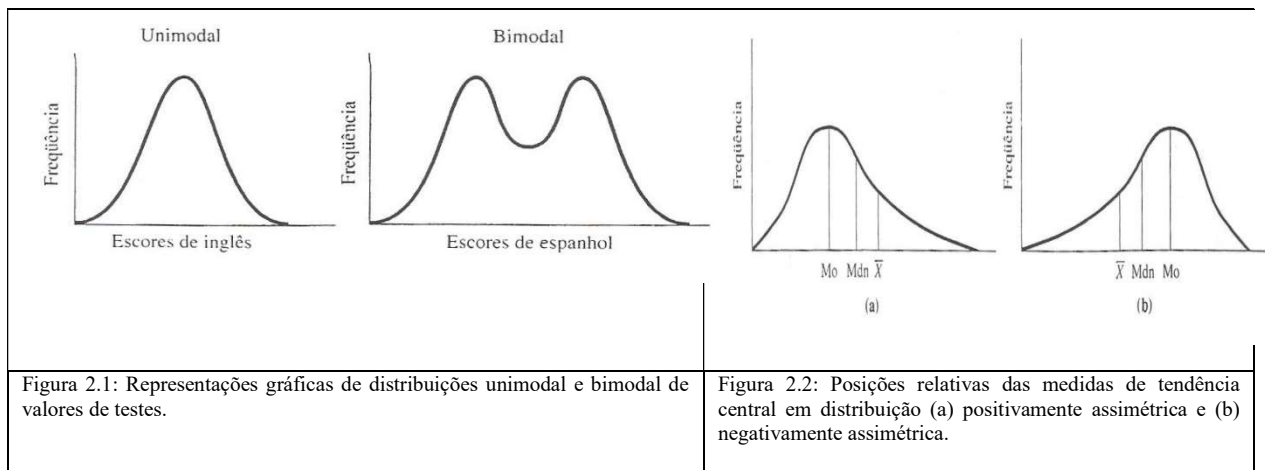
O valor que ocorre com maior frequência é a moda (Mo). Neste exemplo a moda é 1 (1 é o valor que ocorre com maior frequência).

ii) Escreva em ordem crescente ou decrescente os dados seguintes: 11, 25, 12, 20, 13, 17, 16.

iii) Identifique o valor mais ao centro (que divide a distribuição em duas partes iguais) da distribuição.

O valor mais ao centro da distribuição (que divide a distribuição em duas partes iguais) é a mediana (Mdn), o ponto do meio de uma distribuição. Neste exemplo a mediana é 16.

Quando temos um ímpar de elementos, a mediana é o caso que recai exactamente no meio da distribuição.



A posição da mediana pode ser determinada por inspecção ou pela fórmula:

$$\text{Posição da mediana} = \frac{N + 1}{2}$$

iv) Determine a posição da mediana dos seguintes dados: 17, 11, 26, 12, 25, 13, 20, 16.

iv.1) Disponha-os em ordem; iv.2 use a fórmula para determinar o lugar onde se situa a mediana.

Neste exemplo a mediana situa-se a meio caminho entre o quarto e o quinto valor. Neste caso $N=8$ (que é um número par). Portanto, somamos os dois valores centrais e dividimos por dois para obter

o valor da mediana. Neste exemplo a mediana é $16,5 \left(= \frac{16 + 17}{2} \right)$.

v) Separatrizes: quartil, decil, percentil

.....

vi) Calcule a média aritmética $\bar{X}$ para cada uma das distribuições anteriores.

Para obtê-la, soma-se um conjunto de valores e divide-se a soma pelo número de valores. A

fórmula: $\bar{X} = \frac{\sum X}{N}$

Onde $\bar{X}$ = média (lê-se 'X barra')

$\sum$ = soma (letra grega maiúscula sigma)

X= valor bruto num conjunto de valores

N= número total de valores no conjunto.

A moda (Mo), mediana (Mdn) e a média ($\bar{X}$) são medidas de tendência central. Elas consistem em números únicos calculados que representam o que é médio ou típico dum conjunto de dados. A maneira conveniente para descrever um grupo como um todo, em torno duma questão em estudo, é achar as medidas de tendência central.

Podemos calcular uma “*média de médias*” isto é, calcular uma média total para vários grupos diferentes. Por exemplo, temos três turmas, cada uma com a sua média e seu número de alunos:

Turma1: $\bar{X}_1 = 85$; $N_1 = 28$; Turma2: $\bar{X}_2 = 72$; $N_2 = 28$ Turma3: $\bar{X}_3 = 79$; $N_3 = 28$.

Como todas turmas têm o mesmo número de alunos calculamos média total (a média de médias)

usando a fórmula: $\frac{\bar{X}_1 + \bar{X}_2 + \bar{X}_3}{3} = \frac{85 + 72 + 79}{3} = 78,67$

Na maioria dos casos os grupos têm tamanhos diferentes. Neste caso, devemos ponderar cada média do grupo pelo tamanho do grupo (N). A média ponderada pode ser calculada multiplicando-se primeiro cada média de grupo por seu valor de N respectivo e dividindo-se a soma desses produtos pelo número total em todos grupos.

$$\bar{X}_p = \frac{\sum N_{\text{grupo}} \bar{X}_{\text{grupo}}}{N_{\text{total}}}$$

Onde:

$\bar{X}_{\text{grupo}}$ = média de um grupo determinado

N_{grupo} = número em um grupo determinado

N_{total} = número em todos os grupos combinados

$\bar{X}_p$ = média ponderada

Por exemplo, temos as seguintes turmas com as respectivas as médias e tamanhos:

Turma1: $\bar{X}_1 = 85$; $N_1 = 95$; Turma2: $\bar{X}_2 = 72$; $N_2 = 25$ Turma3: $\bar{X}_3 = 79$; $N_3 = 18$. A média ponderada é calculada com base na fórmula seguinte.

