UnB – Universidade de Brasília
FGA – Faculdade UnB Gama
Graduação – ciclo básico

Probabilidade e Estatística aplicada à Engenharia

Profa. Marília Miranda

- **Medidas de posição e dispersão para dados agrupados (Parte II)**

RELEMBRANDO...

- Dados brutos: 24, 26, 24, 21, 27, 27, 30, 41, 32, 38

Classe	Freqüência
15 $\vdash$ 25	3
25 $\vdash$ 35	5
35 $\vdash$ 45	2

UnB – Universidade de Brasília
FGA – Faculdade UnB Gama
Graduação – ciclo básico

Probabilidade e Estatística aplicada à Engenharia

Profa. Marília Miranda

- **Diagrama de Caixas (Box-plot)**
- **Coeficiente de Assimetria**

FORMA

- Descreve como os dados estão distribuídos.
- Medidas de forma:
  - Simétrica: média = mediana = moda
  - Assimétrica:

(1) à esquerda ou negativa:

média < mediana < moda

(2) à direita ou positiva:

MEDIDAS PARA TABELAS DE FREQUÊNCIA

- Média e Desvio Padrão ponderados pelas frequências:

$$\overline{X} = \frac{\sum(X \cdot f)}{n}$$

$$S = \sqrt{\frac{\sum\left(X^2 \cdot f\right) - n \cdot \overline{X}^2}{n-1}}$$

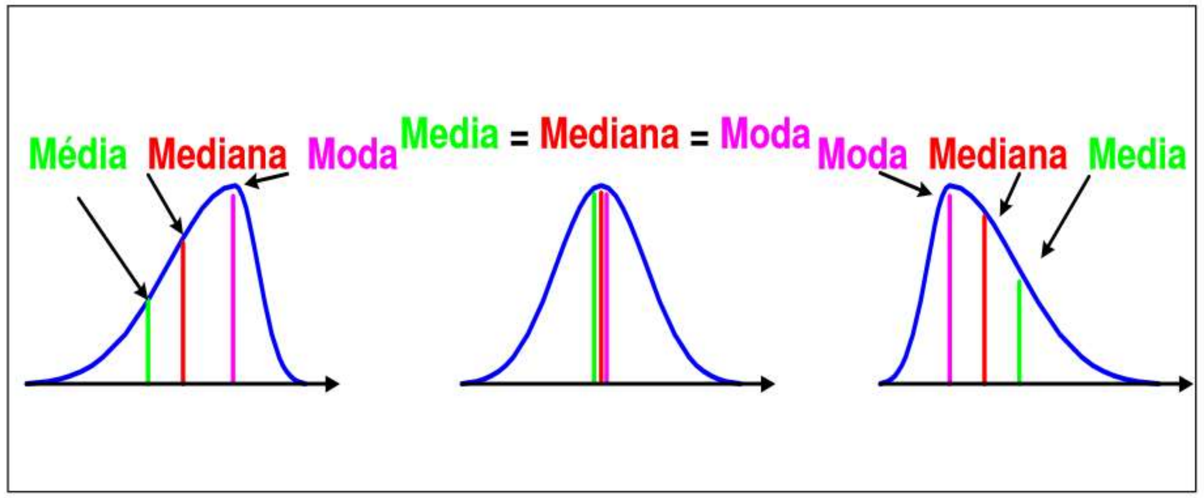
COEFICIENTE DE VARIAÇÃO

- Medida de dispersão relativa;
- Pode ser expresso como uma %;
- Mostra a variação relativa a média;
- Usado para comparar 2 ou mais grupos.
- Fórmula:

$$CV = \left(\frac{S}{\overline{X}}\right) \cdot 100\%$$

média > mediana > moda

FORMA



ESQUEMA DOS CINCO NÚMEROS

- Para ter uma idéia melhor da assimetria de um conjunto de dados, tomamos as seguintes medidas:

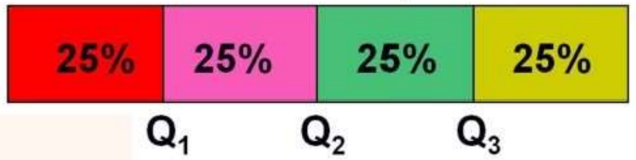
  - Mediana (Md)
  - Extremos (o menor e o maior valor do conjunto de dados)
  - Quartis (Q_1 e Q_3)

