

ESTATÍSTICA DESCRITIVA

Prof. Almeida Júnior

professor.almeidajunior@gmail.com

 @profalmeidajunior

INSTRUTOR

- Prof Almeida Júnior
- Engenheiro Eletrônico – UFBA
- Pós Graduações – Criptografia (UFF); Automação Industrial (UERJ)
- Mestrado – Sistemas Computacionais (UNIFACS)
- Certificações: PMP, ITIL



MÉDIA ARITMÉTICA SIMPLES

$$\overline{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

[Questão]

(VUNESP 2018/CM INDAIATUBA) A tabela a seguir mostra o número de funcionários que faltaram ao trabalho nos cinco dias de uma semana em determinada empresa.

A média diária do número de funcionários que faltaram ao trabalho na 2ª e 6ª feiras, desta semana, supera a média diária do número de funcionários que faltaram na 3ª, 4ª e 5ª feiras, da mesma semana, em

DIAS DA SEMANA	Número de Funcionários Faltantes
2ª feira	13
3ª feira	9
4ª feira	6
5ª feira	18
6ª feira	23

(A) 7. (B) 8. (C) 9. (D) 10. (E) 11.

[Questão]

(VUNESP 2018/CM INDAIATUBA) A tabela a seguir mostra o número de funcionários que faltaram ao trabalho nos cinco dias de uma semana em determinada empresa.

A média diária do número de funcionários que faltaram ao trabalho na 2ª e 6ª feiras, desta semana, supera a média diária do número de funcionários que faltaram na 3ª, 4ª e 5ª feiras, da mesma semana, em

DIAS DA SEMANA	Número de Funcionários Faltantes
2ª feira	13
3ª feira	9
4ª feira	6
5ª feira	18
6ª feira	23

(A) 7. (B) 8. (C) 9. (D) 10. (E) 11.

[Questão]

(CESPE - Perito Criminal Federal/Área 6/2018) Considerando que a análise de uma amostra de minério de chumbo tenha apresentado os seguintes resultados percentuais (%): 8,10; 8,32; 8,12; 8,22; 7,99; 8,31, julgue o item a seguir, relativo a esses dados.

O valor médio do teor de chumbo presente na amostra foi superior a 8%.

[Questão]

(CESPE - Perito Criminal Federal/Área 6/2018) Considerando que a análise de uma amostra de minério de chumbo tenha apresentado os seguintes resultados percentuais (%): 8,10; 8,32; 8,12; 8,22; 7,99; 8,31, julgue o item a seguir, relativo a esses dados.

O valor médio do teor de chumbo presente na amostra foi superior a 8%.

CERTO

MÉDIA PONDERADA

$$\overline{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots + x_n f_n}{f_1 + f_2 + f_3 + \dots + f_n}$$

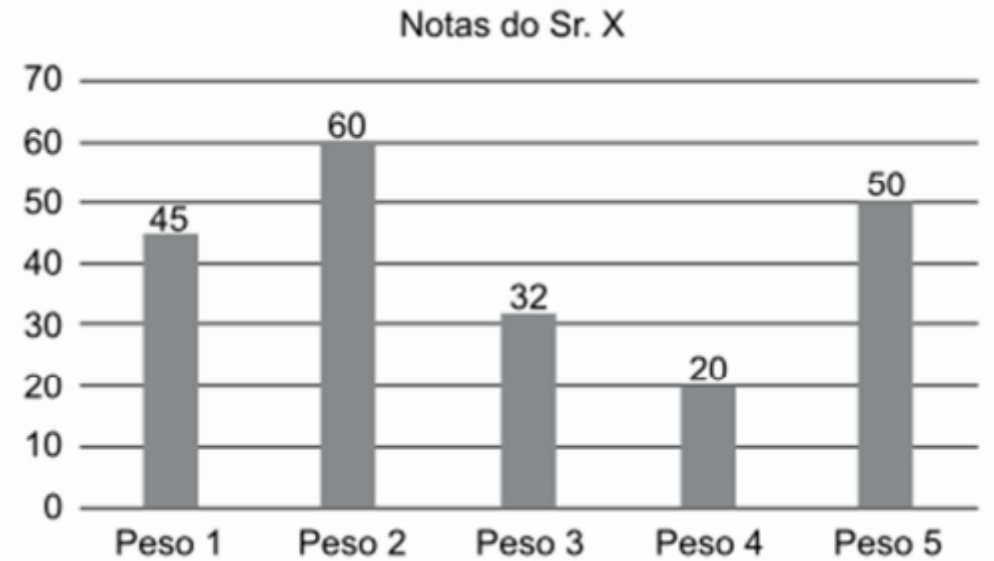
EXEMPLO

Nota	Peso
10	3
5	2
2	4
2	1

[Questão]

