By @kakashi_copiador

Índice

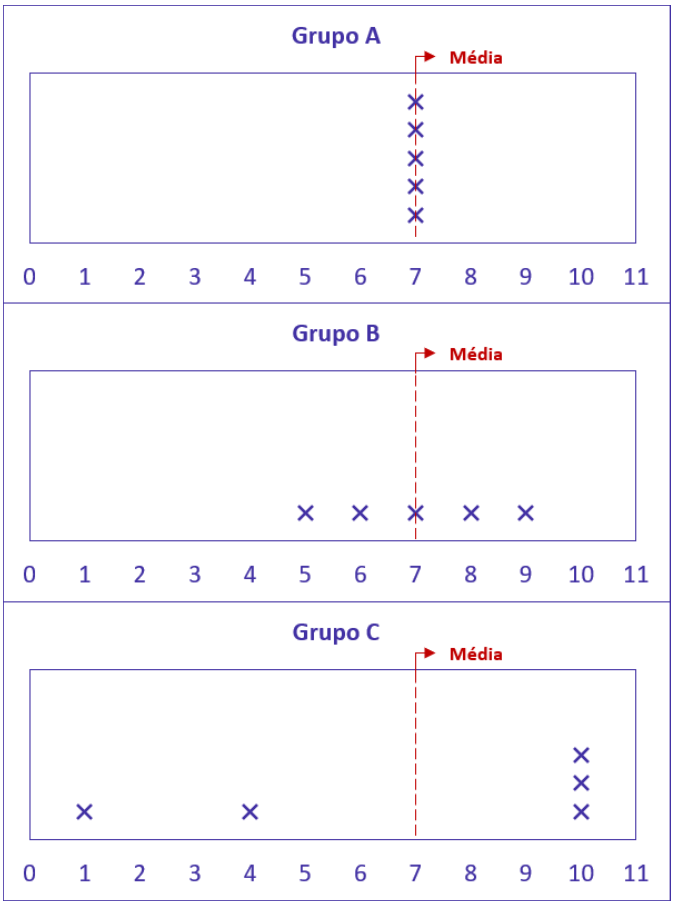
MEDIDAS DE VARIABILIDADE

Nas aulas anteriores, estudamos mecanismos para encontrar valores (média, mediana e moda) que sintetizam o comportamento dos elementos de um conjunto de dados. Esses valores fornecem parâmetros significativos para uma análise dos dados, porém, também é importante identificarmos como variam ou como se diferenciam as características dos elementos de um conjunto.

Imagine, por exemplo, que você precise avaliar três grupos de alunos, cada um com cinco elementos, no que diz respeito ao domínio de uma determinada matéria. Os testes mostraram os seguintes resultados:

Grupos
$\boldsymbol{A = 7, 7, 7, 7, 7}$
$\boldsymbol{B = 5, 6, 7, 8, 9}$
$\boldsymbol{C = 1, 4, 10, 10, 10}$

Para analisar esses dados, podemos, inicialmente, calcular a média aritmética dos três grupos. Concluímos, então, que todos possuem a mesma média aritmética ($\bar{x} = 7$). Contudo, ao observarmos a variação dos dados, percebemos que os grupos se comportam de maneira diferente, apesar de todos possuírem a mesma média.



Nesse caso, **a média**, embora seja uma medida representativa do conjunto, **não indica o grau de homogeneidade ou heterogeneidade existente entre os valores que compõem o conjunto**. Desse modo, precisamos recorrer a procedimentos matemáticos que possibilitem a compreensão da discrepância existente entre os valores do conjunto.

As medidas de dispersão (ou variabilidade) são justamente métricas que mostram a variação dos dados de um conjunto. Elas podem ser divididas em **dois grupos**:

a) medidas de dispersão absoluta:
- amplitude total;
- amplitude interquartílica;
- desvio médio;
- variância; e
- desvio-padrão.

b) medidas de dispersão relativa:
- coeficiente de variação (de Pearson); e
- variância relativa.

Nessa aula, aprenderemos a medir o grau de concentração ou dispersão dos dados em torno da média. Para isso, estudaremos as principais medidas de dispersão, que são: amplitude total, amplitude interquartílica, desvio médio, variância, desvio padrão, coeficiente de variação e variância relativa.

HORA DE PRATICAR!

(COC-UFAC/UFAC/2019) Analise as seguintes assertivas:

I. A moda e o desvio padrão são medidas de dispersão,

II. O desvio médio e a média são medidas de dispersão,

III. O coeficiente de variação e a variância são medidas de dispersão,

IV. A moda, a média e o desvio padrão são medidas de posição.

Pode-se afirmar que estão corretas:

a) Apenas I. e II.

b) Apenas II. e III.

c) Apenas III.

d) Apenas IV.

e) Apenas I. e IV.

Comentários:

As medidas de posição consistem em valores que representam a tendência de concentração dos dados observados. As medidas de posição mais importantes são as medidas de tendência central. Nesse grupo, encontram-se as medidas mais utilizadas: média aritmética, moda e mediana.

Já as medidas de dispersão medem o grau de variabilidade dos elementos de uma distribuição. A dispersão aumenta à proporção que o valor da medida de dispersão também aumenta. As principais medidas de dispersão são amplitude, desvio médio, variância, desvio padrão e coeficiente de variação.

Gabarito: C.

(VUNESP/MPE-SP/2016) 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.

Comentários:

As medidas de tendência central estudam o centro da amostra. As medidas de tendência central mais utilizadas são a média aritmética, a mediana e a moda.

Por sua vez, as medidas de separatrizes dividem os dados em grupos com a mesma quantidade de elementos, sendo representadas pelos quartis, decis e percentis.

Por fim, as medidas de dispersão têm a finalidade de identificar o quanto os dados estão dispersos em torno da média de uma amostra. São dadas pelos coeficientes de variação, desvio padrão, amplitude e variância.

Gabarito: D.

AMPLITUDE TOTAL

A **amplitude total** (ou simplesmente amplitude) é a **diferença entre os valores extremos de um conjunto de observações**, ou seja, **a diferença entre o maior e o menor elemento desse conjunto**:

$$\boldsymbol{A_T = x_{máx} - x_{mín}}$$

Essa medida de dispersão chama atenção por ser extremamente simples e muito fácil de se calcular. Contudo, há uma certa restrição quanto ao seu uso por conta de sua grande instabilidade, vez que leva em consideração apenas os valores extremos da série.

Por exemplo, vamos comparar os conjuntos A e B da tabela a seguir:

Conjunto	Média	Amplitude total
$\boldsymbol{A = 5, 7, 8, 9, 10, 11, 55}$	$\bar{x} = 15$	$A_T = 55 - 5 = 50$
$\boldsymbol{B = 12, 13, 14, 15, 16, 17, 18}$	$\bar{x} = 15$	$A_T = 18 - 12 = 6$

Reparem que as médias aritméticas dos dois conjuntos são iguais a 15. Portanto, no que diz respeito a essa medida de posição, podemos considerá-los idênticos. Porém, ao calcularmos a amplitude total, verificamos que os valores do conjunto A apresentam um grau de dispersão bem maior que os do conjunto B.

Isso acontece porque, **no cálculo da amplitude total, desconsideramos os valores da série que se encontram entre os extremos, o que pode conduzir a interpretações equivocadas**. Com frequência, um valor discrepante pode afetar a medida de maneira acentuada. É o caso, por exemplo, do último valor (55) do conjunto A, sensivelmente maior que seu antecessor (11), que elevou a magnitude da amplitude total para 50.

Além disso, **a amplitude total também é sensível ao tamanho de amostra**. Normalmente, a amplitude total tende a aumentar com o incremento do tamanho da amostra, ainda que não proporcionalmente. Ainda, **a amplitude total pode apresentar muita variação de uma amostra para outra, ainda que extraídas de uma mesma população.**

Apesar das limitações dessa medida, há situações em que ela pode ser aplicada de forma satisfatória. É o caso, por exemplo, da variação da temperatura em um dia. Também é o caso de quando uma compreensão rápida dos dados é mais relevante que a exatidão de um procedimento complexo.

Amplitude Total para dados não-agrupados

Para dados não agrupados, o cálculo da amplitude total pode ser expresso pela seguinte fórmula:

$$\boldsymbol{A_T = x_{máx} - x_{mín}}$$

em que $x_{máx}$ é o maior elemento; e $x_{mín}$ é o menor elemento do conjunto.

EXEMPLIFICANDO

Calcular a amplitude total dos conjuntos apresentados a seguir:

$$A = 50, 50, 50, 50, 50, 50, 50$$

$$B = 47, 48, 49, 50, 51, 52, 53$$

$$C = 20, 30, 40, 50, 60, 70, 80$$

Aplicando a fórmula anterior para esses dados, obtemos os seguintes resultados:

$$A_{T_A} = x_{máx} - x_{mín} = 50 - 50 = 0$$

$$A_{T_B} = x_{máx} - x_{mín} = 53 - 47 = 6$$

$$A_{T_C} = x_{máx} - x_{mín} = 80 - 20 = 60$$

Nesse caso, podemos observar que o conjunto A obteve uma amplitude total igual a 0, ou seja, uma dispersão nula. Então, significa que os valores não variam entre si. O conjunto B, por sua vez, obteve uma amplitude igual a 6. Já a variável C teve uma amplitude total igual a 60.

Embora o valor da amplitude total seja diferente para os conjuntos A, B e C, todos possuem a mesma média aritmética (50). Independentemente da média, verificamos que o conjunto A possui elementos mais homogêneos do que os conjuntos B e C. E, também, que os elementos do conjunto B são mais homogêneos do que os do conjunto C.

Amplitude Total para dados agrupados sem intervalos de classes

Para dados **agrupados SEM intervalos de classe**, a fórmula usada para a identificação da **amplitude total** é similar à adotada para dados não-agrupados. **A única diferença consiste na identificação dos valores mínimo e máximo, que agora ocorre por meio de uma tabela de frequências.**

EXEMPLIFICANDO

Calcular a amplitude total da tabela de frequências apresentada a seguir.

x_i	f_i
1	10
3	15
5	10
7	8
9	7

Nesse caso, como 1 e 9 são os valores mínimo e máximo da variável x_i, temos o seguinte resultado:

$$A_T = x_{máx} - x_{mín}$$

$$A_T = 9 - 1 = 8$$

É importante ressaltar que esses valores foram selecionados independentemente da frequência associada a eles.

Amplitude Total para dados agrupados em classes

Para dados **agrupados em intervalos de classe**, podemos definir a **amplitude total** de duas formas:

1) pela diferença entre o limite superior da última classe ($\boldsymbol{L_{sup}}$) e o limite inferior da primeira classe ($\boldsymbol{l_{inf}}$), conforme expresso na fórmula a seguir:

$$\boldsymbol{A = L_{sup} - l_{inf}}$$

2) pela diferença entre o ponto médio da última classe ($\boldsymbol{PM_{\text{últ}}}$) e o ponto médio da primeira classe ($\boldsymbol{PM_{pri}}$), conforme expresso na fórmula a seguir:

$$\boldsymbol{A = PM_{\text{últ}} - PM_{pri}}$$

EXEMPLIFICANDO

Calcular a amplitude total da distribuição de frequências apresentada a seguir:

Classes	PM_i	f_i
1 ⊢ 5	3	5
5 ⊢ 9	7	10
9 ⊢ 13	11	15
13 ⊢ 17	15	10
17 ⊢ 21	19	5
Total		**45**

Pelo primeiro método, temos que o limite superior da última classe é 21, enquanto o limite inferior da primeira classe é 1. Portanto, temos a seguinte amplitude:

$$A = L_{sup} - l_{inf}$$

$$A = 21 - 1 = 20$$

Pelo segundo método, temos que o ponto médio da última classe é 19, enquanto o ponto médio da primeira classe é 3. Portanto, temos a seguinte amplitude:

$$A = PM_{últ} - PM_{pri}$$

$$A = 19 - 3 = 16$$

Observe que a amplitude é menor pelo segundo método, porque os extremos da distribuição são desconsiderados.

HORA DE PRATICAR!

(COPEVE (UFAL)/Pref. Maceió/2012) Um registro em saúde epidemiológica apresenta os dados: 3, 4, 7, 8 e 8. Se calcularmos 8 – 3 = 5, estaremos determinando:

a) a amplitude total.

b) o primeiro quartil.

c) o desvio médio.

d) a distância interquartílica.

e) o terceiro quartil.

Comentários:

A amplitude total (ou simplesmente amplitude) é a diferença entre os valores extremos de um conjunto de observações, ou seja, a diferença entre o maior e o menor elemento desse conjunto.

Gabarito: A.

(VUNESP/Pref. de São José dos Campos/2012) A diferença entre o maior e o menor valor em um conjunto de dados é denominado (a)

a) curva normal.

b) amplitude total.

c) média.

d) média ponderada.

e) moda.

Comentários:

A diferença entre o maior e o menor valor em um conjunto de dados é denominada de amplitude (ou amplitude total).

Gabarito: B.

Propriedades da Amplitude Total

Nesse tópico, estudaremos as principais propriedades da amplitude total:

1ª Propriedade

- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, a amplitude do conjunto não é alterada.**

EXEMPLIFICANDO

Vamos tomar como exemplo a sequência $\{x_n\} = \{3, 6, 8, 9, 10\}$, cuja amplitude total é:

$$A = x_{máx} - x_{mín}$$
$$A = 10 - 3 = 7$$

Se adicionarmos o número 5 a cada um dos termos da sequência, obteremos uma nova lista $\{y_n\} = \{x_n + 5\} = \{8, 11, 13, 14, 15\}$, cuja amplitude total é:

$$A = y_{máx} - y_{mín}$$
$$A = 15 - 8 = 7$$

Logo, a adição do número 5 a cada um dos termos da sequência fez com que a amplitude permanecesse inalterada.

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, a amplitude do conjunto fica multiplicada (ou dividida) por essa constante.**

EXEMPLIFICANDO

Vamos tomar como exemplo a sequência $\{x_n\} = \{8, 11, 13, 14, 15\}$, cuja amplitude total é:

$$A = x_{máx} - x_{mín}$$
$$A = 15 - 8 = 7$$

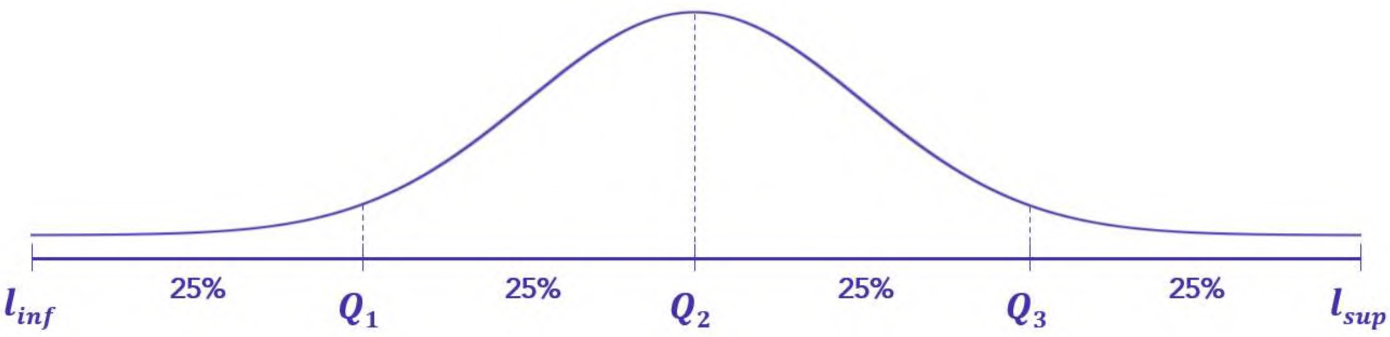
Se multiplicarmos cada um dos termos da sequência por 5, obteremos uma nova lista $\{y_n\} = \{x_n \times 5\} = \{40, 55, 65, 70, 75\}$, cuja amplitude total é:

$$A = y_{máx} - y_{mín}$$
$$A = 75 - 40 = 35$$

Logo, a multiplicação de cada um dos termos da sequência por 5 fez com que a amplitude total do conjunto também fosse multiplicada por 5.

AMPLITUDE INTERQUARTÍLICA

Como já sabemos, denominamos de **quartis** os valores de uma série que a dividem em **quatro partes iguais**, isto é, **quatro partes contendo o mesmo número de elementos (25%).** A imagem a seguir mostra os quartis de uma distribuição hipotética:



Temos, então, **3 quartis** (Q_1 , Q_2 e Q_3) para dividir uma série em **quatro partes** iguais:

- Q_1: o **primeiro quartil** corresponde à separação dos primeiros 25% de elementos da série;
- Q_2: o **segundo quartil** corresponde à separação de metade dos elementos da série, **coincidindo com a mediana ($Q_2 = M_d$)**;
- Q_3: o **terceiro quartil** corresponde à separação dos primeiros 75% de elementos da série, ou dos últimos 25% de elementos da série.

A **amplitude interquartílica** (ou distância interquartílica, ou intervalo interquartílico) é o resultado da subtração entre o terceiro quartil e o primeiro quartil:

$$\boxed{A_{IQ} = Q_3 - Q_1}$$

A **amplitude semi-interquartílica** (ou desvio quartílico) é definida como a metade desse valor, sendo calculada pela expressão apresentada a seguir:

$$\boxed{D_Q = \frac{Q_3 - Q_1}{2}}$$

FIQUE ATENTO!

Reparem que a fórmula da **amplitude interquartílica** (ou distância interquartílica) é muito parecida com a fórmula da **amplitude semi-interquartílico (ou desvio quartílico)**, podendo ser facilmente confundida.

HORA DE PRATICAR!

(AOCP/SUSIPE-PA/2018) Quartis são valores que dividem os dados de uma amostra em quatro grupos, cada um deles contendo 1/4 do tamanho total da amostra. Em relação ao assunto, informe se é verdadeiro (V) ou falso (F) o que se afirma a seguir e assinale a alternativa com a sequência correta.

