

6. MEDIDAS DESCRITIVAS DE DADOS OBSERVADOS: TENDÊNCIA CENTRAL E DE POSIÇÃO

1. MÉDIA ARITMÉTICA (ou simplesmente Média)

1.1. Somatória de todos os valores (x_i) da amostra dividido pelo número de valores (n):

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

1.2. Medida afetada fortemente por “outliers” – valores discrepantes

1.2.1. Distância da cidade de origem (a maioria é regional, mas um é de Manaus)

1.3. Obs: quando a média se refere à população, utiliza-se a notação:

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

1.4. Caso especial: acidentes com operadores utilizando ou não capacetes:

1.4.1. Encontre a média amostral da classificação de severidade de acidentes para motociclistas usando capacete e não usando capacete (obs: acidentes similares)

Classification of accident	Interpretation	Frequency of driver with helmet	Frequency of driver without helmet
0	No head injury	248	227
1	Minor head injury	58	135
2	Moderate head injury	11	33
3	Severe, not life-threatening	3	14
4	Severe and life-threatening	2	3
5	Critical, survival uncertain at time of ac	8	21
6	Fatal	1	6
		331	439

Solution

The sample mean for those wearing helmets is

$$\bar{x} = \frac{0.248 + 1.58 + 2.11 + 3.3 + 4.2 + 5.8 + 6.1}{331} = \frac{143}{331} = 0.432$$

The sample mean for those who did not wear a helmet is

$$\bar{x} = \frac{0.227 + 1.135 + 2.33 + 3.14 + 4.3 + 5.21 + 6.6}{439} = \frac{396}{439} = 0.902$$

Conclusão: a severidade é maior para aqueles que não utilizam capacete (...)

2. MEDIANA (ou valor do meio)

2.1. É o valor que está em posição intermediária na distribuição de valores, isto é, tem tantos elementos acima quanto abaixo de sua posição. Representada por **M** ou \tilde{x}

2.2. Para determinar-se a Mediana, primeiramente deve-se ordenar os **n** valores em ordem crescente (**n** é o número de elementos da amostra)

2.3. Quando **n** é ímpar, a Mediana é um valor real da amostra:

2.3.1. A Mediana é o valor que está na posição:

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

2.3.2. Exemplo: 33 elementos na amostra ($n=33$) – a mediana está na posição 17

2.4. Quando ***n*** é par, a Mediana é um valor fictício entre dois valores reais da amostra

2.4.1. O valor da Mediana é dado pela média dos valores intermediários:

$$M = \frac{Vp_{\frac{n}{2}} + Vp_{\frac{n}{2}+1}}{2}$$

2.4.2. Exemplo: 34 elementos na amostra ($n=34$) – a mediana é a média dos valores que estão nas posições 17 e 18.

3. MODA (ou valor mais frequente)

3.1. É o valor que aparece mais vezes nas observações da amostra.

3.2. Uma amostra pode ser:

3.2.1. Amodal: não ter moda (não há valores repetidos)

3.2.2. Unimodal: apenas um valor com repetições expressiva (homogênea)

3.2.3. Bimodal: dois valores têm repetições expressivas (heterogênea)

3.2.4. Trimodal: três valores têm repetições expressivas (heterogênea)

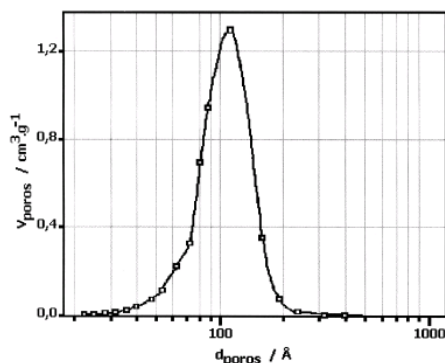


Figura 8: Curva BJH da solução sólida zircônia-ítria.
[Figure 8: BJH curve of the zirconia-yttria solid solution.]

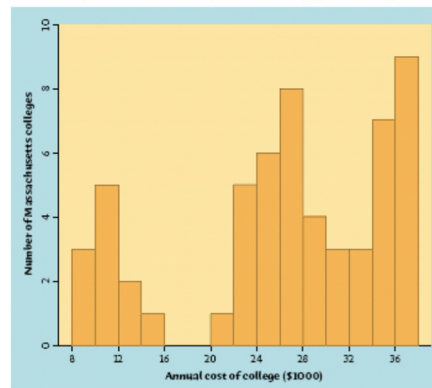


Figure 15 Histogram of the estimated costs (in thousands of dollars) for four-year colleges in Massachusetts. The two clusters distinguish public from private institutions.