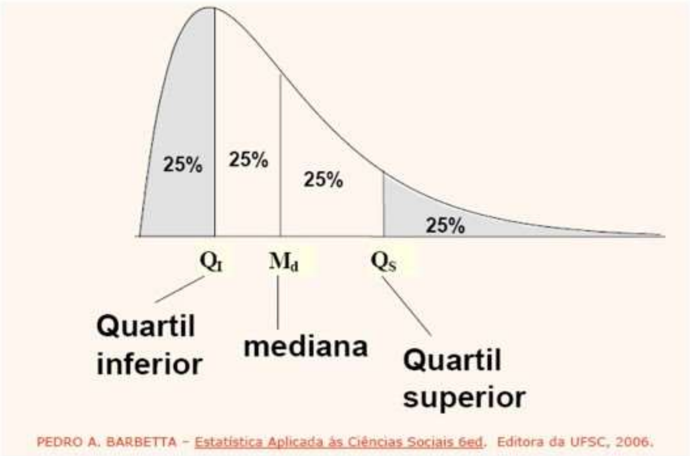
	$n = 40$		
Md		79	
Q	65,5		87,5
E	45		95

MEDIDAS PARA TABELAS DE FREQUÊNCIA

QUARTIS

- Divide os dados ordenados em 4 quartos





MEDIDAS PARA TABELAS DE FREQUÊNCIA

- Determinação dos quartis:

1º) Calcula-se a POSIÇÃO do quartil Qi:

Pos(Q1) = n/4

Pos(Q2) = 2n/4

Pos(Q3) = 3n/4

2º) Identifica-se a classe que o quartil pertence (olharemos para f_i)

MEDIDAS PARA TABELAS DE FREQUÊNCIA

DIAGRAMA DE CAIXAS (BOX-PLOT)

- Apresentação gráfica de dados usando o esquema dos 5 números;
- Maneira de apresentar aspectos relevantes de uma distribuição de frequências;
- Tem a vantagem de não ser tão sensível a valores extremos;
- Explica as propriedades numéricas dos dados;

DIAGRAMA DE CAIXAS (BOX-PLOT)

- Descreve medidas resumo:

  - Tendência central
  - Variação
  - Forma

- Analisa dados quantitativos usando medidas resumo.

DIAGRAMA DE CAIXAS (BOX-PLOT)

3º) Aplicamos a fórmula para encontrar o valor do quartil desejado:

$$Qi = a + h * \left[\frac{b - c}{d}\right]$$

Onde:

a = limite inferior da classe que pertence Qi

b = posição do quartil Qi

c = somatório das f_i anteriores à classe que pertence Qi

d = fi da classe que pertence Qi

h = amplitude da classe que pertence Qi

EXEMPLO

Classes	f_i
7 |--- 17	6
17 |--- 27	15
27 |--- 37	20
37 |--- 47	10
47 |--- 57	5
n	*56*

EXEMPLO

- Cálculo do primeiro quartil = Q1

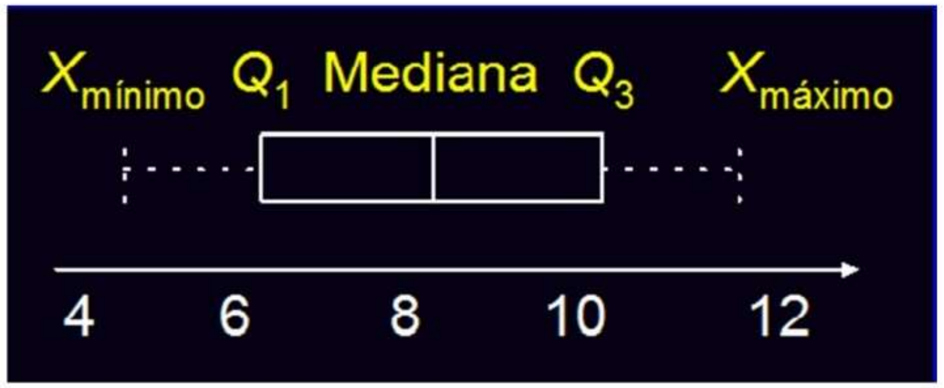
1º) Pos(Q1) = n/4 = 56/4 = 14

(BOX-PLOT)

- É um gráfico de dados que consiste em uma reta que se prolonga do menor ao maior valor, e um retângulo com retas traçadas no primeiro quartil (Q_1), na mediana e no terceiro quartil (Q_3);
- Representação gráfica bastante útil para a comparação de dois ou mais conjuntos de dados.