( ) O primeiro quartil Q1 tem 1/4 dos dados acima dele e 3/4 dos dados abaixo dele.

( ) O terceiro quartil Q3 tem 3/4 dos dados abaixo dele e 1/4 dos dados acima dele.

( ) O quartil Q3 é a própria mediana.

( ) A distância interquartílica é dada por DIQ = Q3 – Q1.

a) V – F – F – V.

b) F – V – F – V.

c) F – V – V – V.

d) V – V – F – V.

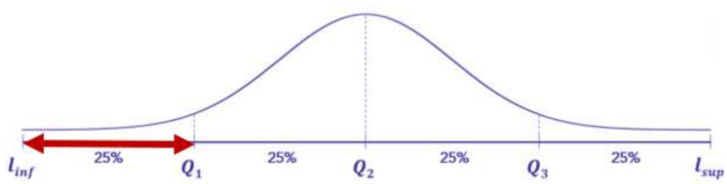
e) F – V – F – F.

Comentários:

Vamos analisar cada assertiva:

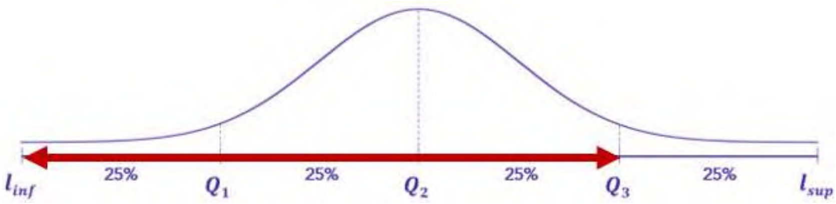
Item 1 - O primeiro quartil Q1 tem 1/4 dos dados acima dele e 3/4 dos dados abaixo dele.

Falso. O primeiro quartil (Q1) tem 1/4 (25%) dos dados **abaixo** dele e 3/4 (75%) **acima** dele.



Item 2 - O terceiro quartil Q3 tem 3/4 dos dados abaixo dele e 1/4 dos dados acima dele.

Verdadeiro. De fato, o terceiro quartil (Q3) tem 3/4 (75%) dos dados abaixo dele e 1/4 (25%) acima dele



Item 3 - O quartil Q3 é a própria mediana.

Falso. A terceira assertiva é falsa, pois a mediana é equivalente ao segundo quartil (Q2).

4 - A distância interquartílica é dada por $DIQ = Q_3 - Q_1$.

Verdadeiro. Essa é a exata definição de distância interquartílica.

Gabarito: B.

Propriedades da Amplitude Interquartílica

A seguir, veremos que a amplitude interquartílica e o desvio quartílico possuem as mesmas propriedades da amplitude total.

1ª Propriedade

- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, a amplitude interquartílica (e o desvio quartílico) do conjunto não é alterada.**

EXEMPLIFICANDO

Vamos tomar como exemplo a sequência $\{x_n\} = \{1, 3, 5, 7, 9, 11, 13\}$, cuja amplitude interquartílica é:

$$A_{IQ} = Q_3 - Q_1$$

$$A_{IQ} = 11 - 3 = 8$$

Se adicionarmos o número 5 a cada um dos termos da sequência, iremos obter uma nova lista $\{y_n\} = \{x_n + 5\} = \{6, 8, 10, 12, 14, 16, 18\}$, cuja amplitude interquartílica é:

$$A_{IQ} = Q_3 - Q_1$$

$$A_{IQ} = 16 - 8 = 8$$

Logo, a adição do número 5 a cada um dos termos da sequência fez com que a amplitude interquartílica permanecesse inalterada.

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, a amplitude interquartílica (e o desvio quartílico) do conjunto fica multiplicada (ou dividida) por essa constante.**

EXEMPLIFICANDO

Vamos tomar como exemplo a sequência $\{x_n\} = \{1, 3, 5, 7, 9, 11, 13\}$, cuja amplitude interquartílica é:

$$A_{IQ} = Q_3 - Q_1$$

$$A_{IQ} = 11 - 3 = 8$$

Se multiplicarmos cada um dos termos da sequência por 5, iremos obter uma nova lista $\{y_n\} = \{x_n \times 5\} = \{5, 15, 25, 35, 45, 55, 65\}$, cuja amplitude interquartílica é:

$$A_{IQ} = Q_3 - Q_1$$

$$A_{IQ} = 55 - 15 = 40$$

Logo, a multiplicação de cada um dos termos da sequência por 5 fez com que a amplitude interquartílica do conjunto também fosse multiplicada por 5.

DESVIOS EM RELAÇÃO À MÉDIA ARITMÉTICA E MEDIANA

Antes de apresentarmos as fórmulas para o cálculo do desvio médio e da variância, precisamos compreender qual o conceito de desvio em estatística. **Um desvio é a distância entre qualquer observação do conjunto de dados e uma medida descritiva desse conjunto**:

$$desvio = observação - medida$$

Em especial, destacamos os desvios em relação à média aritmética e em relação à mediana:

$$\boldsymbol{d_i = x - \bar{x}} \quad (média)$$

ou

$$\boldsymbol{d_i = x - M_d} \quad (mediana)$$

É natural pensarmos que, **quando os desvios em relação a uma medida descritiva são pequenos, as observações estão concentradas em torno dessa medida e, portanto, a variabilidade dos dados é pequena**. Agora, **quando os desvios são maiores, significa que as observações estão dispersas e, portanto, a variabilidade dos dados é grande.**

HORA DE PRATICAR!

(VUNESP/TJ-SP/2015) Leia o texto a seguir para responder à questão. Uma pequena empresa que emprega apenas cinco funcionários paga os seguintes salários mensais (em mil reais):

0,9	1,2	1,4	1,5	2,0

Considerando-se a média dos salários, o valor do desvio do salário de quem ganha R$ 1.400,00 mensais é

a) –1.000.

b) –400.

c) 0.

d) 200.

e) 400.

Comentários:

Para responder a questão, primeiro teremos que calcular a média:

$$\bar{x} = \frac{0,9 + 1,2 + 1,4 + 1,5 + 2}{5}$$

$$\bar{x} = \frac{7}{5} = 1,4\ mil$$

Então, o desvio em relação ao salário de R$ 1.400 é:

$$d_i = x_i - \bar{x}$$

$$1400 - 1400 = 0$$

Gabarito: C.

(VUNESP/TJ-SP/2015) Leia o texto a seguir para responder à questão. Uma pequena empresa que emprega apenas cinco funcionários paga os seguintes salários mensais (em mil reais):

0,9	1,2	1,4	1,5	2,0

Somando-se os valores absolutos dos desvios individuais dos salários tomados em relação à média, encontra-se o valor de

a) 1.400,00.

b) 1.200,00.

c) 1.000,00.

d) 800,00.

e) 0.

Comentários:

Como vimos na questão anterior, a média dos salários é:

$$\bar{x} = \frac{0,9 + 1,2 + 1,4 + 1,5 + 2}{5}$$

$$\bar{x} = \frac{7}{5} \Longrightarrow \bar{x} = 1,4\ mil$$

Agora, calcularemos os desvios para cada valor apresentado:

Valor	Desvio	Desvio absoluto
$0,9$	$0,9 - 1,4 = -0,50$	$0,5$
$1,2$	$1,2 - 1,4 = -0,2$	$0,2$
$1,4$	$1,4 - 1,4 = 0$	0
$1,5$	$1,5 - 1,4 = 0,1$	$0,1$
2	$2 - 1,4 = 0,6$	$0,6$
Total		**1,4**

Portanto, a soma dos desvios absolutos é 1,4 mil.

Gabarito: A.

Propriedades dos Desvios em Relação à Média Aritmética e Mediana

Nesse tópico, revisaremos algumas propriedades importantes dos desvios sobre as quais discutimos quando estudamos sobre a média e a mediana.

1ª Propriedade

- **A soma algébrica dos desvios em relação à média é nula.**

EXEMPLIFICANDO

Vamos tomar como exemplo a sequência $\{x_n\} = \{1, 2, 3, 4, 5, 6, 7\}$, com média $\bar{x} = 4$. O desvio em relação à média é a diferença entre cada elemento da sequência e a média aritmética. Como a sequência possui 7 elementos, teremos o mesmo número de desvios para calcular. Logo, basta encontrarmos a diferença entre cada elemento e a média:

$$d_1 = x_1 - \bar{x} = 1 - 4 = -3$$

$$d_2 = x_2 - \bar{x} = 2 - 4 = -2$$

$$d_3 = x_3 - \bar{x} = 3 - 4 = -1$$

$$d_4 = x_4 - \bar{x} = 4 - 4 = 0$$

$$d_5 = x_5 - \bar{x} = 5 - 4 = 1$$

$$d_6 = x_6 - \bar{x} = 6 - 4 = 2$$

$$d_7 = x_7 - \bar{x} = 7 - 4 = 3$$

Agora, somaremos todos esses desvios:

$$\sum_{i=1}^{7} d_i = d_1 + d_2 + d_3 + d_4 + d_5 + d_6 + d_7$$

$$\sum_{i=1}^{7} d_i = (-3) + (-2) + (-1) + 0 + 1 + 2 + 3$$

$$\sum_{i=1}^{7} d_i = 0$$

Portanto, não importa qual a sequência de números, a soma dos desvios em relação à média é sempre igual a zero.

2ª Propriedade

- **A soma dos quadrados dos desvios da sequência de números $\{x_i\}$, em relação a um número a, é mínima se a for a média aritmética dos números.**

EXEMPLIFICANDO

Novamente, vamos tomar como exemplo a sequência $\{x_n\} = \{1, 2, 3, 4, 5, 6, 7\}$, com média $\bar{x} = 4$. Já calculamos os desvios desses números em relação à média:

$$d_1 = x_1 - \bar{x} = 1 - 4 = -3$$

$$d_2 = x_2 - \bar{x} = 2 - 4 = -2$$

$$d_3 = x_3 - \bar{x} = 3 - 4 = -1$$

$$d_4 = x_4 - \bar{x} = 4 - 4 = 0$$

$$d_5 = x_5 - \bar{x} = 5 - 4 = 1$$

$$d_6 = x_6 - \bar{x} = 6 - 4 = 2$$

$$d_7 = x_7 - \bar{x} = 7 - 4 = 3$$

Na propriedade anterior, vimos que a soma dos desvios é sempre igual a zero. Agora, calcularemos a soma dos quadrados desses desvios. Em outras palavras, vamos elevar cada um deles ao quadrado e somar todos os resultados:

$$\sum_{i=1}^{7} d_i = d_1^2 + d_2^2 + d_3^2 + d_4^2 + d_5^2 + d_6^2 + d_7^2$$

$$\sum_{i=1}^{7} d_i = (-3)^2 + (-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2 + 3^2$$

$$\sum_{i=1}^{7} d_i = 9 + 4 + 1 + 0 + 1 + 4 + 9$$

$$\sum_{i=1}^{7} d_i = 28$$

A propriedade nos garante que, para essa sequência numérica, o valor 28 é o menor valor possível. Isto é, se encontrarmos os desvios em relação a outro número (diferente da média) e, em seguida, calcularmos a soma dos quadrados dos desvios, o valor obtido será maior que 28. Vamos ver o que acontece ao calcularmos o desvio em relação ao número 6:

$$d_1 = x_1 - \bar{x} = 1 - 6 = -5$$

$$d_2 = x_2 - \bar{x} = 2 - 6 = -4$$

$$d_3 = x_3 - \bar{x} = 3 - 6 = -3$$

$$d_4 = x_4 - \bar{x} = 4 - 6 = -2$$

$$d_5 = x_5 - \bar{x} = 5 - 6 = -1$$

$$d_6 = x_6 - \bar{x} = 6 - 6 = 0$$

$$d_7 = x_7 - \bar{x} = 7 - 6 = 1$$

Agora, calcularemos a soma dos quadrados desses números:

$$\sum_{i=1}^{7} d_i = (-5)^2 + (-4)^2 + (-3)^2 + (-2)^2 + (-1)^2 + 0^2 + 1^2$$

$$\sum_{i=1}^{7} d_i = 25 + 16 + 9 + 4 + 1 + 0 + 1 = 56$$

Como esperávamos, o resultado foi maior do que 28.

3ª Propriedade

- **A soma dos desvios absolutos de uma sequência de números, em relação a um número a, é mínima quando a é a mediana dos números.**

EXEMPLIFICANDO

Vamos tomar como exemplo a série $\{x_n\} = \{1, 2, 3, 4, 6, 7, 8, 9\}$. Como o número de termos é par, a mediana será, por convenção, a média aritmética dos dois termos centrais:

$$M_d = \frac{4+6}{2} = 5.$$

O desvio em relação à mediana corresponde à diferença entre cada elemento da sequência e a mediana. Como são 8 números, temos a mesma quantidade de desvios para calcular. Logo, basta encontrarmos a diferença entre cada número e a mediana:

$$d_1 = x_1 - M_d = 1 - 5 = -4$$

$$d_2 = x_2 - M_d = 2 - 5 = -3$$

$$d_3 = x_3 - M_d = 3 - 5 = -2$$

$$d_4 = x_4 - M_d = 4 - 5 = -1$$

$$d_5 = x_5 - M_d = 6 - 5 = 1$$

$$d_6 = x_6 - M_d = 7 - 5 = 2$$

$$d_7 = x_7 - M_d = 8 - 5 = 3$$

$$d_8 = x_8 - M_d = 9 - 5 = 4$$

Agora, precisamos somar os valores absolutos desses desvios:

$$\sum_{i=1}^{8}|d_i| = |d_1| + |d_2| + |d_3| + |d_4| + |d_5| + |d_6| + |d_7| + |d_8|$$

$$\sum_{i=1}^{8}|d_i| = |-4| + |-3| + |-2| + |-1| + |1| + |2| + |3| + |4|$$

$$\sum_{i=1}^{8}|d_i| = 4 + 3 + 2 + 1 + 1 + 2 + 3 + 4 = 20$$

A propriedade garante que, ao calcularmos a soma dos desvios absolutos em relação à mediana, o menor valor que encontraremos para essa sequência será 20.

Há um detalhe importante que precisamos esclarecer. Como vimos anteriormente, **quando o número de elementos do conjunto é ímpar, o valor da mediana é único e igual ao termo central**. Porém, **quando o número de elementos é par, a mediana pode ser qualquer valor entre os termos centrais, havendo infinitos valores possíveis para a mediana**. **Por convenção, contudo, adotamos a média aritmética dos valores centrais**.

Certo, o que isso tem a ver com a propriedade que estamos estudando? Significa dizer que, se calcularmos a soma dos desvios absolutos para qualquer valor entre 4 e 6, que são os termos centrais, o valor dos desvios absolutos em relação a mediana também será mínimo. A título exemplificativo, vamos calcular os desvios em relação ao valor 4,5:

$$d_1 = x_1 - 4{,}5 = 1 - 4{,}5 = -3{,}5$$

$$d_2 = x_2 - 4{,}5 = 2 - 4{,}5 = -2{,}5$$

$$d_3 = x_3 - 4{,}5 = 3 - 4{,}5 = -1{,}5$$

$$d_4 = x_4 - 4{,}5 = 4 - 4{,}5 = -0{,}5$$

$$d_5 = x_5 - 4{,}5 = 6 - 4{,}5 = 1{,}5$$

$$d_6 = x_6 - 4{,}5 = 7 - 4{,}5 = 2{,}5$$

$$d_7 = x_7 - 4{,}5 = 8 - 4{,}5 = 3{,}5$$

$$d_8 = x_8 - 4{,}5 = 9 - 4{,}5 = 4{,}5$$

Somando os valores absolutos desses desvios:

$$\sum_{i=1}^{8}|d_i| = |d_1| + |d_2| + |d_3| + |d_4| + |d_5| + |d_6| + |d_7| + |d_8|$$

$$\sum_{i=1}^{8} |d_i| = |-3{,}5| + |-2{,}5| + |-1{,}5| + |-0{,}5| + |1{,}5| + |2{,}5| + |3{,}5| + |4{,}5|$$

$$\sum_{i=1}^{8} |d_i| = 3{,}5 + 2{,}5 + 1{,}5 + 0{,}5 + 1{,}5 + 2{,}5 + 3{,}5 + 4{,}5 = 20$$

Como havíamos previsto, o valor também foi igual ao valor mínimo, 20.

Por último, a propriedade também garante que, para qualquer valor fora do intervalo entre 4 e 6, encontraremos um valor maior que o mínimo. Por exemplo, vamos calcular os desvios em relação ao valor 7:

$$d_1 = x_1 - 7 = 1 - 7 = -6$$

$$d_2 = x_2 - 7 = 2 - 7 = -5$$

$$d_3 = x_3 - 7 = 3 - 7 = -4$$

$$d_4 = x_4 - 7 = 4 - 7 = -3$$

$$d_5 = x_5 - 7 = 6 - 7 = -1$$

$$d_6 = x_6 - 7 = 7 - 7 = 0$$

$$d_7 = x_7 - 7 = 8 - 7 = 1$$

$$d_8 = x_8 - 7 = 9 - 7 = 2$$

Somando os valores absolutos desses desvios:

$$\sum_{i=1}^{8} |d_i| = |d_1| + |d_2| + |d_3| + |d_4| + |d_5| + |d_6| + |d_7| + |d_8|$$

$$\sum_{i=1}^{8} |d_i| = |-6| + |-5| + |-4| + |-3| + |-1| + |0| + |1| + |2|$$

$$\sum_{i=1}^{8} |d_i| = 6 + 5 + 4 + 3 + 1 + 0 + 1 + 2 = 22$$

Portanto, como havíamos previsto anteriormente, o valor foi maior que o mínimo.