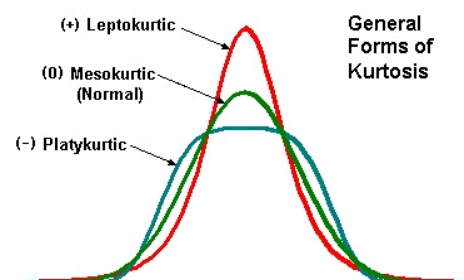
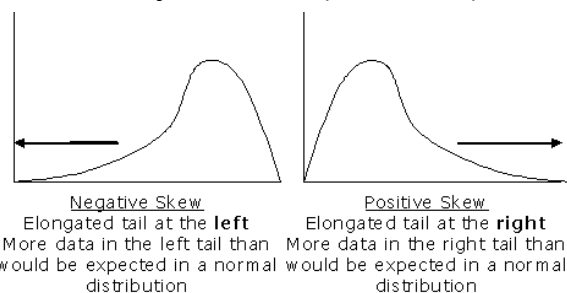
4. RELAÇÃO MÉDIA – MEDIANA

4.1. Distribuições simétricas têm média e mediana próximas

4.2. Distribuições assimétricas tem média e mediana mais distantes

4.3. Inclinação à esquerda (lado maior): média menor que mediana

4.4. Inclinação à direita (lado maior): média maior que mediana



5. Medidas de Posição: PERCENTIL, DECIL, QUARTIL

5.1. São utilizadas para descrever a localização de uma observação particular em relação ao restante do conjunto de dados

5.1.1. Percentil divide as observações (dados) em 100 partes iguais

5.1.2. Decil divide as observações (dados) em 10 partes iguais

5.1.3. Quartil divide as observações em 4 partes iguais

5.1.4. Mediana: divide as observações em 2 partes iguais

5.2. Inicialmente é necessário ordenar as ***n*** observações em ordem crescente (ranking)

6. PERCENTIL

6.1. O percentil para a observação x é encontrado dividindo-se o número de observações que são menores que x pelo número total de observações e multiplicando-se por 100

6.1.1. Se o valor encontrado for inteiro, este é o percentil

6.1.2. Se o valor encontrado não for inteiro, deve ser arredondado ao superior.

6.2. Significado do percentil de uma observação:

6.2.1. Percentil 25: 25% dos valores são menores e ~75% são maiores (grandes n 's)

6.3. Exemplo: encontrar o percentil de 66 nos valores abaixo

43, 54, 56, 61, 62, 66, 68, 69, 69, 70, 71, 72, 77, 78, 88, 89, 93, 98, 99, 99.

6.3.1. Temos 20 valores nessa sequência – o valor 66 tem 5 valores abaixo dele:

$$P = \frac{n^{\circ} \text{ elementos} < \text{que o valor}}{n^{\circ} \text{ valores}} \times 100$$

$$P = \frac{5}{20} \times 100 = P25$$

6.4. Para encontrar o percentil P de uma distribuição, multiplica-se o valor do percentil pelo número de observações da amostra e divide-se por 100 (obs. percentil: $n \gg 100$)

6.4.1. Se o valor encontrado for inteiro, o percentil é a média da observação que está nessa posição do ranking com o valor do próximo no ranking

6.4.2. Se o valor encontrado não for inteiro, arredonda-se para o superior mais próximo e procura-se a observação que está nessa posição no ranking.

$$n^{\circ} \text{ elementos} < \text{que o valor} = \frac{P \times n^{\circ} \text{ valores}}{100}$$

$$n^{\circ} \text{ elementos} < \text{que o valor} = \frac{90 \times 20}{100} = 18 \text{ elementos menores} = \frac{98+99}{2} = 98,5$$

6.5. Interpretação do Percentil

6.5.1. 10% renda menor que 10 mil

6.5.2. 20% renda menor que 17 mil

6.5.3. 50% renda menor que 42 mil

6.5.4. 80% renda menor que 83 mil

6.5.5. 90% renda menor que 116 mil

6.5.6. 95% renda menor que 150 mil

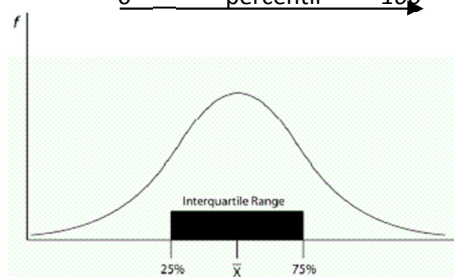
Table 2-1 U.S. Household Income for 2001	
Percentile	2001 Household Income
10th	\$ 10,913
20th	\$ 17,970
50th	\$ 42,228
80th	\$ 83,500
90th	\$ 116,105
95th	\$ 150,499

6.6. Cálculo do Percentil:

6.6.1. Percentil do valor = $100 \times \frac{n.\text{ordem}-1}{n-1}$

1 — n. de ordem — n

0 — percentil — 100



7. DECIL – QUARTIL – MEDIANA

7.1. Posições especiais de Quartil

7.2. Decil: 10% – 20% – 30% etc.

7.3. Quartil: 25% - 75%

7.4. Mediana: 50%

7.5. Cálculo do Quartil – regra prática

7.5.1. Calcular mediana – separar os valores inferiores e os superiores à mediana

7.5.2. Quartil 1 (25%): calcular a mediana dos valores inferiores

7.5.3. Quartil 3 (75%): calcular a mediana dos valores superiores