(VUNESP 2018/CM de Indaiatuba/Controlador Interno) O gráfico a seguir mostra as cinco notas do Sr. X e os respectivos pesos atribuídos a cada uma das provas.

Para ser aprovado, o Sr. X precisava que sua média aritmética ponderada por esses pesos fosse maior ou igual a 48 pontos. Com essas notas, o Sr. X não foi aprovado e sua média ficou abaixo de 48, em uma quantidade de pontos igual a

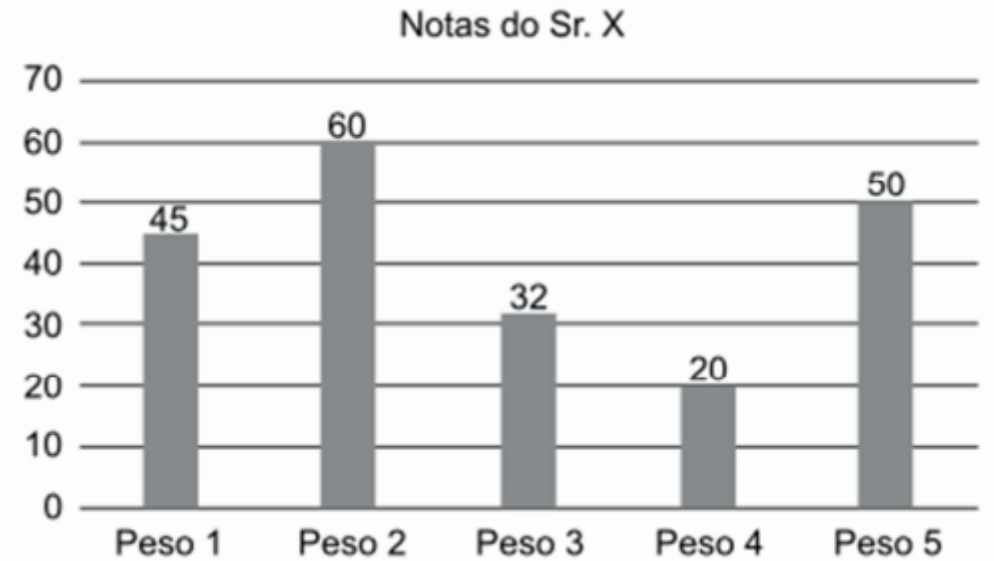


(A) 6,6. (B) 7,9. (C) 8,6. (D) 9,5. (E) 10,8.

[Questão]

(VUNESP 2018/CM de Indaiatuba/Controlador Interno) O gráfico a seguir mostra as cinco notas do Sr. X e os respectivos pesos atribuídos a cada uma das provas.

Para ser aprovado, o Sr. X precisava que sua média aritmética ponderada por esses pesos fosse maior ou igual a 48 pontos. Com essas notas, o Sr. X não foi aprovado e sua média ficou abaixo de 48, em uma quantidade de pontos igual a



(A) 6,6. (B) 7,9. **(C) 8,6.** (D) 9,5. (E) 10,8.

MÉDIA PARA DADOS AGRUPADOS EM CLASSES

Altura	Frequência
1,50 – 1,54	10
1,54 – 1,58	3
1,58 – 1,62	2
1,62 – 1,66	1
Total	

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \dots + x_n f_n}{f_1 + f_2 + f_3 + \dots + f_n}$$

MÉDIA PARA DADOS AGRUPADOS EM CLASSES

Altura	Frequência
1,50 – 1,54	10
1,54 – 1,58	10
1,58 – 1,62	1
1,62 – 1,66	1
Total	

[Questão]

(VUNESP 2014/TJ-PA/Analista Judiciário) A Universidade Alfa quer ter informações sobre a situação sócio-financeira dos alunos que vêm de fora da cidade onde está instalada e o aluguel residencial é uma das variáveis estudadas. Para informações sobre isso, colheu-se uma amostra de 32 alunos, cujos resultados são mostrados na tabela de distribuição de frequência que segue:

O valor médio dos aluguéis está:

- (A) abaixo de 1 000 reais.
- (B) entre 1 000 e 1 200 reais.
- (C) entre 1 200 e 1 500 reais.
- (D) entre 1 500 e 1 600 reais.
- (E) acima de 1 600 reais.