Podemos resumir as propriedades dos desvios da seguinte forma:

1ª) a soma dos desvios em relação à média aritmética é sempre nula;

2ª) a soma dos quadrados dos desvios em relação à média aritmética é mínima; e

3ª) a soma dos módulos dos desvios em relação à mediana é mínima.

Caso o número de elemento seja par, a soma dos módulos também será mínima se os desvios forem calculados em relação a um dos valores centrais. Isto é, também será mínima a soma dos módulos dos desvios calculados em relação a qualquer termo no intervalo $\left[x_{\frac{n}{2}}, x_{\frac{n}{2}+1}\right]$, em que $x_{\frac{n}{2}}$ e $x_{\frac{n}{2}+1}$ são os termos centrais.

(FCC/TRE-SP/2012) Dado um conjunto de observações, indicadas por X_i $(i = 1, 2, 3, \dots, n)$, o desvio $\boldsymbol{e1}$ da i-ésima observação em relação a um valor $\boldsymbol{\alpha}$ é $e_i = X_i - \alpha$ e $|e_i|$ é o valor absoluto de $\boldsymbol{ei}$. Considere as seguintes afirmações para qualquer conjunto de observações:

I. O valor de $\sum e_i{}^2$ é mínimo se $\boldsymbol{\alpha}$ for igual à média aritmética das observações.

II. O valor de $\sum |e_i|$ é mínimo se $\boldsymbol{\alpha}$ for igual à mediana das observações.

III. O valor de $\sum e_i$ é nulo se $\boldsymbol{\alpha}$ for igual à moda das observações.

IV. O valor de $\sum |e_i|$ é nulo se $\boldsymbol{\alpha}$ for igual à média aritmética das observações.

Então, são corretas APENAS

a) I e II.

b) I e III.

c) II e III.

d) II e IV.

e) II, III e IV.

Comentários:

Vamos analisar cada assertiva:

A sentença I é verdadeira, pois a soma dos quadrados dos desvios é mínima quando os desvios são calculados em relação à média aritmética.

A sentença II também é verdadeira, pois a soma dos módulos dos desvios é mínima quando os desvios são calculados em relação à mediana. Em qualquer situação, quando o desvio é calculado em relação à mediana, a soma dos desvios absolutas é mínima.

A sentença III é falsa, vez que a soma dos módulos dos desvios é nula se os desvios são calculados em relação à média. Somente seria verdadeira caso a moda fosse igual à média.

A sentença IV é falsa, pois a soma dos desvios absolutos em relação à média somente é nula quando todos os desvios também são nulos, ou seja, se todos os números fossem iguais e não houvesse dispersão dos dados.

Gabarito: A.

DESVIO ABSOLUTO MÉDIO

O **desvio absoluto médio**, ou simplesmente desvio médio, **mede a dispersão entre os valores da distribuição e a média dos dados coletados**. Para compreender essa medida, vamos supor que o Estratégia Concursos tenha realizado uma semana de revisão para estudantes da área fiscal, obtendo os seguintes números de visualizações:

Dia da semana	Número de visualizações
Domingo	2.000
Segunda	4.000
Terça	5.200
Quarta	6.300
Quinta	5.400
Sexta	4.100
Sábado	2.400
Total	$\sum f_i = 29.400$

Isso significa que a semana de revisão teve uma média diária de 4.200 visualizações. Esse resultado, porém, não retrata a realidade com fidedignidade, pois alguns dias tiveram mais visualizações do que a média; enquanto outros não. Por isso, é importante sabermos o quão distante a média está em relação aos valores reais por ela representados.

Para calculá-los, basta subtrairmos o valor da média de cada observação, conforme mostrado a seguir:

Dia da semana	Número de visualizações	$x_i - \bar{x}$
Domingo	2.000	$2.000 - 4.200 = -2.200$
Segunda	4.000	$4.000 - 4.200 = -200$
Terça	5.200	$5.200 - 4.200 = 1000$
Quarta	6.300	$6.300 - 4.200 = 2.100$
Quinta	5.400	$5.400 - 4.200 = 1.200$
Sexta	4.100	$4.100 - 4.200 = -100$
Sábado	2.400	$2.400 - 4.200 = -1.800$

Total	$\sum f_i = 29.400$	**0**

Notem que, ao calcularmos o desvio médio, obtemos resultados positivos e negativos, que se anulam ao serem somados. Percebam que existem valores de observações que estão muito próximos da média, enquanto outros estão mais distantes.

Como a soma de todos os desvios médios é sempre igual a zero para qualquer conjunto de dados (1.ª propriedade dos desvios), sabemos que $\sum_{i=1}^{n}(x - \bar{x})$ não nos fornecerá nenhuma informação relevante nem nos ajudará a compreender o que está acontecendo com essa variável.

Para superar essa dificuldade, podemos utilizar apenas os resultados positivos dos desvios calculados. A fórmula do cálculo do desvio médio se apresenta da seguinte maneira:

$$\boxed{\boldsymbol{D_m = \frac{\sum_{i=1}^{n}|x_i - \bar{x}|}{n}}}$$

em que D_m representa o desvio médio, $|x_i - \bar{x}|$ representa o módulo da diferença entre uma determinada observação e a média calculada, f_i representa a frequência de um determinado valor para a variável da distribuição, e n representa o total de elementos formados pela distribuição.

O desvio médio é uma medida de dispersão mais robusta do que a amplitude total e a amplitude interquartílica, pois leva em consideração todos os valores do conjunto. O inconveniente dessa medida é a operação de módulo, que, por conta de suas características matemáticas, torna difícil o estudo de suas propriedades.

Desvio Médio para dados não-agrupados

O desvio absoluto médio (D_m), de um conjunto de n observações $x_1, \cdots, x_n$, é a média dos valores absolutos das diferenças entre as observações e a média. Isto é,

$$\boxed{\boldsymbol{D_m = \frac{\sum_{i=1}^{n}|x_i - \bar{x}|}{n}}}$$

As barras verticais indicam a operação de módulo, que é responsável por transformar qualquer número negativo em um número positivo, isto é, retornar o valor absoluto.

EXEMPLIFICANDO

Calcular o desvio médio do conjunto mostrado a seguir:

$$\{1, 2, 3, 5, 9\}$$

Iniciaremos pelo cálculo da média aritmética:

$$\bar{x} = \frac{1 + 2 + 3 + 5 + 9}{5} = \frac{20}{5} = 4$$

Vamos montar uma tabela para facilitar o cálculo do desvio médio:

$\boldsymbol{x_i}$	$(\boldsymbol{x_i - \bar{x}})$	$\lvert \boldsymbol{x_i - \bar{x}} \rvert$
1	$(1 - 4) = -3$	3
2	$(2 - 4) = -2$	2
3	$(3 - 4) = -1$	1
5	$(5 - 4) = 1$	1
9	$(9 - 4) = 5$	5
		$\boldsymbol{\sum \lvert x_i - \bar{x} \rvert = 12}$

Aplicando a fórmula do desvio médio, temos:

$$D_m = \frac{\sum_{i=1}^{K} |x_i - \bar{x}|}{n} = \frac{\sum_{i=1}^{K} |x_i - 4|}{n} = \frac{12}{5} = 2{,}4$$

(CESPE/ANATEL/2004)

meses	fev	mar	abr	mai	jun	jul	ago	set	out	nov
N	100	70	70	60	50	100	50	50	30	20

A tabela acima mostra os números mensais de reclamações (N) feitas por usuários de telefonia fixa, registradas em uma central de atendimento, entre os meses de fevereiro a novembro de 2003. Considerando esses dados, julgue os itens que se seguem.

O maior desvio absoluto dos números mensais de reclamações registradas é superior a 45.

Comentários:

Iniciaremos pelo cálculo da média:

$$\bar{x} = \frac{100 + 70 + 70 + 60 + 50 + 100 + 50 + 50 + 30 + 20}{10} = \frac{600}{10} = 60$$

Agora, calcularemos o módulo (valor absoluto) de cada um dos desvios.

$$|d_1| = |x_1 - \bar{x}| = |100 - 60| = 40$$
$$|d_2| = |x_2 - \bar{x}| = |70 - 60| = 10$$
$$|d_3| = |x_3 - \bar{x}| = |70 - 60| = 10$$
$$|d_4| = |x_4 - \bar{x}| = |60 - 60| = 0$$
$$|d_5| = |x_5 - \bar{x}| = |50 - 60| = 10$$
$$|d_6| = |x_6 - \bar{x}| = |100 - 60| = 40$$
$$|d_7| = |x_7 - \bar{x}| = |50 - 60| = 10$$
$$|d_8| = |x_8 - \bar{x}| = |50 - 60| = 10$$
$$|d_9| = |x_9 - \bar{x}| = |30 - 60| = 30$$
$$|d_{10}| = |x_{10} - \bar{x}| = |20 - 60| = 40$$

O maior desvio absoluto é 40, portanto, o item está incorreto.

Gabarito: Errado.

(CESPE/ANATEL/2004)

meses	fev	mar	abr	mai	jun	jul	ago	set	out	nov
N	100	70	70	60	50	100	50	50	30	20

A tabela acima mostra os números mensais de reclamações (N) feitas por usuários de telefonia fixa, registradas em uma central de atendimento, entre os meses de fevereiro a novembro de 2003. Considerando esses dados, julgue os itens que se seguem.

O desvio médio absoluto da sequência formada pelos números mensais de reclamações é um valor entre 25 e 35.

Comentários:

Para calcular o desvio absoluto médio, temos que encontrar a média dos valores absolutos (módulos) dos desvios.

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

$$D_m = \frac{40 + 10 + 10 + 0 + 10 + 40 + 10 + 10 + 30 + 40}{10} = \frac{200}{10}$$

$$D_m = 20$$

Gabarito: Errado.

Desvio Médio para dados agrupados sem intervalo de classe

Quando os valores vierem dispostos em uma tabela de frequências, o desvio médio será calculado por meio da seguinte fórmula:

$$\boldsymbol{D_m = \frac{\sum_{i=1}^{m}[\,|x_i - \bar{x}| \times f_i\,]}{\sum f_i}}$$

Em que m indica o número de grupos em que os dados estão organizados; e $|x_i - \bar{x}|$ representa o módulo da diferença entre uma determinada observação e a média calculada.

EXEMPLIFICANDO

Durante uma pesquisa, o Estratégia Concursos registrou a quantidade de filhos de seus professores, obtendo a tabela de frequências apresentada a seguir. Vamos calcular o desvio médio dessa distribuição.

Nº de filhos por professor	f_i	$x_i \times f_i$
0	4	$0 \times 4 = 0$
1	8	$1 \times 8 = 8$
2	4	$2 \times 4 = 8$
3	2	$3 \times 2 = 6$
4	2	$4 \times 2 = 8$
*** Pesquisa populacional**	$\sum f_i = 20$	$\sum x_i \times f_i = 30$

Iniciaremos pelo cálculo da média aritmética:

$$\bar{x} = \frac{\sum x_i \times f_i}{\sum f_i} = \frac{30}{20} = 1{,}50\ filhos\ /\ professor$$

Em seguida, adicionaremos uma nova coluna à tabela anterior, em que calcularemos os produtos dos desvios absolutos por suas respectivas frequências:

Nº de filhos por professor	f_i	$x_i \times f_i$	$\lvert x_i - \bar{x}\rvert \times f_i$
0	4	0	$\lvert 0 - 1{,}5\rvert \times 4 = 6$
1	8	8	$\lvert 1 - 1{,}5\rvert \times 8 = 4$
2	4	8	$\lvert 2 - 1{,}5\rvert \times 4 = 2$
3	2	6	$\lvert 3 - 1{,}5\rvert \times 2 = 3$
4	2	8	$\lvert 4 - 1{,}5\rvert \times 2 = 5$
*** Pesquisa populacional**	$\sum f_i = 20$	$\sum x_i \times f_i = 30$	$\sum \lvert x_i - \bar{x}\rvert \times f_i = 20$

Por fim, aplicando a fórmula do desvio médio, temos:

$$D_m = \frac{\sum|x_i - \overline{x}| \times f_i}{\sum f_i} = \frac{20}{20} = 1$$

Desvio Médio para dados agrupados em classes

Se os dados estiverem agrupados em classe, deveremos adotar a mesma convenção que tomamos para o cálculo da média: vamos assumir que todos os valores coincidem com os pontos médios das suas respectivas classes.

EXEMPLIFICANDO

Durante uma pesquisa, o Estratégia Concursos registrou as estaturas de 40 alunos, obtendo a distribuição de frequências apresentada a seguir. Calcule o desvio médio dessa distribuição.

Estaturas	Frequência (f_i)
150 ⊢ 154	4
154 ⊢ 158	9
158 ⊢ 162	11
162 ⊢ 166	8
166 ⊢ 170	5
170 ⊢ 174	3
*** Pesquisa amostral**	$\sum f_i = 40$

Inicialmente, construiremos uma tabela como a mostrada a seguir:

Estaturas	**Frequência (f_i)**	x_i	$x_i \times f_i$	$(x_i - \bar{x})$	$\lvert x_i - \bar{x} \rvert$	$\lvert x_i - \bar{x} \rvert \times f_i$
150 ⊢ 154	4	152	608	-9	9	36
154 ⊢ 158	9	156	1.404	-5	5	45
158 ⊢ 162	11	160	1.760	-1	1	11
162 ⊢ 166	8	164	1.312	3	3	24
166 ⊢ 170	5	168	840	7	7	35
170 ⊢ 174	3	172	516	11	11	33
*** Pesquisa amostral**	$\sum f_i = 40$		$\sum x_i \times f_i = 6.440$			$\sum \lvert x_i - \bar{x} \rvert \times f_i = 184$

Feito isso, podemos calcular a média da distribuição por meio da seguinte fórmula:

$$\bar{x} = \frac{\sum x_i \times f_i}{\sum f_i} = \frac{6.440}{40} = 161$$

Conhecendo a média, completamos a tabela com as diferenças e os produtos necessários para o cálculo do desvio médio. Assim, aplicando a fórmula do desvio médio, temos:

$$D_m = \frac{\sum_{i=1}^{6} \lvert x_i - \bar{x} \rvert \times f_i}{\sum f_i} = \frac{184}{40} = 4{,}6\ cm$$

Portanto, o desvio médio para essa distribuição de estaturas é $4{,}6\ cm$.

(UEPA/SEFAZ-PA/2013) A tabela abaixo representa as estaturas dos jogadores de voleibol que disputaram a Liga Mundial de 2012.

ESTATURAS (cm)	NÚMERO DE JOADORES
$180 \vdash 190$	10
$190 \vdash 200$	30
$200 \vdash 210$	10
$\sum$	

O desvio médio da estatura dos jogadores é:

a) 2

b) 4

c) 6

d) 8

e) 10

Comentários:

Vamos iniciar pelo cálculo da média. Para isso, construiremos uma coluna com os pontos médios e multiplicaremos cada um pela sua respectiva frequência. Da seguinte forma:

Estaturas	f_i	x_i	$x_i \times f_i$
$180 \vdash 190$	10	185	$10 \times 185 = 1.850$
$190 \vdash 200$	30	195	$30 \times 195 = 5.850$
$200 \vdash 210$	10	205	$10 \times 205 = 2.050$
Total	$\sum f_i = 50$		$\sum x_i \times f_i = 9.750$

Portanto, a média é:

$$\bar{x} = \frac{9.750}{50} = 195$$

Em seguida, adicionaremos uma coluna para calcularmos os módulos dos desvios:

Estaturas	f_i	x_i	$x_i \times f_i$	$\lvert x_i - \bar{x}\rvert$
$180 \vdash 190$	10	185	$10 \times 185 = 1.850$	$\lvert 185 - 195\rvert = 10$
$190 \vdash 200$	30	195	$30 \times 195 = 5.850$	$\lvert 195 - 195\rvert = 0$
$200 \vdash 210$	10	205	$10 \times 205 = 2.050$	$\lvert 205 - 195\rvert = 10$
Total	$\sum f_i = 50$		$\sum x_i \times f_i = 9.750$	

Para calcular o desvio médio, devemos multiplicar cada desvio absoluto pela sua respectiva frequência. Depois, basta somar tudo e dividir por n.

Estaturas	f_i	x_i	$x_i \times f_i$	$\lvert x_i - \bar{x}\rvert$	$\lvert x_i - \bar{x}\rvert \times f_i$
$180 \vdash 190$	10	185	$10 \times 185 = 1.850$	$\lvert 185 - 195\rvert = 10$	$10 \times 10 = 100$
$190 \vdash 200$	30	195	$30 \times 195 = 5.850$	$\lvert 195 - 195\rvert = 0$	$0 \times 30 = 0$
$200 \vdash 210$	10	205	$10 \times 205 = 2.050$	$\lvert 205 - 195\rvert = 10$	$10 \times 10 = 100$
Total	$\sum f_i = 50$		$\sum x_i \times f_i = 9.750$		$\sum \lvert x_i - \bar{x}\rvert \times f_i = 200$

Portanto, o desvio médio é:

$$D_m = \frac{\sum |x_i - \bar{x}| \times f_i}{\sum f_i} = \frac{200}{50} = 4$$

Gabarito: B.

Propriedades do Desvio Médio

Nesse tópico, vamos aprender as principais propriedades do desvio médio.

1ª Propriedade

- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, o desvio médio do conjunto não é alterado.**

EXEMPLIFICANDO

Vamos tomar como exemplo a sequência $\{x_n\} = \{1, 3, 5, 7, 9\}$, cuja desvio médio é:

$$D_m = \frac{|1-5| + |3-5| + |5-5| + |7-5| + |9-5|}{5}$$

$$D_m = \frac{4+2+0+2+4}{5} = \frac{12}{5}$$

Se adicionarmos o número 5 a cada um dos termos da sequência, iremos obter uma nova lista $\{y_n\} = \{x_n + 5\} = \{6, 8, 10, 12, 14\}$, cujo desvio médio é:

$$D_m = \frac{|6-10| + |8-10| + |10-10| + |12-10| + |14-10|}{5}$$

$$D_m = \frac{4+2+0+2+4}{5} = \frac{12}{5}$$

Logo, a adição do número 5 a cada um dos termos da sequência fez com que o desvio médio permanecesse inalterado.