- obs.: é importante utilizarmos a mesma escala, de forma a possibilitar a comparação.

DIAGRAMA DE CAIXAS (BOX-PLOT)



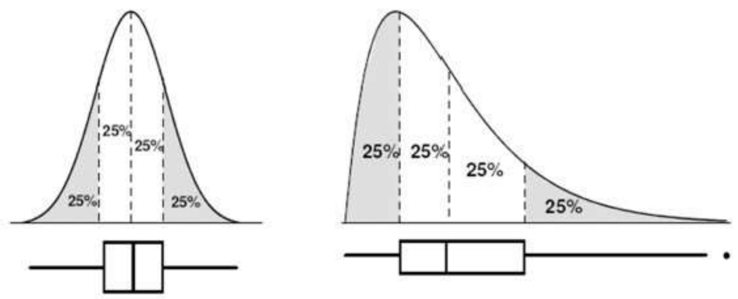


DIAGRAMA DE CAIXAS (BOX-PLOT)

1º) Pos(Q1) = n/4 = 56/4 = 14

2º) classe: 17 |--- 27

3º)

$$\boldsymbol{Q1} = 17 + 10 * \left[\frac{14-6}{15}\right] = 22{,}3$$

EXEMPLO

- Cálculo do segundo quartil = Q2 (MEDIANA)

1º) Pos(Q1) = 2n/4 = (2*56)/4 = 28

2º) classe: 27 |--- 37

3º)

$$\boldsymbol{Q}2 = 27 + 10 * \left[\frac{28-21}{20}\right] = 30{,}5$$

EXEMPLO

- Cálculo do terceiro quartil = Q3

1º) Pos(Q1) = 3n/4 = (3*56)/4 = 42

2º) classe: 37 |--- 47

3º)

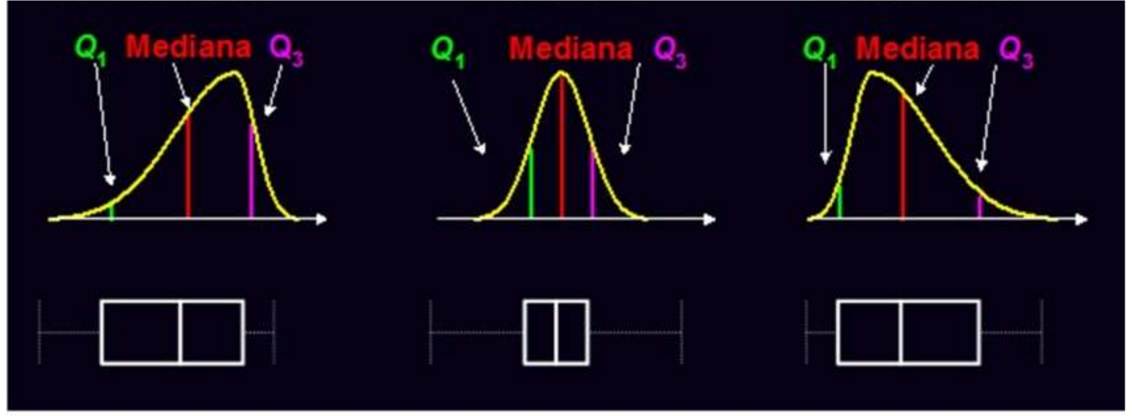


DIAGRAMA DE CAIXAS (BOX-PLOT)