Aluguel mensal (R\$)	Frequência
De 700 a menos de 900	3
De 900 a menos de 1 100	10
De 1 100 a menos de 1 300	10
De 1 300 a menos de 1 500	8
De 1 500 a menos de 1 700	1
Total	32

[Questão]

(VUNESP 2014/TJ-PA/Analista Judiciário) A Universidade Alfa quer ter informações sobre a situação sócio-financeira dos alunos que vêm de fora da cidade onde está instalada e o aluguel residencial é uma das variáveis estudadas. Para informações sobre isso, colheu-se uma amostra de 32 alunos, cujos resultados são mostrados na tabela de distribuição de frequência que segue:

O valor médio dos aluguéis está:

- (A) abaixo de 1 000 reais.
- (B) entre 1 000 e 1 200 reais.**
- (C) entre 1 200 e 1 500 reais.
- (D) entre 1 500 e 1 600 reais.
- (E) acima de 1 600 reais.

Aluguel mensal (R\$)	Frequência
De 700 a menos de 900	3
De 900 a menos de 1 100	10
De 1 100 a menos de 1 300	10
De 1 300 a menos de 1 500	8
De 1 500 a menos de 1 700	1
Total	32

MÉDIA GEOMÉTRICA

$$G = \sqrt[n]{x_1 * x_2 * \dots * x_n}$$

MÉDIA HARMÔNICA

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}}$$

DESIGUALDADE DAS MÉDIAS

$$\frac{1}{n} \sum_{i=1}^n x_i \geq \sqrt[n]{\prod_{i=1}^n x_i} \geq \frac{n}{\sum_{i=1}^n \left(\frac{1}{x_i} \right)}$$

média aritmética é maior ou igual à média geométrica e esta maior ou igual à média harmônica

[Questão]

(CESPE/ANAC/2012) Considerando que uma pesquisa de satisfação referente a um novo terminal de passageiros tenha sido realizada com 50 pessoas e o resultado em uma amostra de notas conforme apresentado na tabela acima, julgue os itens seguintes.

A média geométrica da distribuição das notas foi superior à média aritmética.

nota	contagem
0	27
1	6
2	5
3	7
4	2
5	3

[Questão]

(CESPE/ANAC/2012) Considerando que uma pesquisa de satisfação referente a um novo terminal de passageiros tenha sido realizada com 50 pessoas e o resultado em uma amostra de notas conforme apresentado na tabela acima, julgue os itens seguintes.

A média geométrica da distribuição das notas foi superior à média aritmética.

nota	contagem
0	27
1	6
2	5
3	7
4	2
5	3

ERRADO

POPULAÇÃO E AMOSTRA

- Uma população é um conjunto de pessoas, itens ou eventos sobre os quais você quer fazer inferências (censo)
- Uma amostra é um conjunto reduzido da população que mantém todas as características essenciais da mesma (amostragem)

VARIÁVEL

- Característica dos elementos da amostra que nos interessa averiguar estatisticamente
- Classificação
 - Qualitativas
 - Ordinais (Ex: Grau de Instrução)
 - Nominais (Ex: Cores)
 - Quantitativas
 - Discretas (Ex: Número de filhos de um casal)
 - Contínuas (Ex: Altura, Peso)

DADOS BRUTOS, ROL E AMPLITUDE TOTAL

- Dados Brutos: Conjunto de dados que ainda não foram numericamente organizados, obtidos após a crítica dos valores.
- Rol: É um arranjo dos dados brutos em ordem crescente.
- Amplitude Total (AT): É a diferença entre o maior e o menor valor observado.