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, o desvio médio do conjunto fica multiplicado (ou dividido) por essa constante.**

EXEMPLIFICANDO

Vamos tomar como exemplo a sequência $\{x_n\} = \{1, 3, 5, 7, 9\}$, cujo desvio médio é:

$$D_m = \frac{|1-5| + |3-5| + |5-5| + |7-5| + |9-5|}{5}$$

$$D_m = \frac{4+2+0+2+4}{5} = \frac{12}{5}$$

Se multiplicarmos cada um dos termos da sequência por 5, iremos obter uma nova lista $\{y_n\} = \{x_n \times 5\} = \{5, 15, 25, 35, 45\}$, cujo desvio médio é:

$$D_m = \frac{|5-25| + |15-25| + |25-25| + |35-25| + |45-25|}{5}$$

$$D_m = \frac{20+10+0+10+20}{5} = \frac{60}{5}$$

Logo, a multiplicação de cada um dos termos da sequência por 5 fez com que o desvio médio do conjunto também fosse multiplicado por 5.

VARIÂNCIA (σ^2)

Existem outras formas de se eliminar o problema com os números negativos. Além da operação de módulo, podemos trabalhar com potências pares. A utilização de potências de expoente par, como o número dois, além de transformar números negativos em positivos, simplifica o cálculo.

A variância é determinada pela média dos quadrados dos desvios em relação à média aritmética. Por meio dessa medida de dispersão ou variabilidade, podemos avaliar o quanto os dados estão dispersos em relação à média aritmética. Nesse sentido, **quanto maior a variância, maior a dispersão dos dados**.

A **variância** leva em consideração a totalidade dos valores da variável em estudo, e não apenas os valores extremos, como faz a amplitude total. Por isso, essa medida de variabilidade **é considerada muito estável**. Além disso, a variância complementa as informações obtidas pelas medidas de tendência central.

Até o momento, as medidas que estudamos não sofriam nenhuma alteração quando o cálculo era realizado para uma amostra. Contudo, para a variância, devemos levar em consideração essa informação, pois há uma **pequena diferença** entre o cálculo da **variância populacional** e da **variância amostral**.

A **variância populacional** é simbolizada pela letra grega σ (sigma), sendo **calculada usando todos os elementos da população**, pela seguinte fórmula:

$$\boldsymbol{\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}}$$

em que: x_i é o valor de ordem i assumido pela variável; μ é a média populacional de x; σ^2 é a variância populacional; e n é o número de dados da população.

A **variância amostral** é simbolizada pela letra s, sendo **calculada a partir de uma amostra da população**, pela seguinte fórmula:

$$\boldsymbol{s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

em que: x_i é o valor de ordem i assumido pela variável; $\bar{x}$ é a média amostral de x; s^2 é a variância amostral; e n é o número de dados da amostra.

Normalmente, uma população possui uma grande quantidade de elementos, o que inviabiliza a realização de um estudo aprofundado de suas medidas, chamadas de **parâmetros populacionais**. Nesse caso, recorremos ao estudo de amostras representativas dessa população, buscando obter indícios do valor correto do parâmetro populacional desconhecido. Esse valor amostral é denominado de **estimador** do parâmetro populacional.

Em nosso caso, a variância populacional cumpre o papel de **parâmetro populacional**, enquanto a variância amostral atua como um **estimador**. Já vimos a variância populacional e a variância amostral são representadas por símbolos diferentes: σ^2 e s^2. O mesmo acontece com a média populacional e a média amostral, que também possuem símbolos diferentes: μ (parâmetro populacional) e $\bar{x}$ (estimador).

Reparem que, quando a variância representa uma descrição da amostra e não da população, caso mais frequente em estatística, o denominador das expressões deve ser $n - 1$, em vez de n. Isso ocorre porque a utilização do divisor $(n - 1)$ resulta em uma melhor estimativa do parâmetro populacional.

Além disso, como a soma dos desvios em relação à média aritmética é sempre nula, apenas $(n - 1)$ dos desvios $(x_i - \bar{x})$ são independentes, vez que $(n - 1)$ desvios determinam automaticamente o valor desconhecido. Para amostras grandes ($n > 30$), não há diferença significativa entre os resultados proporcionados pela utilização de qualquer dos dois divisores, n ou $(n - 1)$.

Em determinadas situações, a aplicação dessas fórmulas pode requerer um esforço considerável. É o caso do que acontece quando a média não é um número natural, situação em que a obtenção da soma dos quadrados dos desvios se torna muito trabalhosa. Por isso, é importante aprendermos outras fórmulas que podem nos ajudar no cálculo da variância.

Já ouviram dizer que **a variância é igual à média dos quadrados menos o quadrado da média**? Pois bem, essa é a fórmula que expressa a **variância populacional**:

$$\boldsymbol{\sigma^2 = \overline{x^2} - (\bar{x})^2}$$

em que $\overline{x^2}$ é a média dos quadrados; e $(\bar{x})^2$ é o quadrado da média.

Como vimos, para encontrarmos a fórmula da variância amostral, basta substituirmos n por $(n - 1)$. Isso é equivalente a multiplicarmos a variância populacional por $\left(\frac{n}{n-1}\right)$. É exatamente o que faremos agora:

$$\boldsymbol{s^2 = \left[\overline{x^2} - (\bar{x})^2\right] \times \left(\frac{n}{n - 1}\right)}$$

em que $\overline{x^2}$ é a média dos quadrados; $(\bar{x})^2$ é o quadrado da média; e n é o tamanho da amostra.

Por fim, é importante ressaltarmos que, por ser calculada a partir dos quadrados dos desvios, a **variância é um número em unidade quadrada em relação à variável em questão**, o que pode ser considerado um inconveniente. Por isso, essa medida tem pouca utilidade na estatística descritiva, mas é extremamente importante na inferência estatística e em combinações de amostras. Por exemplo, se os dados estiverem expressos em quilogramas (Kg), a variância estará expressa em quilogramas ao quadrado (Kg²).

TOME NOTA!

Símbolo da variância populacional:

$$\sigma^2$$

Símbolo da variância amostral:

$$s^2$$

FIQUE ATENTO!

A variância de um conjunto é zero quando todos os elementos são iguais. Se todos os elementos são iguais, a média aritmética do conjunto coincide com o valor dos elementos e todos os desvios também são iguais a zero. Logo, a variância também é zero.

A variância é sempre maior ou igual a zero, isto é, sempre tem valor positivo.

RESUMINDO

Fórmula da variância populacional:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} \quad \text{ou} \quad \sigma^2 = \overline{x^2} - (\bar{x})^2$$

Fórmula da variância amostral:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} \quad \text{ou} \quad s^2 = \left[\overline{x^2} - (\bar{x})^2\right] \times \left(\frac{n}{n-1}\right)$$

Variância para dados não-agrupados

Para dados não agrupados, a variância pode ser expressa por meio das seguintes fórmulas:

a) para populações

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n} \quad ou \quad \sigma^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}}{n}$$

b) para amostras

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n - 1} \quad ou \quad s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}}{n - 1}$$

A relação entre a variância amostral (s^2) e a variância populacional (σ^2) é dada por:

$$s^2 = \left(\frac{n}{n - 1}\right) \times \sigma^2$$

EXEMPLIFICANDO

Calcular a **variância amostral** do conjunto de números mostrado a seguir:

$$\{1, 2, 3, 5, 9\}$$

Iniciaremos pelo cálculo da média aritmética:

$$\bar{x} = \frac{1 + 2 + 3 + 5 + 9}{5} = \frac{20}{5} = 4$$

Agora, vamos montar uma tabela para facilitar o cálculo da variância:

x_i	$(x_i - \overline{x})^2$
1	$(1-4)^2 = 9$
2	$(2-4)^2 = 4$
3	$(3-4)^2 = 1$
5	$(5-4)^2 = 1$
9	$(9-4)^2 = 25$
	$\sum (x_i - \bar{x})^2 = 40$

Por fim, aplicando a fórmula da **variância amostral**, temos:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n - 1} = \frac{40}{5-1} = 10$$

HORA DE PRATICAR!

(VUNESP/TJ-SP/2015) Dados os valores de uma variável: 5, 10, 15, 20, 25, as variâncias amostral e populacional são, respectivamente,

a) 14,7 e 15.

b) 125 e 250.

c) 62,5 e 50.

d) 29,4 e 30,8.

e) 83,3 e 85.

Comentários:

Vamos começar calculando a média:

$$\frac{5 + 10 + 15 + 20 + 25}{5} = 15$$

Agora, vamos encontrar os desvios em relação à média:

$$d_1 = 5 - 15 = -10$$

$$d_2 = 10 - 15 = -5$$

$$d_3 = 15 - 15 = 0$$
$$d_4 = 20 - 15 = 5$$
$$d_5 = 25 - 15 = 10$$

Para calcular a variância (populacional ou amostral), precisamos calcular a soma dos quadrados dos desvios, isto é:

$$\sum d_i^2 = (-10)^2 + (-5)^2 + 0^2 + 5^2 + 10^2$$
$$\sum d_i^2 = 250$$

Nesse momento, dividiremos esse valor por n para encontrar a variância populacional e por $n - 1$ para encontrar a variância amostral:

$$s^2 = \frac{\sum d_i^2}{n - 1} = \frac{250}{5 - 1} = \frac{250}{4} = 62{,}5 \text{ (variância amostral)}$$

$$\sigma^2 = \frac{\sum d_i^2}{n} = \frac{250}{5} = 50 \text{ (variância populacional)}$$

Gabarito: C.

Variância para dados agrupados sem intervalos de classes

Quando os valores vierem dispostos em uma tabela de frequências, a variância será calculada por meio de uma das seguintes fórmulas:

a) para populações

$$\boldsymbol{\sigma^2 = \frac{\sum_{i=1}^{m}(X_i - \mu)^2 \times f_i}{n} \quad ou \quad \sigma^2 = \frac{\sum_{i=1}^{m}(X_i^2 \times f_i) - \frac{(\sum_{i=1}^{m} X_i \times f_i)^2}{n}}{n}}$$

b) para amostras

$$\boldsymbol{s^2 = \frac{\sum_{i=1}^{m}(X_i - \bar{x})^2 \times f_i}{n - 1} \quad ou \quad s^2 = \frac{\sum_{i=1}^{m}(X_i^2 \times f_i) - \frac{(\sum_{i=1}^{m} X_i \times f_i)^2}{n}}{n - 1}}$$

Em que $n = \sum_{i=1}^{m} f_i$ e $\bar{x} = \frac{\sum_{i=1}^{m} X_i \times f_i}{n}$.

EXEMPLIFICANDO

Durante uma pesquisa, o Estratégia Concursos registrou a quantidade de filhos por professor, obtendo a tabela de frequências apresentada a seguir. Sendo assim, calcule a variância amostral dessa tabela.

Nº de filhos por professor	f_i	$x_i \times f_i$
0	4	$0 \times 4 = 0$
1	8	$1 \times 8 = 8$
2	4	$2 \times 4 = 8$
3	2	$3 \times 2 = 6$
4	2	$4 \times 2 = 8$
*** Pesquisa populacional**	$\sum f_i = 20$	$\sum x_i \times f_i = 30$

Iniciaremos pelo cálculo da média aritmética:

$$\bar{x} = \frac{\sum x_i \times f_i}{\sum f_i} = \frac{30}{20} = 1{,}50\ filhos\ /\ professor$$

Em seguida, adicionaremos uma nova coluna à tabela anterior, em que calcularemos os produtos dos quadrados dos desvios por suas respectivas frequências:

Nº de filhos por professor	f_i	$x_i \times f_i$	$(x_i - \bar{x})^2 \times f_i$
0	4	0	$(0 - 1{,}5)^2 \times 4 = 9$
1	8	8	$(1 - 1{,}5)^2 \times 8 = 2$
2	4	8	$(2 - 1{,}5)^2 \times 4 = 1$
3	2	6	$(3 - 1{,}5)^2 \times 2 = 4{,}5$
4	2	8	$(4 - 1{,}5)^2 \times 2 = 12{,}5$
*** Pesquisa populacional**	$\sum f_i = 20$	$\sum x_i \times f_i = 30$	$\sum (x_i - x)^2 \times f_i = 29$

Por fim, aplicando a fórmula do desvio padrão amostral, temos:

$$s^2 = \frac{\sum(x_i - \bar{x})^2 \times f_i}{n-1} = \frac{29}{19} = 1{,}52$$

Variância para dados agrupados em classes

Para dados contínuos agrupados em classes, a variância é calculada por meio das seguintes expressões:

a) para populações

$$\boldsymbol{\sigma^2 = \frac{\sum_{i=1}^{k}(PM_i - \mu)^2 \times f_i}{n} \quad ou \quad \sigma^2 = \frac{\sum_{i=1}^{k} PM_i^2 \times f_i - \frac{\left(\sum_{i=1}^{k} PM_i \times f_i\right)^2}{n}}{n}}$$

b) para amostras

$$\boldsymbol{s^2 = \frac{\sum_{i=1}^{k}(PM_i - \bar{x})^2 \times f_i}{n - 1} \quad ou \quad s^2 = \frac{\sum_{i=1}^{k} PM_i^2 \times f_i - \frac{\left(\sum_{i=1}^{k} PM_i \times f_i\right)^2}{n}}{n - 1}}$$

Em que $n = \sum_{i=1}^{k} f_i$ e $\bar{x} = \frac{\sum_{i=1}^{k} PM_i \times f_i}{n}$.

Observem que as fórmulas são praticamente iguais as apresentadas no subtópico anterior. A diferença básica é que agora vamos utilizar o ponto médio das k classes.

EXEMPLIFICANDO

Durante uma pesquisa, o Estratégia Concursos registrou as estaturas de 40 alunos, obtendo a distribuição de frequências apresentada a seguir. Vamos calcular a variância amostral dessa distribuição.

Estaturas	Frequência (f_i)
150 ⊢ 154	4
154 ⊢ 158	9
158 ⊢ 162	11
162 ⊢ 166	8
166 ⊢ 170	5
170 ⊢ 174	3
*** Pesquisa amostral**	$\sum f_i = 40$

Inicialmente, construiremos uma tabela como a mostrada a seguir:

Estaturas	Frequência (f_i)	x_i	$x_i \times f_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \times f_i$
150 ⊢ 154	4	152	608	-9	81	324
154 ⊢ 158	9	156	1.404	-5	25	225
158 ⊢ 162	11	160	1.760	-1	1	11
162 ⊢ 166	8	164	1.312	3	9	72
166 ⊢ 170	5	168	840	7	49	245
170 ⊢ 174	3	172	516	11	121	363
*** Pesquisa populacional**	$\sum f_i = 40$		$\sum x_i \times f_i = 6.440$			$\sum (x_i - \bar{x})^2 \times f_i = 1.240$

Feito isso, podemos calcular a média da distribuição por meio da seguinte fórmula:

$$\bar{x} = \frac{\sum x_i \times f_i}{\sum f_i} = \frac{6.440}{40} = 161$$

Conhecendo a média, completamos a tabela com as diferenças e os produtos necessários para o cálculo da variância. Agora, aplicando a fórmula da variância amostral, temos:

$$s^2 = \frac{\sum_{i=1}^{k}(x_i - \bar{x})^2 \times f_i}{n - 1} = \frac{\sum_{i=1}^{6}(x_i - 161)^2 \times f_i}{40 - 1} = \frac{1.240}{39} = 31{,}79\ cm^2$$

A variância amostral das estaturas é $31,79\ cm^2$.

Propriedades do Variância

Nesse tópico, vamos aprender as principais propriedades da variância.

1ª Propriedade

- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, a variância do conjunto não é alterada.**

EXEMPLIFICANDO

Vamos tomar como exemplo a sequência $\{x_n\} = \{1, 3, 5, 7, 9\}$, cuja variância é:

$$\sigma^2 = \frac{(1-5)^2 + (3-5)^2 + (5-5)^2 + (7-5)^2 + (9-5)^2}{5}$$

$$\sigma^2 = \frac{16+4+0+4+16}{5} = 8$$

Se adicionarmos o número 5 a cada um dos termos da sequência, iremos obter uma nova lista $\{y_n\} = \{x_n + 5\} = \{6, 8, 10, 12, 14\}$, cuja variância é:

$$\sigma^2 = \frac{(6-10)^2 + (8-10)^2 + (10-10)^2 + (12-10)^2 + (14-10)^2}{5}$$

$$\sigma^2 = \frac{16+4+0+4+16}{5} = 8$$

Logo, a adição do número 5 a cada um dos termos da sequência fez com que a variância permanecesse inalterada.

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, a variância do conjunto fica multiplicada (ou dividida) pelo QUADRADO dessa constante.**

EXEMPLIFICANDO

Vamos tomar como exemplo a sequência $\{x_n\} = \{1, 3, 5, 7, 9\}$, cuja variância é:

$$\sigma^2 = \frac{(1-5)^2 + (3-5)^2 + (5-5)^2 + (7-5)^2 + (9-5)^2}{5}$$

$$\sigma^2 = \frac{16 + 4 + 0 + 4 + 16}{5} = 8$$

Se multiplicarmos cada um dos termos da sequência por 5, iremos obter uma nova lista $\{y_n\} = \{x_n \times 5\} = \{5, 15, 25, 35, 45\}$, cuja variância é:

$$\sigma^2 = \frac{(5-25)^2 + (15-25)^2 + (25-25)^2 + (35-25)^2 + (45-25)^2}{5}$$

$$\sigma^2 = \frac{400 + 100 + 0 + 100 + 400}{5} = 200$$

Logo, a multiplicação de cada um dos termos da sequência por 5 fez com que a variância do conjunto fosse multiplicada por $5^2 = 25$.