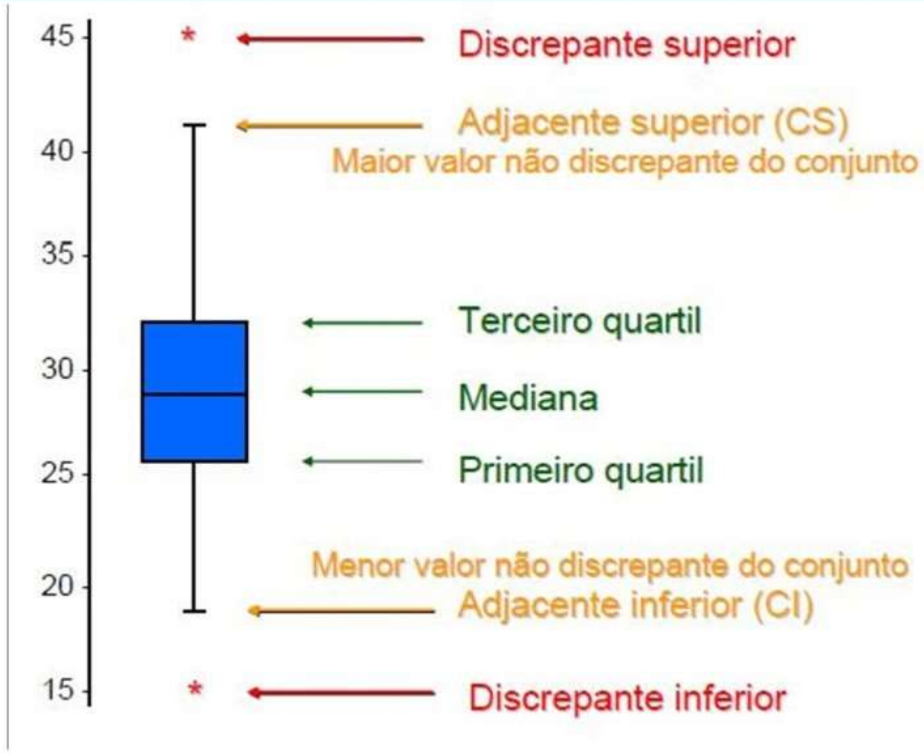
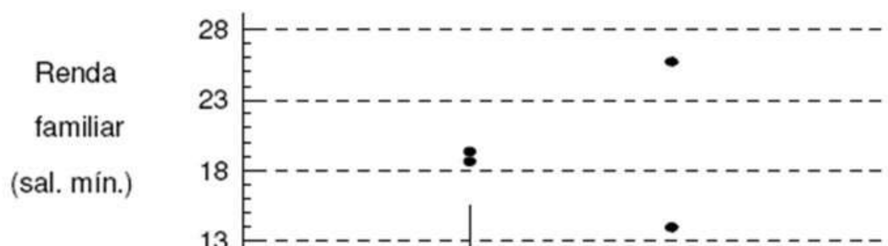


DIAGRAMA DE CAIXAS (BOX-PLOT)



3º)

$$Q3 = 37 + 10 * \left[\frac{42-41}{10}\right] = 38$$

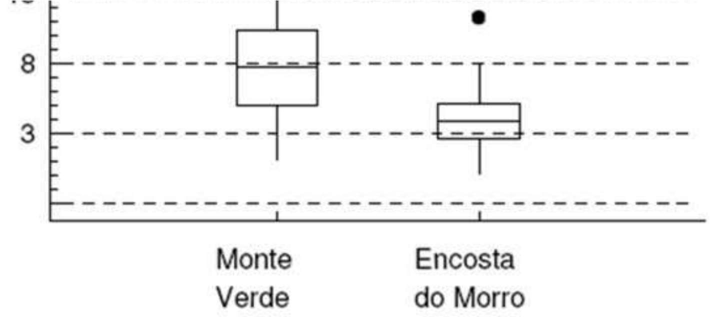


DIAGRAMA DE CAIXAS (BOX-PLOT)