DADOS TABULADOS

- Frequência absoluta ou simples (f_i): é o número de observações
- Dados tabulados
 - Não há perda de informação

DISTRIBUIÇÃO DE FREQUÊNCIAS

Classe	Frequência (fi)

- Frequência absoluta ou simples (fi): é o número de observações que se encontra presente em uma classe ou intervalo específico
- Há perda de informação

ELEMENTOS DE UMA DISTRIBUIÇÃO DE FRÊQUÊNCIAS

- Classe
- Limites de Classe
- Intervalo de Classe
 - |-----
 - -----|
 - |-----|
 - -----
- Amplitude do intervalo
- Amplitude da Distribuição

Classes	Freqüências
41 ----- 45	7
45 ----- 49	3
49 ----- 53	4
53 ----- 57	1
57 ----- 61	5
Total	20

TIPOS DE FREQUÊNCIA

- Absoluta (f) e Relativa (F)
 - Simples
 - Acumulada
 - Crescente e Decrescente

Classe	f	fac	fad	F	Fac	Fad
20 --- 30	2					
30 --- 40	3					
40 --- 50	4					
50 --- 60	1					

TIPOS DE FREQUÊNCIA

- Absoluta (f) e Relativa (F)
 - Simples
 - Acumulada
 - Crescente e Decrescente

Classe	f	fac	fad	F (%)	Fac (%)	Fad (%)
20 --- 30	2	2	10	20%	20%	100%
30 --- 40	3	5	8	30%	50%	80%
40 --- 50	4	9	5	40%	90%	50%
50 --- 60	1	10	1	10%	100%	10%

INTERPOLAÇÃO LINEAR DA OGIVA

Classe	Frequência (fi)
0 ----- 10	5
10 ----- 20	8
20 ----- 30	11
30 ----- 40	4

Quantos elementos estão abaixo do valor 12 ?

R: $5 + X$

$X = 1,6$

R: 6,6

INTERPOLAÇÃO LINEAR DA OGIVA

Classe	Frequência (fi)
0 ----- 10	5
10 ----- 20	8
20 ----- 30	11
30 ----- 40	4

Quantos elementos estão acima do valor 15 ?

R: $X + 11 + 4$

$X = 4$

R: 19

MODA - ROL

- É o valor com maior frequência
 - Cuidado: não é a frequência desse valor!
- Podemos ter mais de uma moda
- A moda pode não existir

MODA — DADOS TABULADOS

x_i	f_i
1	5
7	8
9	11
10	4

MODA — DISTRIBUIÇÃO DE FREQUÊNCIAS

Classe	Frequência (fi)
0 ----- 10	5
10 ----- 20	8
20 ----- 30	11
30 ----- 40	4

- Moda Bruta (Ponto Médio)
- Moda de Czuber (Regra)
- King

$$Mo = \frac{Linf + Lsup}{2}$$

MODA — CZUBER

Classe	Frequência (fi)
0 ----- 10	5
10 ----- 20	8
20 ----- 30	11
30 ----- 40	4

$$Mo = Linf + \frac{\Delta a}{\Delta a + \Delta p} h$$

$$Mo = Linf + \frac{\Delta a}{\Delta a + \Delta p} h = 20 + \frac{3}{3 + 7} 10 = 23$$

MODA – KING

Classe	Frequência (fi)
0 ----- 10	5
10 ----- 20	8
20 ----- 30	11
30 ----- 40	4

$$Mo = Linf + \frac{fp}{fa + fp}h$$

$$Mo = Linf + \frac{fp}{fa + fp}h = 20 + \frac{4}{8 + 4}10 = 23,33..$$

ATENÇÃO PARA DIFERENÇA ENTRE AS FÓRMULAS

$$Mo = Linf + \frac{\Delta a}{\Delta a + \Delta p} h$$

$$Mo = Linf + \frac{fp}{fa + fp} h$$

MEDIANA - ROL

É aquele valor que está no meio

- 1) Colocar os elementos em ordem
- 2) Se quantidade for impar, termo central
- 3) Se quantidade for par, calcular a média

[Questão]

(FCC 2018/ALE-SE) Em um grupo de pessoas encontramos as seguintes idades: 20, 30, 50, 39, 20, 25, 41, 47, 36, 45, 41, 52, 18, 41.

A mediana é

- a) 36.
- b) 40.
- c) 41.
- d) 42.
- e) 39.