$$\begin{aligned}
\bar{X}_p &= \frac{N_1 \bar{X}_1 + N_2 \bar{X}_2 + N_3 \bar{X}_3}{N_{total}} \\
&= \frac{95(85) + 25(72) + 18(79)}{138} \\
&= \frac{8.075 + 1.800 + 1.422}{138} \\
&= \frac{11.297}{138} \\
&= 81,86
\end{aligned}$$

Deste modo a média ponderada para todas turmas combinadas, foi 81,86.

vii) *Decisão na escolha duma medida de tendência central para uma situação de pesquisa particular*

A decisão envolve diversos factores, incluindo:

1. Nível de mensuração;
2. Forma da distribuição dos dados e;
3. Objectivo da pesquisa.

Como a *moda (Mo)* exige apenas uma contagem de frequências, pode ser aplicada a qualquer conjunto de dados, seja no nível nominal, ordinal ou intervalar de mensuração.

O cálculo da *mediana (Mdn)* exige a ordenação das categorias. Assim, a mediana só pode ser obtida para dados ordinais ou intervalares, mas não para dados nominais. Não tem sentido calcular a mediana da filiação religiosa, do sexo ou do país de origem porque não se pode fazer uma disposição em postos ou um escalonamento.

O uso da *média ($\bar{X}$)* está restrito exclusivamente a dados intervalares.

A forma de distribuição é outro factor que pode influenciar o pesquisador na escolha duma medida de tendência central. Quando o pesquisador social lida com uma distribuição simétrica, sua escolha da medida de tendência central se baseia principalmente nos objectivos específicos da pesquisa e no nível em que os dados são medidos. A média tende a ser preferida em relação a mediana porque pode ser usada facilmente em análises estatísticas mais avançadas. Numa distribuição assimétrica, a mediana se situa sempre entre a média e a moda. É essa característica que torna a mediana a medida mais conveniente de tendência central para descrever uma distribuição assimétrica de valores. A mediana tende a proporcionar uma representação equilibrada de valores extremos. Para descrever adequadamente as distribuições bimodais, em geral é conveniente identificar ambas as modas, pois utilizando apenas a mediana ou a média poderemos obscurecer aspectos importantes dessas distribuições.

viii) *Cálculo da moda, mediana e média duma distribuição de frequências simples*

Numa distribuição de frequências simples a moda é o valor que aparece com maior incidência na coluna de frequências da tabela. Observe a tabela 4. Usa-se a fórmula aprendida anteriormente para achar a posição da mediana. Para os dados da tabela 4 a mediana situa-se no 13º lugar.

$$\text{Posição da mediana} = \frac{N+1}{2} = \frac{25+1}{2} = 13$$

Vemos que o valor que se encontra no 13º (na coluna de frequências acumuladas) é 22. Esta é a mediana para os dados da tabela 4. Começando com o valor mais baixo (18), adicionamos as frequências até atingir um valor que represente o décimo terceiro valor na distribuição, o que se consegue procurando o menor valor que tenha uma frequência acumulada de, no mínimo, 13. A média para uma distribuição de frequências simples é dada pela fórmula:

$$\bar{X} = \frac{\sum fX}{N}$$

Onde:

$\bar{X}$ = média

X = um valor do escore na distribuição

f = frequência da ocorrência de X

N = número total de escores

$$\bar{X} = \frac{\sum fX}{N} = \frac{575}{25} = 23$$

Tabela 4: Idade ao primeiro casamento de 25 adultos

	X	f	fa	fX
	31	1	25	31
	30	1	24	30
	29	1	23	29
	28	0	22	0
	27	2	22	54
	26	3	20	78
	25	1	17	25
	24	1	16	24
	23	2	15	46
	22	2	13	44
	21	2	11	42
	20	3	9	60
Mo ->	19	4	6	76
	18	2	2	36

$$\sum fX = 575$$

ix) *Cálculo da moda, mediana e média numa distribuição de frequências agrupadas*

Numa distribuição de frequências agrupadas, a *moda* é o ponto médio do intervalo de classe que tem a maior frequência. A moda é 19 para os dados da tabela 5.

Podemos determinar a mediana a partir da fórmula abaixo.

$$Mediana = I + \left(\frac{N/2 - fa_h}{f} \right) h$$

Onde:

Tabela 5: Idades agrupadas de adultos no primeiro casamento

Intervalo de classe	<i>m</i>	<i>f</i>	<i>fa</i>	<i>fm</i>
30 - 32	31	2	25	62
27 - 29	28	3	23	84
24 - 26	25	5	20	125
21 - 23	22	6	15	132
18 - 20	19	9	9	171
		N= 25	$\Sigma fm = 574$	

N = número de casos na distribuição

fa_h = frequência acumulada (numericamente) abaixo do limite inferior do intervalo crítico

I = limite inferior do intervalo crítico

f = frequência dentro do intervalo crítico

h = tamanho do intervalo de classe

$$Mediana = I + \left(\frac{N/2 - fa_h}{f} \right) h = 20,5 + \left(\frac{12,5 - 9}{6} \right) 3 = 20,5 + 1,75 = 22,25$$

A mediana para os dados da tabela 5 é 22,25.

Para obter a média duma distribuição de frequências agrupadas usamos a fórmula seguinte:

$$\bar{X} = \frac{\sum fm}{N}$$

Onde:

$\bar{X}$ = média

m = ponto médio de um intervalo de classe

f = frequência de um intervalo de classe

N = número total de escores

$$\bar{X} = \frac{\sum fm}{N} = \frac{574}{25} = 22,96$$

Para os dados da tabela 5 a média é 22,96