DESVIO-PADRÃO (σ)

O desvio padrão (s ou σ) é definido como sendo a raiz quadrada da média aritmética dos quadrados dos desvios e, dessa forma, **é determinado pela raiz quadrada da variância. É uma das medidas de variabilidade mais utilizadas porque é capaz de apontar de forma mais precisa a dispersão dos valores em relação à média aritmética.**

Valores muito próximos da média resultarão em um desvio-padrão pequeno, enquanto valores mais espalhados levarão a desvios maiores. Essa medida será sempre maior ou igual a zero. Ela será igual a zero quando todos os elementos do conjunto forem iguais.

O desvio padrão é utilizado para comparar a variabilidade de dois conjuntos de dados diferentes quando as médias forem aproximadamente iguais e quando as unidades de medidas para os dois conjuntos forem idênticas.

A fórmula para o cálculo do desvio padrão populacional é:

$$\boldsymbol{\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}}}$$

Para o desvio padrão amostral, a fórmula é a seguinte:

$$\boldsymbol{s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}}}$$

Como vimos no tópico anterior, a utilização do divisor $(n-1)$ resulta em uma melhor estimativa do parâmetro populacional. Além disso, como a soma dos desvios em relação à média aritmética é sempre nula, apenas $(n-1)$ dos desvios $(x_i - \bar{x})$ são independentes, uma vez que esses $(n-1)$ desvios determinam automaticamente o valor desconhecido.

Por fim, **o desvio-padrão é expresso nas mesmas unidades dos dados originais**. Tanto o desvio padrão como a variância são usados como medidas de dispersão ou variabilidade. O uso de uma medida ou de outra dependerá da finalidade que se tiver em mente.

TOME NOTA!

Símbolo do desvio-padrão populacional:

$$\sigma$$

Símbolo do desvio-padrão amostral:

$$s$$

FIQUE ATENTO!

O desvio-padrão será igual a zero quando todos os elementos forem iguais. Se todos os elementos forem iguais, a média aritmética do conjunto será igual ao valor dos elementos e todos os desvios também serão iguais a zero. Logo, o desvio-padrão também será zero.

O desvio-padrão é sempre maior ou igual a zero, isto é, sempre tem valor positivo.

RESUMINDO

Fórmula do desvio-padrão populacional:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}}$$

Fórmula do desvio-padrão amostral:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

(VUNESP/ARTESP/2018) Numa série composta por n dados, todos de mesmo valor x ($x \neq 0$), o valor do desvio padrão s é:

a) $s = \frac{n}{x}$

b) $s = 0$

c) $s = \frac{nx}{2}$

d) $s = x$

e) $s = 1$

Comentários:

Como todos os dados são iguais, todos os desvios são nulos. Consequentemente, os quadrados dos desvios também são nulos. Logo, a variância e o desvio-padrão serão iguais a zero.

Gabarito: B.

(UFMT/Pref. de Cáceres-MT/2017) Um conjunto de dados sobre a plaquetopenia de pacientes com dengue tem variância igual a zero. Pode-se concluir que também vale zero

a) a média.

b) o desvio padrão.

c) a mediana.

d) a moda.

Comentários:

O desvio-padrão é a raiz quadrada da variância. Nesse caso, como a variância é igual a zero, então o desvio-padrão vale:

$$\sigma = \sqrt{0} = 0.$$

Gabarito: B.

Desvio-padrão para dados não-agrupados

Para dados não agrupados, o desvio-padrão pode ser expresso por meio das seguintes fórmulas:

a) para populações

$$\boldsymbol{\sigma = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}}}$$

b) para amostras

$$\boldsymbol{s = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}}$$

EXEMPLIFICANDO

Vamos calcular o desvio-padrão amostral do conjunto de números mostrado a seguir:

$$\{1, 2, 3, 5, 9\}$$

Iniciaremos pelo cálculo da média aritmética:

$$\bar{x} = \frac{1+2+3+5+9}{5} = \frac{20}{5} = 4$$

Em seguida, montaremos uma tabela para facilitar o cálculo do desvio padrão:

x_i	$(x_i - \bar{x})^2$
1	$(1-4)^2 = 9$
2	$(2-4)^2 = 4$
3	$(3-4)^2 = 1$
5	$(5-4)^2 = 1$
9	$(9-4)^2 = 25$
	$\sum (x_i - \bar{x})^2 = 40$

Por fim, aplicando a fórmula do desvio padrão temos:

$$s = \sqrt{\frac{40}{5-1}} = \sqrt{10} \cong 3{,}16$$

HORA DE PRATICAR!

(FCC/ARTESP/2017) O departamento de operações de uma autarquia do Estado fez um levantamento do número de acidentes em um determinado trecho de rodovia no ano de 2016, conforme tabela a seguir.

Mês	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Nº de Acidentes	36	28	12	5	3	2	2	4	9	11	22	38

Os números indicam que há uma dispersão significativa, portanto, o desvio padrão para esta amostra é representado por

a) 13,30.

b) 14,33.

c) 12,74.

d) 10,40.

e) 11,50.

Comentários:

Vamos iniciar calculando a média:

$$\bar{x} = \frac{36+28+12+5+3+2+2+4+9+11+22+38}{12}$$

$$\bar{x} = \frac{172}{12}$$

$$\bar{x} = \frac{43}{3} = 14{,}33$$

Agora, vamos montar uma tabela para simplificar o cálculo da média dos quadrados:

Valor (x)	X^2
36	1.296
28	784
12	144
5	25
3	9
2	4
2	4
4	16
9	81
11	121
22	484
38	1.444
Total	**4.412**

Portanto, a média dos quadrados é:

$$\overline{x^2} = \frac{4.412}{12} = 367{,}67$$

A variância populacional é dada pela diferença entre a média dos quadrados e o quadrado da média:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

$$\sigma^2 = 367{,}67 - (14{,}33)^2$$

$$\sigma^2 = 367{,}67 - 205{,}35$$

$$\sigma^2 = 162{,}32$$

Se multiplicarmos a variância populacional por $\frac{n}{n-1}$, encontraremos a variância amostral:

$$s^2 = 162{,}32 \times \frac{12}{11}$$

$$s^2 = 162{,}32 \times 1{,}09$$

$$s^2 = 177{,}07$$

$$s = \sqrt{177{,}07}$$

$$s = 13{,}30$$

Gabarito: A.

(CESPE/Polícia Federal/2018)

	dia				
	1	2	3	4	5
X (quantidade diária de drogas apreendidas, em kg)	10	22	18	22	28

Tendo em vista que, diariamente, a Polícia Federal apreende uma quantidade X, em kg, de drogas em determinado aeroporto do Brasil, e considerando os dados hipotéticos da tabela precedente, que apresenta os valores observados da variável X em uma amostra aleatória de 5 dias de apreensões no citado aeroporto, julgue o próximo item.

O desvio padrão amostral da variável *X* foi inferior a 7

Comentários:

Começaremos calculando a média:

$$\frac{10+22+18+22+28}{5} = 20$$

Agora, vamos encontrar os desvios:

$$d_1 = 10 - 20 = -10$$

$$d_2 = 22 - 20 = 2$$

$$d_3 = 18 - 20 = -2$$

$$d_4 = 22 - 20 = 2$$

$$d_5 = 28 - 20 = 8$$

Para calcular a variância (populacional ou amostral), precisamos calcular a soma dos quadrados dos desvios, isto é:

$$\sum d_i^2 = (-10)^2 + 2^2 + (-2)^2 + 2^2 + 8^2$$

$$\sum d_i^2 = 176$$

Nesse momento, dividiremos esse valor por $n-1$ para encontrarmos a variância amostral:

$$s^2 = \frac{\sum d_i^2}{n-1} = \frac{176}{5-1} = \frac{176}{4} = 44$$

E, por fim, o desvio padrão é a raiz quadrada da variância:

$$s = \sqrt{44}$$

O enunciado diz que esse valor é menor do que 7kg. De fato, sabemos que $7^2 = 49$, logo $\sqrt{44} < 7$.

Gabarito: Certo.

Desvio-padrão para dados agrupados sem intervalo de Classe

Quando os valores vierem dispostos em uma tabela de frequências, o desvio-padrão será calculado por meio de uma das seguintes fórmulas:

a) para populações

$$\boldsymbol{\sigma} = \sqrt{\frac{\sum_{i=1}^{m}(d_i^2 \times f_i)}{n}} = \sqrt{\frac{\sum_{i=1}^{m}[(X_i - \mu)^2 \times f_i]}{n}}$$

b) para amostras

$$\boldsymbol{s} = \sqrt{\frac{\sum_{i=1}^{m}(d_i^2 \times f_i)}{n-1}} = \sqrt{\frac{\sum_{i=1}^{m}[(X_i - \overline{x})^2 \times f_i]}{n-1}}$$

Em que $n = \sum_{i=1}^{m} f_i$ e $\bar{x} = \frac{\sum_{i=1}^{m} X_i \times f_i}{n}$.

EXEMPLIFICANDO

Durante uma pesquisa, o Estratégia Concursos registrou a quantidade de filhos de seus professores, obtendo a tabela de frequências apresentada a seguir. Vamos calcular o desvio-padrão amostral dessa distribuição.

Nº de filhos por professor	f_i	$x_i \times f_i$
0	4	$0 \times 4 = 0$
1	8	$1 \times 8 = 8$
2	4	$2 \times 4 = 8$
3	2	$3 \times 2 = 6$
4	2	$4 \times 2 = 8$
*** Pesquisa populacional**	$\sum f_i = 20$	$\sum x_i \times f_i = 30$

Iniciaremos pelo cálculo da média aritmética:

$$\bar{x} = \frac{\sum X_i \times f_i}{\sum f_i} = \frac{30}{20} = 1{,}50\ filhos\ /\ professor$$

Em seguida, adicionaremos uma nova coluna à tabela anterior, em que calcularemos os produtos dos quadrados dos desvios por suas respectivas frequências:

Nº de filhos por professor	f_i	$x_i \times f_i$	$(x_i - \overline{x})^2 \times f_i$
0	4	0	$(0 - 1{,}5)^2 \times 4 = 9$
1	8	8	$(1 - 1{,}5)^2 \times 8 = 2$
2	4	8	$(2 - 1{,}5)^2 \times 4 = 1$
3	2	6	$(3 - 1{,}5)^2 \times 2 = 4{,}5$
4	2	8	$(4 - 1{,}5)^2 \times 2 = 12{,}5$
*** Pesquisa populacional**	$\sum f_i = 20$	$\sum x_i \times f_i = 30$	$\sum (x_i - x)^2 \times f_i = 29$

Por fim, aplicando a fórmula do desvio padrão amostral, temos:

$$s = \sqrt{\frac{\sum (x_i - \overline{x})^2 \times f_i}{n - 1}}$$

$$s = \sqrt{\frac{29}{19}} = \sqrt{1{,}52} \cong 1{,}23$$

Desvio-padrão para dados agrupados em classes

Quando tivermos que calcular o desvio-padrão para dados agrupados em classes, usaremos as mesmas fórmulas para dados sem intervalos de classes, utilizando para x_i os pontos médios de cada classe, mas adotando os mesmos procedimentos.

EXEMPLIFICANDO

Durante uma pesquisa, o Estratégia Concursos registrou as estaturas de 40 alunos, obtendo a distribuição de frequências apresentada a seguir. Vamos calcular o desvio-padrão amostral dessa distribuição.

Estaturas	Frequência (f_i)
150 ⊢ 154	4
154 ⊢ 158	9
158 ⊢ 162	11
162 ⊢ 166	8
166 ⊢ 170	5
170 ⊢ 174	3
*** Pesquisa amostral**	$\sum f_i = 40$

Inicialmente, construiremos uma tabela como a mostrada a seguir:

Estaturas	Frequência (f_i)	x_i	$x_i \times f_i$	$(x_i - \overline{x})$	$(x_i - \overline{x})^2$	$(x_i - \overline{x})^2 \times f_i$
150 ⊢ 154	4	152	608	-9	81	324
154 ⊢ 158	9	156	1.404	-5	25	225
158 ⊢ 162	11	160	1.760	-1	1	11
162 ⊢ 166	8	164	1.312	3	9	72
166 ⊢ 170	5	168	840	7	49	245
170 ⊢ 174	3	172	516	11	121	363
*** Pesquisa populacional**	$\sum f_i = 40$		$\sum x_i \times f_i = 6.440$			$\sum (x_i - \overline{x})^2 \times f_i = 1.240$

Feito isso, podemos calcular a média da distribuição por meio da seguinte fórmula:

$$\bar{x} = \frac{\sum PM_i \times f_i}{\sum f_i} = \frac{6.440}{40} = 161$$

Conhecendo a média, completamos a tabela com as diferenças e os produtos necessários para o cálculo do desvio padrão. Agora, aplicando a fórmula do desvio padrão amostral, temos:

$$s = \sqrt{\frac{\sum_{i=1}^{k}(PM_i - \bar{x})^2 \times f_i}{n-1}} = \sqrt{\frac{\sum_{i=1}^{6}(PM_i - 161)^2 \times f_i}{40-1}} = \sqrt{\frac{1.240}{39}} = \sqrt{31{,}79} \cong 5{,}64\ cm$$

O desvio-padrão das estaturas é $5{,}64\ cm$. Vimos anteriormente que o desvio médio, para essa mesma distribuição, foi de $4{,}63\ cm$.

Propriedades do Desvio-padrão

Nesse tópico, vamos estudar as principais propriedades do desvio-padrão.

1ª Propriedade

- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, o desvio-padrão do conjunto não é alterado.**

EXEMPLIFICANDO

Vamos tomar como exemplo a sequência $\{x_n\} = \{1, 3, 5, 7, 9\}$, cujo desvio-padrão é:

$$\sigma = \sqrt{\frac{(1-5)^2 + (3-5)^2 + (5-5)^2 + (7-5)^2 + (9-5)^2}{5}}$$

$$\sigma = \sqrt{\frac{16+4+0+4+16}{5}} = 2\sqrt{2}$$

Se adicionarmos o número 5 a cada um dos termos da sequência, iremos obter uma nova lista $\{y_n\} = \{x_n + 5\} = \{6, 8, 10, 12, 14\}$, cujo desvio-padrão é:

$$\sigma = \sqrt{\frac{(6-10)^2 + (8-10)^2 + (10-10)^2 + (12-10)^2 + (14-10)^2}{5}}$$

$$\sigma = \sqrt{\frac{16+4+0+4+16}{5}} = 2\sqrt{2}$$

Logo, a adição do número 5 a cada um dos termos da sequência fez com que o desvio-padrão permanecesse inalterado.

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, o desvio-padrão do conjunto fica multiplicado (ou dividido) por essa constante.**

EXEMPLIFICANDO

Vamos tomar como exemplo a sequência $\{x_n\} = \{1, 3, 5, 7, 9\}$, cujo desvio-padrão é:

$$\sigma = \sqrt{\frac{(1-5)^2 + (3-5)^2 + (5-5)^2 + (7-5)^2 + (9-5)^2}{5}}$$

$$\sigma = \sqrt{\frac{16 + 4 + 0 + 4 + 16}{5}} = 2\sqrt{2}$$

Se multiplicarmos cada um dos termos da sequência por 5, iremos obter uma nova lista $\{y_n\} = \{x_n \times 5\} = \{5, 15, 25, 35, 45\}$, cujo desvio-padrão é:

$$\sigma = \sqrt{\frac{(5-25)^2 + (15-25)^2 + (25-25)^2 + (35-25)^2 + (45-25)^2}{5}}$$

$$\sigma = \sqrt{\frac{400 + 100 + 0 + 100 + 400}{5}} = \sqrt{\frac{200}{5}} = 10\sqrt{2}$$

Logo, a multiplicação de cada um dos termos da sequência por 5 fez com que o desvio-padrão do conjunto também fosse multiplicado por 5.

Coeficiente de Variação (ou Dispersão Relativa)

O desvio-padrão pode ser utilizado para a comparação de duas ou mais séries de valores, no que diz respeito à variabilidade e dispersão, quando os conjuntos possuem a mesma média e estão expressos na mesma unidade de medida (p.ex., os dois conjuntos em centímetros). Porém, quando os conjuntos de dados estão expressos em unidades diferentes (p.ex., quilogramas e centímetros), precisamos de outra medida.

Para contornar essa limitação do desvio-padrão, podemos caracterizar a dispersão ou variabilidade dos dados de maneira relativa ao seu valor médio. Nesse sentido, **o coeficiente de variação é uma medida de dispersão relativa que fornece a variação dos dados em relação à média**, podendo ser calculado como:

a) para populações

$$\boldsymbol{CV = \frac{\sigma}{\mu} \times 100\ (\%)}$$

b) para amostras

$$\boldsymbol{CV = \frac{s}{\bar{x}} \times 100\ (\%)}$$

em que: σ é o desvio-padrão populacional; μ é a média populacional; s é o desvio-padrão amostral; e $\bar{x}$ é a média amostral.

O **coeficiente de variação** pode ser interpretado por meio de algumas regras empíricas:

a) a distribuição tem **baixa dispersão** se $CV < 15\%$;

b) a distribuição tem **média dispersão** se $15\% < CV < 30\%;\ e$

c) a distribuição tem **elevada dispersão** se $CV > 30\%$.

Além disso, quanto menor for o valor do **coeficiente de variação**, mais homogêneos serão os dados, ou seja, menor será a dispersão em torno da média. Por isso, podemos classificar as distribuições em homogêneas ou heterogêneas, da seguinte forma:

a) a distribuição é **homogênea** quando possui dispersão baixa ou média ($CV < 30\%$);

b) a distribuição é **heterogênea** quando possui dispersão elevada ($CV > 30\%$).