[Questão]

(FCC 2018/ALE-SE) Em um grupo de pessoas encontramos as seguintes idades: 20, 30, 50, 39, 20, 25, 41, 47, 36, 45, 41, 52, 18, 41.

A mediana é

- a) 36.
- b) 40.**
- c) 41.
- d) 42.
- e) 39.

MEDIANA — DISTRIBUIÇÃO DE FREQUÊNCIAS

Classe	f	fac
0 ----- 10	5	5
10 ----- 20	8	13
20 ----- 30	10	23
30 ----- 40	5	

- Calcular n e n/2
- fac (até ser $\geq n/2$)
- Aplicar fórmula

$$Md = Linf + \frac{\frac{n}{2} - fac_A}{f} h$$

$$Md = 20 + \frac{\frac{28}{2} - 13}{10} 10 = 20 + 14 - 13 = 21$$

QUARTIL

- Os quartis dividem os dados em 4 partes
- São 3 quartis: Q_1 , Q_2 , Q_3
- Q_2 é a mediana
- $Q_3 - Q_1$ é chamado de distância/amplitude interquartílica
- Semi-interquartílica é metade amplitude interquartílica
 - Também chamado de desvio quartílico



[Questão]

(CESPE 2018/IPHAN) Uma pesquisa a respeito das quantidades de teatros em cada uma de 11 cidades brasileiras selecionadas apresentou o seguinte resultado: $\{1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4\}$. Com referência a esses dados, julgue os itens seguintes.

- 01. A mediana do conjunto é igual a 3.
- 02. O valor do primeiro quartil do conjunto de dados (Q_1) é igual a 3
- 03. O valor do terceiro quartil do conjunto de dados (Q_3) é igual a 4

[Questão]

(CESPE 2018/IPHAN) Uma pesquisa a respeito das quantidades de teatros em cada uma de 11 cidades brasileiras selecionadas apresentou o seguinte resultado: $\{1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4\}$. Com referência a esses dados, julgue os itens seguintes.

- 01. A mediana do conjunto é igual a 3.
- 02. O valor do primeiro quartil do conjunto de dados (Q_1) é igual a 3
- 03. O valor do terceiro quartil do conjunto de dados (Q_3) é igual a 4