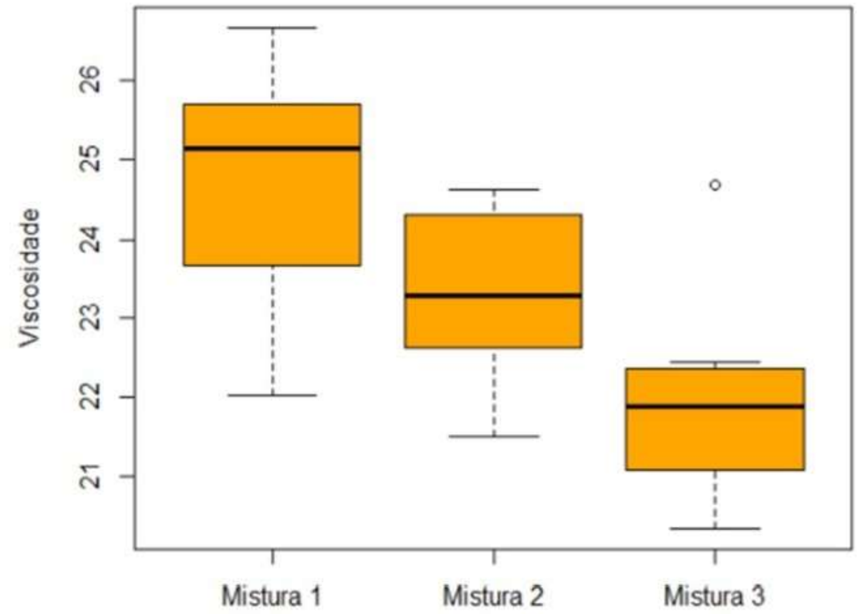
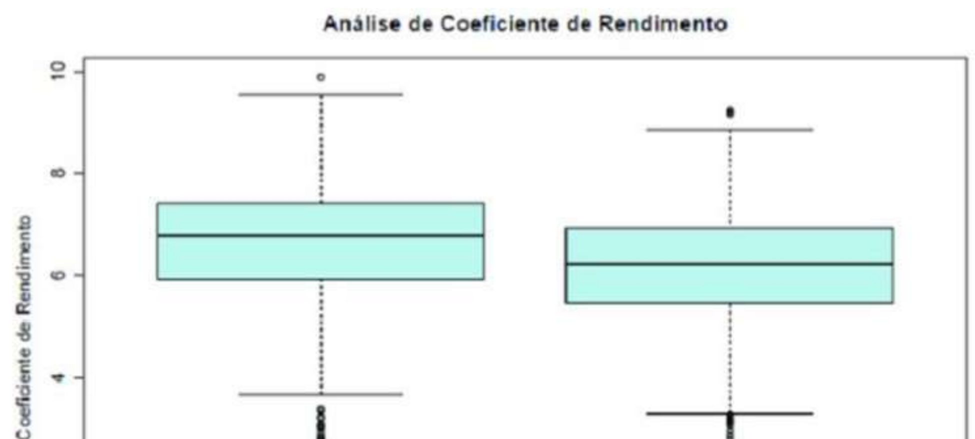


DIAGRAMA DE CAIXAS (BOX-PLOT)



Fonte: Bispo (2016)

DIAGRAMA DE CAIXAS (BOX-PLOT)