EXEMPLIFICANDO

Em uma empresa de tecnologia, o salário médio dos homens é de R$ 1800,00 com desvio-padrão de R$ 810,00 e o salário médio das mulheres é de R$ 1500,00 com desvio padrão de R$ 705,00. A dispersão relativa dos salários dos homens é maior que a das mulheres?

Vamos identificar os dados do problema:

a) para os homens:

$$\begin{cases} \mu_H = 1800 \\ \sigma_H = 810 \end{cases}$$

b) para as mulheres:

$$\begin{cases} \mu_M = 1500 \\ \sigma_M = 705 \end{cases}$$

Agora, vamos calcular os respectivos coeficientes de variação:

a) para os homens:

$$CV = \frac{\sigma_H}{\mu_H} \times 100 = \frac{810}{1800} = 45{,}0\%$$

b) para as mulheres:

$$CV = \frac{\sigma_M}{\mu_M} \times 100 = \frac{705}{1500} = 47{,}0\%$$

Portanto, os salários das mulheres apresentam uma dispersão relativa maior que os salários dos homens. Além disso, as duas distribuições possuem uma alta dispersão ($CV > 30\%$).

(FCC/ALAP/2020) O número de empregados de uma empresa é igual a 200, sendo que 60% são homens e o restante mulheres. Nesta empresa, a média aritmética dos salários da população formada pelos salários dos homens é igual a 5 mil reais, com um coeficiente de variação igual a 30%, e a média aritmética dos salários da população formada pelos salários das mulheres também é igual a 5 mil reais, porém com um coeficiente de variação igual a 20%. Considerando a população formada por todos os 200 empregados da empresa, obtém-se que a variância, em mil reais ao quadrado, dos respectivos salários é igual a

a) 1,69

b) 1,75

c) 1,30

d) 2,50

e) 3,25

Comentários:

Para responder essa questão, encontraremos os dados considerando separadamente os homens e depois faremos o mesmo processo para as mulheres. Ao final, acharemos o que foi pedido para a população $N = 200$.

Segundo a questão, a população tem tamanho igual a 200, isto é, $N = 200$. Dessa população de empregados, temos que 60% são homens, ou seja:

$$60\% \times 200 = 120\, homens.$$

Consequentemente, o número de mulheres será:

$$200 - 120 = 80\, mulheres.$$

De acordo com o enunciado, a média aritmética dos salários da população tem coeficiente de variação igual a 30%, isto é, $CV_{homens} = 30\%$. Esse coeficiente é calculado por meio da seguinte fórmula:

$$CV_{homens} = \frac{\sigma}{\mu}$$

A questão nos informou que a média salarial dos homens é de 5 mil reais, ou seja, $\mu = 5\ (mil\ reais)$. Logo, usando a fórmula acima, conseguiremos encontrar o desvio padrão "populacional" dos homens:

$$30\% = \frac{\sigma_{homens}}{5}$$

$$\sigma_{homens} = 30\% \times 5$$

$$\sigma_{homens} = 1{,}5\ (mil\ reais)$$

Sabemos que a variância é o quadrado do desvio padrão, então:

$$\sigma^2_{homens} = (1{,}5)^2 = 2{,}25(mil\ reais)^2$$

Adotaremos o mesmo procedimento para as mulheres. A questão nos informou que a média salarial das mulheres é de 5 mil reais, ou seja, $\mu = 5\ (mil\ reais)$. A única diferença é que o coeficiente de variação das mulheres é igual a 20%, $CV_{mulheres} = 20\%$.

$$CV_{mulheres} = \frac{\sigma_{mulheres}}{\mu}$$

$$20\% = \frac{\sigma_{mulheres}}{5}$$

$$\sigma_{mulheres} = 1{,}00\ (mil\ reais)$$

A variância é o quadrado do desvio padrão, então:

$$\sigma^2_{mulheres} = (1{,}00)^2 = 1{,}00(mil\ reais)^2$$

Agora, consideraremos toda a população $N = 200$. A média populacional dos salários dos 200 empregados será 5 mil, já que tanto a média salarial dos homens quanto a média salarial das mulheres é igual é igual a 5 mil reais. Portanto:

$$\bar{x} = 5\ (mil\ reais)$$

Agora, para encontrar a variância, vamos utilizar a fórmula clássica da variância:

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{N}$$

Buscaremos o termo $\sum(x_i - \mu)^2$ para homens e mulheres, lembrando sempre que a média é igual a $5\ (mil\ reais)$, tanto para homens quanto para mulheres.

Calculando para os homens:

$$\sigma^2_{homens} = \frac{\sum(x_i - \bar{x})^2}{N}$$

$$2{,}25 = \frac{\sum(x_i - \bar{x})^2}{120}$$

$$\sum(x_i - \mu)^2 = 2{,}25 \times 120 = 270$$

Calculando para as mulheres:

$$\sigma^2_{mulheres} = \frac{\sum(x_i - \bar{x})^2}{N}$$

$$1{,}00 = \frac{\sum(x_i - \bar{x})^2}{80}$$

$$\sum(x_i - \mu)^2 = 1{,}00 \times 80 = 80$$

Agora, substituiremos esses valores na variância de toda a população, considerando $N = 200$:

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{N}$$

$$\sigma^2 = \frac{270 + 80}{200} = \frac{350}{200} = 1{,}75\ (mil\ reais)^2$$

Gabarito: B.

(FCC/Metrô-SP/2019) Uma empresa possui 40 funcionários dos quais F1 são mulheres e F2 são homens. Sabe-se que a média salarial das mulheres é de 8 salários mínimos, que a média salarial dos homens é de 10 salários mínimos e que a média salarial de todos os 40 funcionários é de 8,6 salários mínimos. Se a variância dos salários dos funcionários do sexo masculino é igual a (F2 + 4) (salários mínimos)², o coeficiente de variação desses funcionários do sexo masculino é igual a

a) 32%.

b) 25%.

c) 36%.

d) 40%.

e) 15%

Comentários:

Conforme o enunciado, uma empresa possui um total de 40 funcionários, sendo um subtotal F_1 de mulheres e um subtotal F_2 de homens. Logo,

$$F_1 + F_2 = 40 \ (Equação\ 1)$$

De acordo com a questão, a média salarial das mulheres é 8, enquanto a média salarial dos homens é 10.

$$\bar{x}_{mulheres} = 8$$

$$\bar{x}_{homens} = 10$$

$$\bar{x} = 8{,}6$$

Calculando a média dos salários para homens e mulheres :

$$\frac{\Sigma(salário\ mulheres)}{F_1} = 8$$

$$\Sigma(salário\ mulheres) = 8 \times F_1$$

$$\frac{\Sigma(salário\ homens)}{F_2} = 10$$

$$\Sigma(salário\ homens) = 10 \times F_2$$

A média total pode ser calculada por meio da seguinte expressão:

$$\bar{x} = \frac{\Sigma(salário\ mulheres) + \Sigma(salário\ homens)}{F_1 + F_2}$$

$$8{,}6 = \frac{8 \times F_1 + 10 \times F_2}{F_1 + F_2}$$

$$8{,}6 \times F_1 + 8{,}6 \times F_2 = 8 \times F_1 + 10 \times F_2$$

$$0{,}6 \times F_1 = 1{,}4 \times F_2 \ (Equação\ 2)$$

Chegamos, portanto, a uma situação em que temos duas equações e duas incógnitas (F_1 e F_2). Podemos isolar a variável F_1 na Equação 2 e, em seguida, substituí-la na Equação 1, chegando ao valor de F_2.

$$F_1 = \frac{1{,}4 \times F_2}{0{,}6} = \left(\frac{7}{3}\right) \times F_2$$

Substituindo a variável F_1 na Equação 1, chegamos ao valor de F_2.

$$F_1 + F_2 = 40 \ (Equação\ 1)$$

$$\left(\frac{7}{3}\right) \times F_2 + F_2 = 40$$

Multiplicando todos os termos por 3, temos:

$$7 \times F_2 + 3 \times F_2 = 120$$

$$10 \times F_2 = 120$$

$$F_2 = 12$$

Portanto, o número de homens é 12.

O enunciado também forneceu a variância, que é equivalente à expressão (F2+4). Isto é:

$$\sigma^2 = F_2 + 4 = 12 + 4 = 16.$$

Então, o desvio padrão será a raiz quadrada da variância:

$$\sigma = 4.$$

O coeficiente de variação (CV) para os homens será:

$$CV = \frac{\sigma}{\bar{x}_{homens}} = \frac{4}{10} = 40\%$$

Gabarito: D.

(FCC/TRT 20ª Região/2016) Em uma associação de determinada carreira profissional é realizado um censo em que foram apurados os salários de todos os seus 320 associados em número de salários mínimos (S.M.). O coeficiente de variação correspondente foi de 16% e a soma dos quadrados de todos os salários, em (S.M.)2, foi de 8.204,80. O desvio padrão dos salários destes associados é, em S.M., de

a) 0,80

b) 0,64

c) 0,96

d) 0,40

e) 1,60

Comentários:

O coeficiente de variação foi informado na questão. Sabemos que ele é resultado da divisão entre o desvio padrão e a média, então:

$$\frac{16}{100} = \frac{\sigma}{\bar{x}}$$

$$\bar{x} = \frac{100\sigma}{16}$$

A variância resulta da diferença entre a média dos quadrados e o quadrado da média. Vamos aplicar o valor da média na fórmula da variância:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

$$\sigma^2 = \frac{8.204,80}{320} - \left(\frac{100\sigma}{16}\right)^2$$

$$\sigma^2 = \frac{8.204,80}{320} - \frac{10.000\sigma^2}{256}$$

$$\sigma^2 = \frac{32.819,2 - 50.000\sigma^2}{1280}$$

$$1280\sigma^2 = 32.819,2 - 50.000\sigma^2$$

$$51280\sigma^2 = 32.819,2$$

$$\sigma^2 = \frac{32.819,2}{51280}$$

$$\sigma^2 = 0,64$$

$$\sigma = \sqrt{0,64}$$

$$\sigma = 0,8$$

Gabarito: A.

VARIÂNCIA RELATIVA

A variância relativa é uma medida de dispersão relativa que resulta do quociente entre a variância absoluta e o quadrado da média. É basicamente o quadrado do coeficiente de variação. Isto é:

a) para populações

$$VR = \left(\frac{\sigma}{\mu}\right)^2 = \frac{\sigma^2}{\mu^2}$$

b) para amostras

$$VR = \left(\frac{s}{\bar{x}}\right)^2 = \frac{s^2}{\bar{x}^2}$$

A variância relativa, assim como o coeficiente de variação, é uma medida adimensional, ou seja, não tem uma unidade de medida. Repare que tanto o numerador (variância) quanto o denominador (quadrado da média) são expressos na mesma unidade de medida, de modo a se cancelarem no momento da divisão.

EXEMPLIFICANDO

Em uma empresa de tecnologia, o salário médio dos homens é de R$ 1800,00 com desvio-padrão de R$ 810,00 e o salário médio das mulheres é de R$ 1500,00 com desvio padrão de R$ 705,00. A **variância relativa** dos salários dos homens é maior que a das mulheres?

Vamos identificar os dados do problema:

a) para os homens:

$$\begin{cases} \mu_H = 1800 \\ \sigma_H = 810 \end{cases}$$

b) para as mulheres:

$$\begin{cases} \mu_M = 1500 \\ \sigma_M = 705 \end{cases}$$

Agora, vamos calcular as respectivas variâncias relativas:

a) para os homens:

$$VR = \left(\frac{\sigma_H}{\mu_H}\right)^2 = \left(\frac{810}{1800}\right)^2 \cong 0{,}20$$

b) para as mulheres:

$$VR = \left(\frac{\sigma_M}{\mu_M}\right)^2 = \left(\frac{705}{1500}\right)^2 \cong 0{,}22$$

Portanto, os salários das mulheres apresentam uma variância relativa maior que os salários dos homens.

HORA DE PRATICAR!

(FCC/SEFAZ-BA/2019) O coeficiente de variação de Pearson correspondente a uma população P1 com média aritmética igual a 20 e tamanho 20 é igual a 30%. Decide-se excluir de P1, em um determinado momento, dois elementos iguais a 11 cada um, formando uma nova população P2. A variância relativa de P2 é igual a

a) 10/147.

b) 4/49.

c) 16/147.

d) 8/49.

e) 4/441.

Comentários:

O coeficiente de variação de Pearson é a razão entre o desvio padrão e a média.

$$CV_{P_1} = \frac{\sigma_1}{\overline{X}_1}$$

$$0{,}3 = \frac{\sigma_1}{20}$$

$$\sigma_1 = 20 \times 0{,}3 = 6$$

Logo, a variância de P_1 é:

$$\sigma_1^2 = 6^2 = 36$$

Como a variância é a média dos quadrados menos o quadrado das médias, temos:

$$\sigma_1^2 = \overline{X_1^2} - (\overline{X_1})^2$$

$$36 = \overline{X_1^2} - 20^2$$

$$36 = \overline{X_1^2} - 400$$

$$\overline{X_1^2} = 436$$

Com isso, podemos calcular a soma dos termos:

$$\overline{X_1} = \frac{\sum X_i}{20}$$

$$\overline{X_1} = \frac{\sum_{i=1}^{20} X_i}{20}$$

$$\sum_{i=1}^{20} X_i = 20 \times \overline{X_1} = 20 \times 20 = 400$$

De igual forma, temos:

$$\overline{X_1^2} = \frac{\sum_{i=1}^{20} X_i^2}{20}$$

$$\sum_{i=1}^{20} X_i^2 = 20 \times \overline{X_1^2} = 20 \times 436 = 8.720$$

O enunciado afirma que dois elementos iguais a 11 serão retirados, formando uma nova população P_2. Dessa forma, as novas somas serão iguais a:

$$\sum_{i=1}^{18} X_i = 400 - 2 \times 11 = 378$$

$$\sum_{i=1}^{18} X_i^2 = 8720 - 2 \times 11^2 = 8.478$$

Assim, as novas médias são iguais a:

$$\overline{X_2} = \frac{\sum_{i=1}^{18} X_i}{18} = \frac{378}{18} = 21$$

$$\overline{X_2^2} = \frac{\sum_{i=1}^{18} X_i^2}{18} = \frac{8.478}{18} = 471$$

De posse dessas informações, podemos calcular a nova variância absoluta:

$$\sigma_2^2 = \overline{X_2^2} - (\overline{X_2})^2$$

$$\sigma_2^2 = 471 - (21)^2$$

$$\sigma_2^2 = 471 - 441$$

$$\sigma_2^2 = 30$$

Finalmente, temos que a variância relativa é a razão entre a variância e o quadrado da média:

$$VR_{P_2} = \frac{\sigma_2^2}{(\overline{X_2})^2} = \frac{30}{441}$$

Simplificando por 3, temos:

$$VR_{P_2} = \frac{30}{441} = \frac{10}{147}$$

Gabarito: A.

Questões Comentadas

Medidas de Dispersão

1. (FCC/TJ SC/2021) O número de processamento de documentos da Divisão de Pesquisa e Informação do Tribunal de Justiça requer uma análise refinada de variáveis quantitativas. Dentre os procedimentos analíticos, indicam-se as medidas de dispersão, representadas por

a) Média harmônica e Média geométrica.

b) Amplitude total e Desvio médio absoluto.

c) Moda e Mediana.

d) Média aritmética simples e Média aritmética ponderada.

e) Mediana para dados agrupados em classes e Mediana para dados não agrupados.

Comentários:

As medidas de dispersão (ou variabilidade) são justamente métricas que mostram a variação dos dados de um conjunto. Elas podem ser divididas em dois grupos:

a) medidas de dispersão absoluta:
   - amplitude total;
   - amplitude interquartílica;
   - desvio médio;
   - variância; e
   - desvio-padrão.

b) medidas de dispersão relativa:
   - coeficiente de variação (de Pearson); e
   - variância relativa.

Gabarito: B.

2. (VUNESP/MPE-SP/2016) 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.

Comentários:

As medidas de tendência central estudam o centro da amostra. São dadas pela média, mediana e moda.

As medidas de separatrizes separam os dados em grupos com a mesma quantidade. São os quartis, decis e percentis.

As medidas de dispersão têm a finalidade de identificar o quanto os dados estão dispersos em torno da média de uma amostra. São coeficientes de variação, desvio padrão, amplitude e variância.

Gabarito: D.

QUESTÕES COMENTADAS

Amplitude Total

1. (VUNESP/IPSM SJC/2018) Considere as informações a seguir para construir uma distribuição de frequência sem intervalo de classe e responder a questão.

Um dado foi lançado 50 vezes e foram registrados os seguintes resultados:

5 4 6 1 2 5 3 1 3 3

4 4 1 5 5 6 1 2 5 1

3 4 5 1 1 6 6 2 1 1

4 4 4 3 4 3 2 2 2 3

6 6 3 2 4 2 6 6 2 1

A amplitude total é

a) 50.

b) 6.

c) 7.

d) 10.

e) 5.

Comentários:

A amplitude é calculada pela diferença entre os valores máximos e mínimos da amostra. Observando os dados apresentados, temos:

$$Valor\ máximo\ =\ 6$$

$$Valor\ mínimo\ =\ 1$$

$$6-1=5$$

Gabarito: E.

2. (CESPE/SEDF/2017) Um levantamento estatístico, feito em determinada região do país, mostrou que jovens com idades entre 4 e 17 anos assistem à televisão, em média, durante 6 horas por dia. A tabela a seguir apresenta outras estatísticas produzidas por esse levantamento.

	Distribuição dos tempos gastos assistindo televisão (*T, em horas*)
1° quartil	2
2° quartil	4
3° quartil	8
1° decil	1
9° decil	10

Tendo como referência essas informações, julgue o seguinte item.