CERTO, ERRADO, CERTO

QUARTIL – DISTRIBUIÇÃO DE FREQUÊNCIAS

Classe	f	fac
0 ----- 10	5	5
10 ----- 20	8	13
20 ----- 30	10	23
30 ----- 40	5	

- Calcular $n/4$
- fac (até ser $\geq n/4$)
- Aplicar fórmula

$$Q_1 = Linf + \frac{\frac{n}{4} - fac_A}{f} h$$

$$Md = 10 + \frac{\frac{28}{4} - 5}{8} 10 = 10 + (20/8) = 100/8 = 12,5$$

QUARTIL – DISTRIBUIÇÃO DE FREQUÊNCIAS

Classe	f	fac
0 ----- 10	5	5
10 ----- 20	8	13
20 ----- 30	10	23
30 ----- 40	5	

- Calcular $n/2$
- fac (até ser $\geq n/2$)
- Aplicar fórmula

$$Q_2 = Linf + \frac{\frac{n}{2} - fac_A}{f} h$$

$$Md = 20 + \frac{\frac{28}{2} - 13}{10} 10 = 20 + 14 - 13 = 21$$

QUARTIL – DISTRIBUIÇÃO DE FREQUÊNCIAS

Classe	f	fac
0 ----- 10	5	5
10 ----- 20	8	13
20 ----- 30	10	23
30 ----- 40	5	

- Calcular $3n/4$
- fac (até ser $\geq 3n/4$)
- Aplicar fórmula

$$Q_3 = Linf + \frac{\frac{3n}{4} - fac_A}{f} h$$

$$Md = 20 + \frac{\frac{3 \cdot 28}{4} - 13}{10} 10 = 20 + 21 - 13 = 28$$

QUARTIL – FÓRMULA GERAL

$$Q_i = Linf + \frac{\frac{i \cdot n}{4} - fac_A}{f} h$$

$$Q_1 = Linf + \frac{\frac{n}{4} - fac_A}{f} h$$

$$Q_2 = Linf + \frac{\frac{2n}{4} - fac_A}{f} h$$

$$Q_3 = Linf + \frac{\frac{3n}{4} - fac_A}{f} h$$

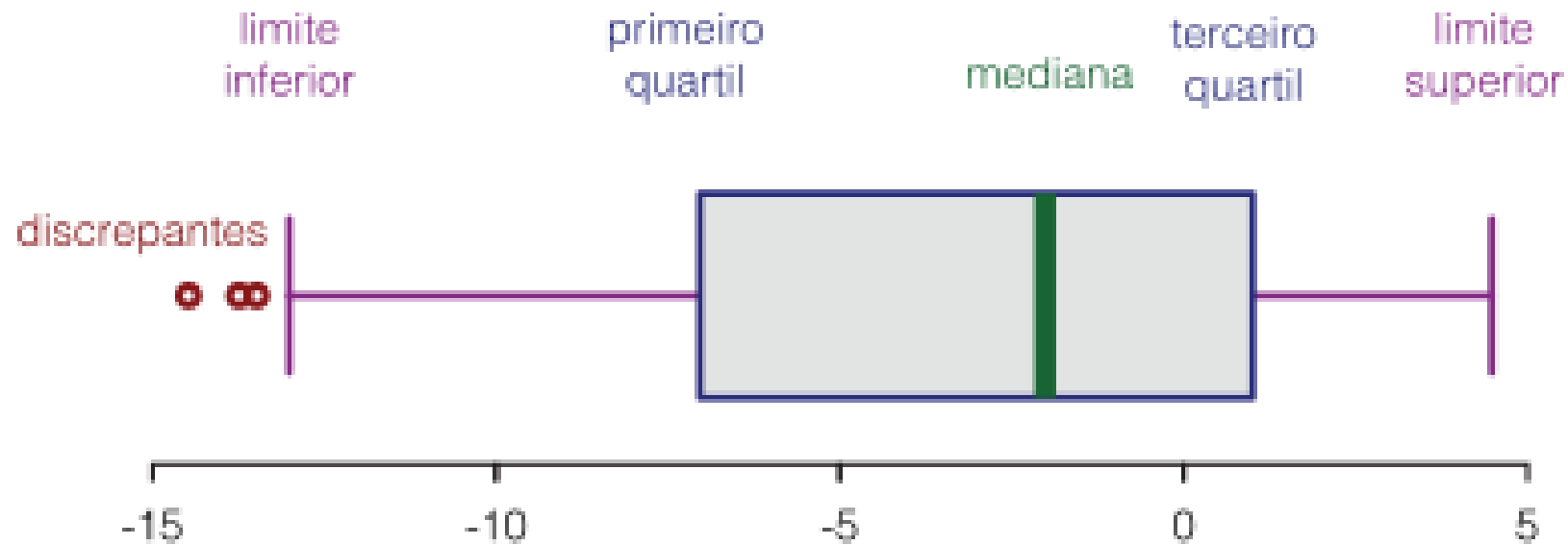
DECIS (D1...D9)

$$D_i = Linf + \frac{\frac{i \cdot n}{10} - fac_A}{f} h$$

PERCENTIS

$$D_i = Linf + \frac{\frac{i \cdot n}{100} - fac_A}{f} h$$

BOX PLOT / DIGRAMA DE CAIXA / BOX-AND-WHISKER PLOT



$$Lsup = \min(\max(dados); Q3 + 1,5DI)$$

$$Linf = \max(\min(dados); Q1 - 1,5DI)$$

MEDIDAS DE DISPERSÃO/VARIABILIDADE

- Amplitude
- Amplitude interquartílica
- Desvio quartílico
- Desvio médio (ou desvio absoluto médio),
- Variância
- Desvio padrão
- Coeficiente de variação
- Variância relativa

[Questão]

(VUNESP/2016/MPE-SP) Na estatística, são considerados medidas de dispersão:

- a) média e moda.
- b) percentil e coeficiente de variação.
- c) amplitude total e percentil.
- d) amplitude total e desvio padrão.
- e) variância e média.

[Questão]

(VUNESP/2016/MPE-SP) Na estatística, são considerados medidas de dispersão:

- a) média e moda.
- b) percentil e coeficiente de variação.
- c) amplitude total e percentil.
- d) amplitude total e desvio padrão.**
- e) variância e média.



DÚVIDAS

Siga-me no Instagram:

@profalmeidajunior