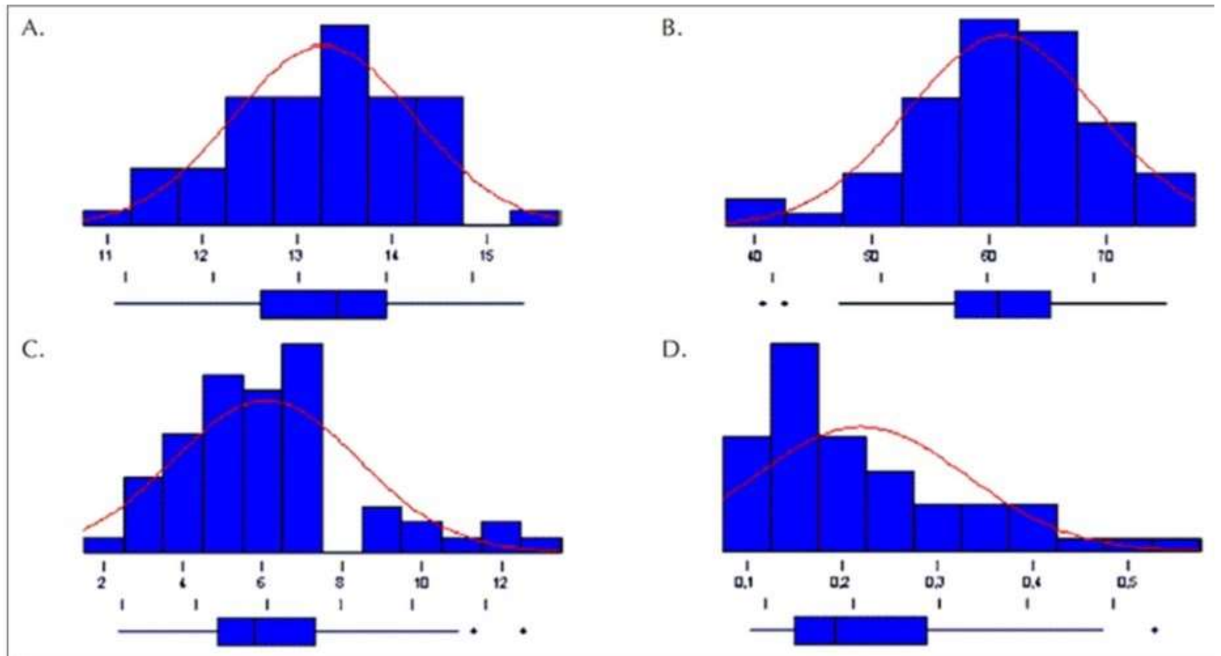
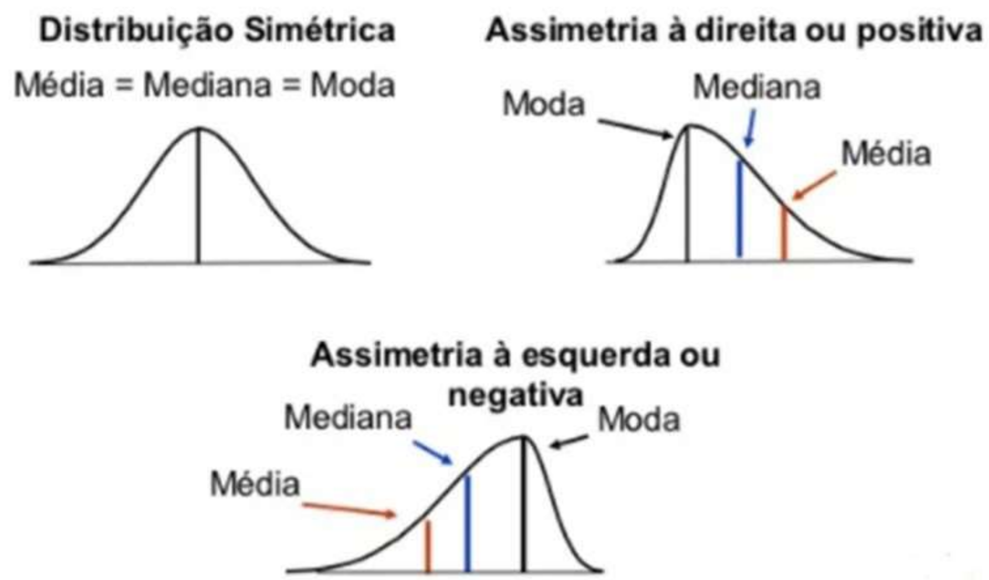


Figura 1. Histograma, curva normal e box-plot de CTC (A) ($cmol_c\ dm^{-3}$), V% (B) (%), fósforo (C) ($mg\ dm^{-3}$) e potássio (D) ($cmol_c\ dm^{-3}$)

COEFICIENTE DE ASSIMETRIA (As)



COEFICIENTE DE ASSIMETRIA (As)

- Simétrica: |As| < 0,15
- Assimétrica moderada: 0,15 ≤ |As| < 1,0
- Assimétrica forte: |As| ≥ 1,0

$$As = \frac{3(média - mediana)}{desvio - padrão}$$

COEFICIENTE DE ASSIMETRIA (As)

- Exemplo: Foram coletados os pesos (kg) de alunos de três turmas da educação infantil de uma escola.

	Turma A	Turma B	Turma C
Média	12	12,9	11,1
Mediana	12	13,5	10,5
Desvio-padrão	4,42	4,20	4,20
n	*60*	*78*	*78*

$$As_A = \frac{3\,(12-12)}{4{,}42} = 0 \qquad As_B = \frac{3\,(12{,}9-13{,}5)}{4{,}20} = -0{,}429$$

$$As_C = \frac{3\,(11{,}1-10{,}5)}{4{,}20} = +0{,}429$$