A amplitude total dos tempos T é igual ou superior a 9 horas.

Comentários:

A amplitude é calculada pela diferença entre os valores máximos e mínimos da amostra. Os valores do 9.º decil e do 1.º decil foram informados na questão. Assim, para resolvermos a questão, basta fazermos a subtração dos valores correspondentes ao 9.º decil e 1.º decil:

$$10 - 1 = 9$$

Com isso, sabemos que a amplitude é de, no mínimo, 9.

Gabarito: Certo.

3. (CESGRANRIO/IBGE/2016) Suponha que, em uma pesquisa on-line sobre as idades dos habitantes de um condomínio, um respondente de 30 anos digite erroneamente sua idade como sendo 300 anos. Considere que esse erro passe despercebido e que não haja outros erros na base de dados.

Nessas condições, a única conclusão que NÃO pode ser formulada é:

a) A média de idades calculada a partir dos dados da base será maior do que a média de idades reais dos respondentes.

b) A mediana de idades calculada a partir dos dados da base será maior do que a mediana de idades reais dos respondentes.

c) A amplitude de idades calculada a partir dos dados da base será maior do que a amplitude de idades reais dos respondentes.

d) O valor máximo das idades calculado a partir dos dados da base será maior do que a idade real do respondente mais velho.

e) A diferença entre as duas maiores idades dos dados da base será maior do que a diferença das idades reais dos dois respondentes mais velhos.

Comentários:

Analisando as alternativas, temos:

Alternativa A: **Correta.** A média é dada pela soma de todos os termos da amostra dividida pelo número de observações. Sendo assim, se o valor no somatório é aumentado, consequentemente, a média também sofre um aumento.

Alternativa B: **Errada.** A mediana corresponde ao termo central da amostra. A depender de quais os termos centrais (ou o termo central) da amostra, a mediana pode mudar ou permanecer a mesma, caso um termo seja aumentado.

Alternativa C: **Correta.** A amplitude é calculada pela diferença entre o maior e o menor valor da amostra.

Alternativa D: **Correta.** O valor máximo das idades será o maior valor informado pelos respondentes.

Alternativa E: **Correta.** Se a maior idade for aumentada, a diferença entre as duas idades também será aumentada.

Gabarito: B.

4. (CESGRANRIO/IBGE/2016) Uma pesquisa em determinado município coletou, dentre outros dados, o número de filhos em cada família. Algumas estatísticas são apresentadas na Tabela abaixo.

Número de filhos	
Média	2
Mediana	1
Moda	0
Desvio-padrão	3
Amplitude	5

Segundo essas estatísticas,

a) metade das famílias tem mais do que 2 filhos.

b) o mais comum é que famílias tenham 2 filhos.

c) mais da metade das famílias não têm filhos.

d) uma família padrão tem em média 3 filhos.

e) de todas as famílias entrevistadas, nenhuma tem 6 filhos.

Comentários:

Vamos analisar cada uma das alternativas:

Alternativa A: **Errada**. Na tabela, temos que a mediana vale 1. Assim, sabemos que metade das famílias tem mais do que 1 filho.

Alternativa B: **Errada**. Na tabela, a moda ou número mais comum vale 0.

Alternativa C: **Errada**. Não podemos afirmar que mais da metade das famílias não têm filhos, pois a mediana do número de filhos vale 1.

Alternativa D: **Errada**. Na tabela, temos que o desvio padrão vale 3. O desvio padrão não pode ser confundido com o conceito de família padrão, que está mais relacionado à moda.

Alternativa E: **Correta**. Temos que a amplitude vale 5, e que o menor número de filhos é 0. Assim, podemos calcular a amplitude:

$$AT = X_{máximo} - X_{mínimo}$$

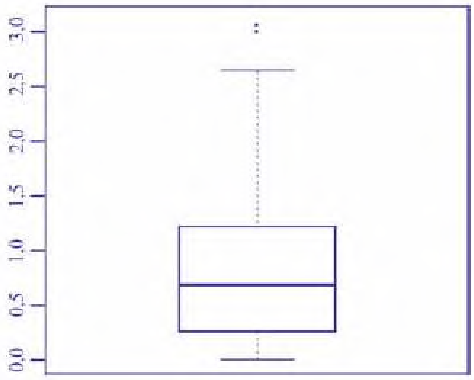
$$5 = X_{máximo} - 0$$

$$X_{máximo} = 5$$

Logo, nenhuma das famílias têm 6 filhos.

Gabarito: E.

5. (CESPE/TCE-PA/2016)



média amostral	0,80
desvio padrão amostral	0,70
primeiro quartil	0,25
mediana	0,70
terceiro quartil	1,20
mínimo	0
máximo	3,10

Um indicador de desempenho X permite avaliar a qualidade dos processos de governança de instituições públicas. A figura mostra, esquematicamente, a sua distribuição, obtida mediante estudo amostral feito por determinada agência de pesquisa. A tabela apresenta estatísticas descritivas referentes a essa distribuição.

Com base nessas informações, julgue o item a seguir.

A amplitude total da amostra é inferior a 3.

Comentários:

A amplitude (ou amplitude total) é a diferença entre o valor máximo e o mínimo. Esses valores foram apresentados na tabela do enunciado. Então, aplicando esses dados na fórmula, temos:

$$A = 3{,}10 - 0 = 3{,}10$$

Gabarito: Certo.

6. (CESPE/TCE-PA/2016)

Número diário de denúncias registradas (X)	Frequência Relativa
0	0,3
1	0,1
2	0,2
3	0,1
4	0,3
Total	**1,0**

A tabela precedente apresenta a distribuição de frequências relativas da variável X, que representa o número diário de denúncias registradas na ouvidoria de determinada instituição pública. A partir das informações dessa tabela, julgue o item seguinte.

A amplitude total da amostra é igual ou superior a 5.

Comentários:

A amplitude (ou amplitude total) é a diferença entre o valor máximo e o mínimo. Esses valores foram apresentados na tabela do enunciado. Então, aplicando esses dados na fórmula, temos:

$$A = 4 - 0 = 4$$

Gabarito: Errado.

7. (FGV/DPE-RJ/2014) Dentre as informações coletadas dos cidadãos através do 1º atendimento da Defensoria Pública estão as variáveis idade, renda e o número de dependentes. Cada uma é classificada em três diferentes níveis A, B e C, com valores de referência conforme a tabela:

Variáveis	A	B	C	D
Renda Média (R$ 100)	9	11	11	17
Idade Média (anos)	20	32	36	48
Dependentes (pessoas)	2	3	3	2

Portanto, as unidades de medida são distintas (R$, anos e pessoas). Mesmo assim, através de uma estatística de amplitude, escolhida convenientemente, aqui representada por VB, é possível comparar as dispersões. Logo, renda, idade e número de dependentes seguem a ordenação

a) VB (renda) < VB (idade) < VB (dependentes).

b) VB (renda) < VB (dependentes) < VB (idade).

c) VB (idade) < VB (dependentes) < VB (renda).

d) VB (idade) < VB (renda) < VB (dependentes).

e) VB (dependentes) < VB (renda) < VB (idade).

Comentários:

A amplitude é a diferença entre o valor máximo e o mínimo. Tomando as informações da tabela, temos:

- amplitude para a renda: $17 - 9 = 8$;
- amplitude para a idade: $48 - 20 = 28$; e
- amplitude para os dependentes: $3 - 2 = 1$.

Logo:

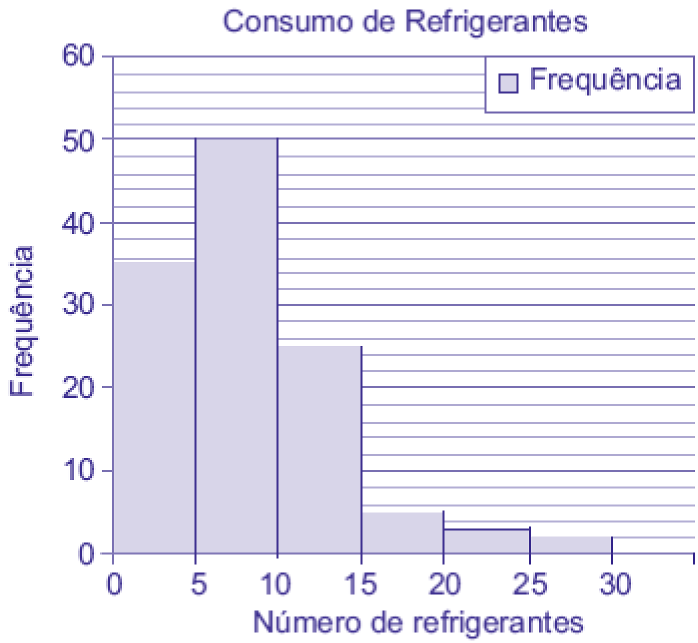
$$dependentes < renda < idade$$

Gabarito: E.

QUESTÕES COMENTADAS

Amplitude Interquartílica

1. (CESGRANRIO/BB/2018) Uma escola de Ensino Médio decide pesquisar o comportamento de seus estudantes quanto ao número de refrigerantes consumidos semanalmente por eles. Para isso, uma amostra aleatória de 120 estudantes foi selecionada, e os dados foram sintetizados no histograma abaixo, em classes do tipo [0, 5), [5, 10), [10, 15), [15, 20), [20, 25) e [25, 30].



Qual o valor da amplitude interquartílica, obtido por meio do método de interpolação linear dos dados agrupados em classes?

a) 15

b) $\frac{15}{2}$

c) $\frac{29}{5}$

d) $\frac{47}{7}$

e) 10

Comentários:

Inicialmente, vamos colocar as informações do gráfico em uma tabela para melhor visualizarmos as informações:

Classes	Nº de refrigerantes	Nº acumulado
$\mathbf{0 \vdash 5}$	35	35
$\mathbf{5 \vdash 10}$	50	85
$\mathbf{10 \vdash 15}$	25	110
$\mathbf{15 \vdash 20}$	5	115
$\mathbf{20 \vdash 25}$	3	118
$\mathbf{25 \vdash 30}$	2	120

Sabemos que cada quartil corresponde a 25% da distribuição acumulada, portanto Q_1 tem até 25% das observações e Q_3 tem 75%.

Se temos um total de 120 observações, então temos que Q_1 ocupa o número acumulado de 30 e Q_3 de 90.

$$Q_1 = \frac{120}{4} = 30$$

$$Q_3 = \frac{3 \times 120}{4} = 90$$

Observando a tabela temos que Q_1está na classe [0,5) e Q_3 está na classe [10,15).

Vamos usar o método da interpolação linear para determinar o valor de cada quartil:

$$Q_k = l_{inf_{Q_k}} + \left[\frac{\frac{k \times \sum f_i}{4} - f_{ac_{ant}}}{f_{Q_k}} \right] \times h_{Q_k}$$

em que:

$l_{inf_{Q_k}}$ = limite inferior da classe do quartil considerado;

$f_{ac_{ant}}$ = frequência acumulada da classe anterior à classe do quartil considerado;

h_{Q_k} = amplitude do intervalo de classe do quartil considerado;

f_{Q_k} = frequência simples da classe do quartil considerado.

Aplicando a fórmula acima para o primeiro quartill:

$$Q_1 = l_{inf_{Q_1}} + \left[\frac{\frac{1 \times \sum f_i}{4} - f_{ac_{ant}}}{f_{Q_1}}\right] \times h_{Q_1}$$

$$Q_1 = 0 + \left[\frac{\left(\frac{1 \times 120}{4}\right) - 0}{35}\right] \times (5 - 0)$$

$$Q_1 = \left[\frac{30 - 0}{35}\right] \times 5$$

$$Q_1 = \frac{150}{35}$$

$$Q_1 = \frac{30}{7}$$

Fazendo o mesmo para o terceiro quartil:

$$Q_3 = l_{inf_{Q_3}} + \left[\frac{\frac{3 \times \sum f_i}{4} - f_{ac_{ant}}}{f_{Q_1}}\right] \times h_{Q_3}$$

$$Q_3 = 10 + \left[\frac{\left(\frac{3 \times 120}{4}\right) - 85}{25}\right] \times (15 - 10)$$

$$Q_3 = 10 + \left[\frac{90 - 85}{25}\right] \times 5$$

$$Q_3 = 10 + \left(\frac{25}{25}\right)$$

$$Q_3 = 10 + 1$$

$$Q_3 = 11$$

Já temos os valores de Q_3 e Q_1, agora é só fazermos a diferença para sabermos a amplitude interquartílica:

$$Q_3 - Q_1 = 11 - \frac{30}{7} = \frac{77 - 30}{7} = \frac{47}{7}$$

Gabarito: D.

2. (CESGRANRIO/PETROBRAS/2018) Você dispõe de um montante para investir em ações e precisa decidir em que empresa(s) vai alocar esse montante. Três empresas lhe parecem interessantes, e você resolve consultar o desempenho delas nos últimos sessenta meses para minimizar possíveis riscos da sazonalidade no movimento da Bolsa de Valores. Os dados revelaram a seguinte distribuição, em %, das rentabilidades mensais das ações:

Medidas Estatísticas	Empresa A	Empresa B	Empresa C

Rentabilidade média mensal	0,50	0,60	0,40
Desvio padrão	1,00	1,20	0,80
Rentabilidade mínima	-1,80	-2,20	-1,20
Rentabilidade máxima	2,20	2,30	1,80
1º quartil	-0,20	-0,30	-0,10
3º quartil	0,80	0,90	0,70

A alocação dos recursos vai ser feita de acordo com a atitude conservadora de não investir em empresa com rentabilidade considerada outlier, entendendo como tal aquela que apresentar valor além de 1,5 desvio quartílico abaixo ou acima dos quartis 1 e 3.

Com base nesse critério, a escolha do investimento deve recair sobre a(s)

a) empresa A, apenas

b) empresa B, apenas

c) empresa C, apenas

d) empresas A e C, apenas

e) três empresas

Comentários:

Geralmente, os limites para consideração de valores outliers se baseiam na amplitude quartílica, que é somente a diferença entre o terceiro e o primeiro quartil. Assim, temos que a amplitude quartílica é dada por:

$$DIQ = Q_3 - Q_1$$

Vamos calcular a amplitude para cada empresa:

$$DIQ_A = Q_3 - Q_1 = 0{,}8 - (-0{,}2) = 1$$

$$DIQ_B = Q_3 - Q_1 = 0{,}9 - (-0{,}3) = 1{,}2$$

$$DIQ_C = Q_3 - Q_1 = 0{,}7 - (-0{,}1) = 0{,}8$$

Agora, vamos multiplicar a amplitude quartílica de cada empresa por 1,5, conforme enunciado:

$$\Delta_A = 1 \times 1{,}5 = 1{,}5$$

$$\Delta_B = 1{,}2 \times 1{,}5 = 1{,}8$$

$$\Delta_C = 0{,}8 \times 1{,}5 = 1{,}2$$

Pronto, já podemos verificar as rentabilidades mínima e máxima para cada empresa:

Medidas Estatísticas	Empresa A	Empresa B	Empresa C
Rentabilidade mínima	$-0{,}2 - 1{,}5 = -1{,}7$	$-0{,}3 - 1{,}8 = -2{,}1$	$-0{,}1 - 1{,}2 = -1{,}3$
Rentabilidade máxima	$0{,}8 + 1{,}5 = 2{,}3$	$0{,}9 + 1{,}8 = 2{,}7$	$0{,}7 + 1{,}2 = 1{,}9$

Portanto, somente a empresa C estaria apta a receber os investimentos.

Gabarito: C.

3. (CESPE/SEE-DF/2017) Um levantamento estatístico, feito em determinada região do país, mostrou que jovens com idades entre 4 e 17 anos assistem à televisão, em média, durante 6 horas por dia. A tabela a seguir apresenta outras estatísticas produzidas por esse levantamento.

	Distribuição dos tempos gastos assistindo televisão (*T, em horas*)
1° quartil	2
2° quartil	4
3° quartil	8
1° decil	1
9° decil	10

Tendo como referência essas informações, julgue o seguinte item.

O desvio quartílico dos tempos T foi igual a 3.

Comentários:

O desvio quartílico é dado por $\frac{Q_3 - Q_1}{2}$, em que Q_3 e Q_1 são o 3º e o 1º quartis, respectivamente. Aplicando os dados da tabela na fórmula, temos:

$$\frac{Q_3 - Q_1}{2} = \frac{8 - 2}{2} = 3$$

Gabarito: Certo.

4. (FGV/IBGE/2017) A população de um estudo é dividida em quatro estratos, sendo o menor com 10% dos indivíduos e os demais com tamanhos acrescidos de dez pontos percentuais, progressivamente. Os estratos se distinguem por classes de renda com amplitude constante, sendo maiores quanto menor a renda. Sobre os estratos sabe-se que:

$$Rd_{Estrato1} = 65\ Rd_{Estrato2} = \ 45\ e\ Rd_{Estrato4} = \ 5$$

Onde os valores acima representam os limites inferiores da renda dos extratos, inclusive.

Portanto, é correto afirmar que:

a) Tomando os pontos médios das classes como representativos, a renda média é igual a Md(Rd) = 38;

b) A mediana da distribuição de renda, Me(Rd), é menor que 45 e maior do que ou igual a 25;

c) Tomando os pontos médios das classes como representativos, a moda da renda é igual a Mo(Rd) = 35;

d) O valor máximo atingido pela renda nessa distribuição é igual a Mx(Rd) = 85;

e) O valor do desvio-interquartílico da distribuição de renda deverá ser superior a 50.

Comentários:

O enunciado informou que a amplitude é igual para todas as classes. Tomando a informação da segunda classe, 65 a 45, concluímos que a amplitude será igual a 20. Assim, temos:

Estrato	Renda	Frequência	Frequência acumulada
4	85⊣65	10%	10%
3	**65⊣45**	**20%**	**30%**
2	45⊣25	30%	60%
1	25⊣5	40%	100%

Vamos analisar cada uma das alternativas:

- letra A: tomando os pontos médios de cada classe, calculamos a média:

  $$\bar{x} = (75 \times 0{,}1 \ + \ 55 \times 0{,}2 \ + \ 35 \times 0{,}3 \ + \ 15 \times 0{,}4) = 35$$

  Logo, a alternativa está incorreta;

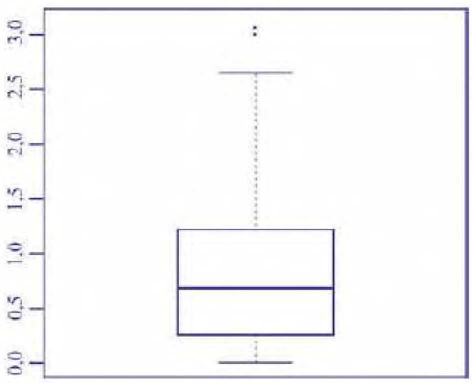
- letra B: a mediana é o valor que ocupa o termo central da amostra, representando 50% do total. Pela frequência acumulada, percebemos que 50% correspondem ao estrato 2, cuja classe é 25 a 45. Portanto, a alternativa está correta;

- letra C: a moda é o termo que mais se repete na amostra. Assim, quem mais se repete representa 40% da frequência relativa. Logo será 15, ponto médio da classe 5 a 25. Alternativa incorreta;

- letra D: a renda máxima é o valor que tende a 85, mas não necessariamente será esse valor exato. Alternativa errada
- letra E: primeiro quartil Q_1 pode assumir qualquer valor entre 5 e 25, e o terceiro quartil Q_3 pode assumir qualquer valor entre 45 e 65. Se pegarmos os pontos médios de cada classe, o desvio interquartílico fica: $Q_3 - Q_1 = 55 - 15 = 40$. Portanto, alternativa incorreta.

Gabarito: B.

5. (CESPE/TCE-PA/2016)



média amostral	0,80
desvio padrão amostral	0,70
primeiro quartil	0,25
mediana	0,70
terceiro quartil	1,20
mínimo	0
máximo	3,10

Um indicador de desempenho X permite avaliar a qualidade dos processos de governança de instituições públicas. A figura mostra, esquematicamente, a sua distribuição, obtida mediante estudo amostral feito por determinada agência de pesquisa. A tabela apresenta estatísticas descritivas referentes a essa distribuição.

Com base nessas informações, julgue o item a seguir.

O intervalo interquartílico da distribuição do indicador X é superior a 1,4.

Comentários:

O intervalo quartílico é dado pela distância entre o terceiro (Q_3) e o primeiro quartil (Q_1), isto é:

$$Q_3 - Q_1 = 1{,}20 - 0{,}25 = 0{,}95$$

Gabarito: Errado.

6. (CESGRANRIO/EPE/2014) A Tabela a seguir apresenta a vazão média em cada mês para um determinado rio.

Mês	Vazão média mensal (m^3/s)
Janeiro	97

Fevereiro	60
Março	50
Abril	60
Maio	70
Junho	85
Julho	60
Agosto	50
Setembro	68
Outubro	117
Novembro	80
Dezembro	43

De acordo com os dados da Tabela, a mediana e a amplitude interquartílica das vazões valem, respectivamente,

a) 70 e 25

b) 64 e 25

C) 64 e 27,5

d) 68 e 27,5

e) 70 e 27,5

Comentários:

Inicialmente, vamos organizar os dados da tabela em ordem crescente:

$$\underbrace{43 \quad 50 \quad 50} \quad \underbrace{60 \quad 60 \quad 60} \quad \underbrace{68 \quad 70 \quad 80} \quad \underbrace{85 \quad 97 \quad 117}$$

Sabemos que cada quartil corresponde a 25% da distribuição, portanto até Q1 tem 25% das observações e Q3 tem 75%.

Para o primeiro quartil, temos que será a média entre 50 e 60:

$$Q_1 = \frac{50 + 60}{2} = 55$$

Para o segundo quartil, temos que é a mesma mediana da amostra e será a média entre 60 e 68:

$$Q_2 = \frac{60 + 68}{2} = 64$$

Para o terceiro quartil, temos que será a média entre 80 e 85:

$$Q_3 = \frac{80 + 85}{2} = 82{,}5$$

Portanto, já sabemos que a mediana vale 64. Calculando a amplitude:

$$Q_3 - Q_1 = 82{,}5 - 55 = 27{,}5$$

Gabarito: C.

7. (FCC/TCE-PR/2011) Atenção: Considere as informações a seguir para responder à questão.

A distribuição dos salários dos 1000 funcionários da companhia A, em número de salários mínimos, está apresentada na tabela abaixo:

Faixa salarial (em número de salários mínimos)	Frequência absoluta
$1 \vdash 3$	200
$3 \vdash 5$	400
$5 \vdash 7$	200
$7 \vdash 9$	200

A distância interquartil desses salários, definida por $Q_3 - Q_1$, onde Q_3 e Q_1 são, respectivamente, os quartis de ordem 3 e 1, calculados pelo método da interpolação linear, em número de salários mínimos, é

a) 2,75.

b) 3,00.

c) 3,25.

d) 3,50.

e) 4,00.

Comentários:

Sabemos que cada quartil corresponde a 25% da distribuição acumulada, portanto, Q_1 tem até 25% das observações e Q_3 tem 75%. Para resolvermos a questão, precisamos calcular as frequências acumuladas:

Classes	Freq. simples	Freq. acumuladas
1 a 3	200	$= 200$
3 a 5	400	$200 + 400 = 600$
5 a 7	200	$600 + 200 = 800$
7 a 9	200	$800 + 200 = 1.000$

Se temos um total de 1000 observações, então temos que Q_1 ocupa a frequência acumulada de posição 250 e Q_3 ocupa a posição 750. Vamos usar o método da interpolação linear para determinar o valor de cada quartil:

$$\frac{Q_1 - 3}{5 - 3} = \frac{250 - 200}{600 - 200}$$

$$\frac{Q_1 - 3}{2} = \frac{50}{400}$$

$$Q_1 - 3 = \frac{100}{400}$$

$$Q_1 - 3 = 0{,}24$$

$$Q_1 = 3{,}24$$

Fazendo o mesmo para o terceiro quartil:

$$\frac{Q_3 - 5}{7 - 5} = \frac{750 - 600}{800 - 600}$$

$$\frac{Q_3 - 5}{2} = \frac{150}{200}$$

$$\frac{Q_3 - 5}{2} = 0{,}75$$

$$Q_3 - 5 = 2 \times 0{,}75 = 1{,}5$$

$$Q_3 = 6{,}5$$

Já temos os valores de Q_3 e Q_1, agora, basta encontrarmos a diferença para sabermos o intervalo interquartílico:

$$Q_3 - Q_1 = 6{,}5 - 3{,}25 = 3{,}25$$

Gabarito: C.

8. (FCC/BACEN/2006) Considere a distribuição de frequências a seguir para resolver a questão.

Salários dos empregados das empresas XYZ em dezembro de 2005

Salários	Frequências simples absolutas
$1.000{,}00 \vdash 2.000{,}00$	2
$2.000{,}00 \vdash 3.000{,}00$	8
$3.000{,}00 \vdash 4.000{,}00$	16
$4.000{,}00 \vdash 5.000{,}00$	10
$5.000{,}00 \vdash 6.000{,}00$	4

A amplitude do intervalo entre o primeiro decil e o terceiro quartil, encontrados pelo método da interpolação linear, é

a) R$ 2 500,00

b) R$ 2 400,00

c) R$ 2 150,00

d) R$ 2 000,00

e) R$ 1 400,00

Comentários:

Sabemos que cada quartil corresponde a 25% da distribuição acumulada, portanto, Q_1 tem até 25% das observações e Q_3 tem 75%. Para resolvermos a questão, precisamos calcular as frequências acumuladas:

Classes	Freq. simples	Freq. acumuladas
1.000 - 2.000	2	$= 2$
2.000 - 3.000	8	$2 + 8 = 10$
3.000 - 4.000	16	$10 + 16 = 26$
4.000 - 5.000	10	$26 + 10 = 36$
5.000 - 6.000	4	$36 + 4 = 40$

Se temos um total de 40 observações, então temos que Q_3 ocupa a frequência acumulada de posição 30, que corresponde a 75% de 40; e o primeiro decil, D_1, corresponderá à frequência acumulada de posição 4, que corresponde a 10% de 40. Usando o método da interpolação linear para determinar o valor de cada quartil, temos:

$$\frac{Q_3 - 4.000}{5.000 - 4.000} = \frac{30 - 26}{36 - 26}$$

$$\frac{Q_3 - 4000}{1000} = \frac{4}{10}$$

$$Q_3 - 4000 = \frac{4000}{10}$$

$$Q_3 - 4000 = 400$$

$$Q_3 = 4.400$$

Fazendo o mesmo para o primeiro decil:

$$\frac{D_1 - 2.000}{3.000 - 2.000} = \frac{4 - 2}{10 - 2}$$

$$\frac{D_1 - 2.000}{1.000} = \frac{2}{8}$$

$$D_1 - 2.000 = \frac{2000}{8}$$

$$D_1 - 2.000 = 250$$

$$D_1 = 2.250$$

Já temos os valores de Q_3 e D_1. Agora, basta encontrarmos a diferença para sabermos a amplitude entre o terceiro quartil e o primeiro decil:

$$Q_3 - D_1 = 4.400 - 2.250 = 2.150$$

Gabarito: C.

Questões Comentadas

Desvios em Relação à Média Aritmética e Mediana

1. (VUNESP/Pref. Campinas/2019) Considere a tabela-1 e o enunciado seguintes para responder à questão.

x_i	f_i	$x_i \times f_i$	$(x_i - \overline{x})^2$
1	2	2	A
3	4	12	4
5	6	30	B
7	4	28	4
9	2	18	C
Totais	18	90	40

Tabela-1

A tabela-1 de distribuição de frequência mostra a organização e síntese de 18 dados x_i colhidos como amostra para um estudo estatístico, onde a coluna f_i é a que registra os valores das frequências, enquanto a coluna $(x_i - \overline{x})^2$ contém os valores dos quadrados dos desvios.

Os valores substituídos pelas letras A, B e C na tabela são, respectivamente:

a) 0, 16, 0.

b) 4, 4, 4.

c) 16, 0, 16.

d) 0, 4, 0.

e) 16, 4, 16.

Comentários:

Vamos inicialmente calcular a média. A tabela já fornece os totais dos dados:

$$\bar{x} = \frac{90}{18} = 5$$

Agora, basta aplicarmos a fórmula já dada na tabela para encontrar os valores de A, B e C: $(x_i - \bar{x})^2$.

$$A = (1 - 5)^2 = 16$$

$$B = (5 - 5)^2 = 0$$

$$C = (9 - 5)^2 = 16$$

Gabarito: C.

2. (VUNESP/TJ SP/2015) Leia o texto a seguir para responder à questão.

Uma pequena empresa que emprega apenas cinco funcionários paga os seguintes salários mensais (em mil reais):

0,9	1,2	1,4	1,5	2,0

Considerando-se a média dos salários, o valor do desvio do salário de quem ganha R$ 1.400,00 mensais é

a) –1.000.

b) –400.

c) 0.

d) 200.

e) 400.

Comentários:

Vamos inicialmente calcular a média:

$$\bar{x} = \frac{0{,}9 + 1{,}2 + 1{,}4 + 1{,}5 + 2}{5}$$

$$\bar{x} = \frac{7}{5} = 1{,}4\, mil$$

O desvio do salário é de 1.400. Então:

$$\sigma = x_i - \bar{x}$$

$$1400 - 1400 = 0$$

Gabarito: C.

3. (VUNESP/TJ SP/2015) Leia o texto a seguir para responder à questão.

Uma pequena empresa que emprega apenas cinco funcionários paga os seguintes salários mensais (em mil reais):

0,9	1,2	1,4	1,5	2,0

Somando-se os valores absolutos dos desvios individuais dos salários tomados em relação à média, encontra-se o valor de

a) 1.400,00.

b) 1.200,00.

c) 1.000,00.

d) 800,00.

e) 0.

Comentários:

Conforme já calculado no exercício anterior, a média é:

$$\bar{x} = \frac{0,9 + 1,2 + 1,4 + 1,5 + 2}{5}$$

$$\bar{x} = \frac{7}{5} \Longrightarrow \bar{x} = 1,4\ mil$$

Agora, calculamos os desvios para cada valor apresentado:

Valor	Desvio	Desvio absoluto
$0,9$	$0,9 - 1,4 = -0,50$	$0,5$
$1,2$	$1,2 - 1,4 = -0,2$	$0,2$
$1,4$	$1,4 - 1,4 = 0$	0
$1,5$	$1,5 - 1,4 = 0,1$	$0,1$
2	$2 - 1,4 = 0,6$	$0,6$
Total		**1,4**

Assim, a soma dos desvios absolutos é 1,4 mil.

Gabarito: A.

4. (FGV/SEFAZ MS/2006) Analise as afirmativas a seguir, a respeito da média aritmética:

I. A soma dos resíduos em relação à média aritmética é sempre igual a zero.

II. É em relação à média aritmética que a soma dos valores absolutos dos resíduos é mínima.

III. É em relação à média aritmética que a soma dos quadrados dos resíduos é mínima.

Assinale:

a) se somente a afirmativa II estiver correta.

b) se somente as afirmativas I e II estiverem corretas.

c) se somente as afirmativas I e III estiverem corretas.

d) se somente as afirmativas II e III estiverem corretas.

e) se todas as afirmativas estiverem corretas.

Comentários:

Vamos analisar cada item:

Item I - **Certo**. Desenvolvendo a propriedade da média:

$$\sum_{i=1}^{n}(x_i - \bar{x}) =$$

$$\sum_{i=1}^{n}(x_i) - n \times \bar{x} =$$

Temos que $\bar{x} = \frac{\sum x_i}{n}$ logo:

$$= \sum_{i=1}^{n}(x_i) - n \times \frac{(\sum x_i)}{n}$$

$$= \sum_{i=1}^{n}(x_i) - \left(\sum x_i\right)$$

$$= 0$$

Item II - **Errado**. Essa propriedade é aplicada à mediana e não à média.

Item III - **Certo**. Se calculamos os desvios em relação a um valor "y" qualquer, a soma dos quadrados dos desvios fica:

$$\sum_{i=1}^{n}(x_i - \mathrm{y})^2 =$$

Minimizando a soma temos:

$$2 \times (-1) \times \sum_{i=1}^{n}(x_i - \mathrm{y}) = 0$$

$$\sum_{i=1}^{n}(x_i - \mathrm{y}) = 0$$

Assim temos que y é tal que a soma dos desvios é nula. Logo:

$$y = \bar{x}$$

Gabarito: C.

5. (FGV/SEFAZ-MS/2006) Analise as afirmativas a seguir, a respeito da mediana:

I. A soma dos resíduos em relação à mediana é sempre igual a zero.

II. É em relação à mediana que a soma dos valores absolutos dos resíduos é mínima.

III. É em relação à mediana que a soma dos quadrados dos resíduos é mínima.

Assinale:

a) se somente a afirmativa II estiver correta.

b) se somente as afirmativas I e II estiverem corretas.

c) se somente as afirmativas I e III estiverem corretas.

d) se somente as afirmativas II e III estiverem corretas.

e) se todas as afirmativas estiverem corretas.

Comentários:

Vamos analisar cada item:

Para I - **Errado**. Essa propriedade é aplicada à média e não à mediana.

Para II - **Certo**. Quando calculados em relação à mediana a soma dos módulos dos desvios de fato é mínima.

Para III - **Errado**. Essa propriedade é aplicada à média e não à mediana.

Gabarito: A.

6. (CESPE/ANATEL/2004)

meses	fev	mar	abr	mai	jun	jul	ago	set	out	nov
N	100	70	70	60	50	100	50	50	30	20

A tabela acima mostra os números mensais de reclamações (N) feitas por usuários de telefonia fixa, registradas em uma central de atendimento, entre os meses de fevereiro a novembro de 2003. Considerando esses dados, julgue o item que se segue.

O maior desvio absoluto dos números mensais de reclamações registradas é superior a 45.

Comentários:

Iniciaremos calculando a média das reclamações:

$$\bar{x} = \frac{100 + 70 + 70 + 60 + 50 + 100 + 50 + 50 + 30 + 20}{10}$$

$$\bar{x} = \frac{600}{10}$$

$$\bar{x} = 60$$

A partir daí, montamos uma tabela para calcular os desvios:

Reclamações (x_i)	Desvio em relação à média $(x_i - \bar{x})$
100	$100 - 60 = 40$
70	$70 - 60 = 10$
60	$60 - 60 = 0$
50	$50 - 60 = -10$
30	$30 - 60 = -30$
20	$20 - 60 = -40$

Portanto, temos que o maior desvio absoluto é 40.

Gabarito: Errado.

QUESTÕES COMENTADAS

Desvio Absoluto Médio

1. (FUNDATEC/SEPOG RS/2022) "O _______________ é uma boa medida de dispersão, porque dá a distância média de cada número em relação à média. Todavia, para muitos propósitos, é mais conveniente elevar ao quadrado cada _________ e tomar a média de todos esses quadrados. Essa grandeza é chamada _______________". (DOWNING; CLARK, 2011).

Assinale a alternativa que preenche, correta e respectivamente, as lacunas do trecho acima.

a) desvio médio absoluto – desvio – coeficiente de variação

b) desvio médio absoluto – desvio – variância

c) desvio médio absoluto – média – variância

d) desvio padrão – desvio – coeficiente de variação

e) desvio padrão – média – variância

Comentários:

O desvio absoluto médio, ou simplesmente desvio médio, mede a dispersão entre os valores da distribuição e a média dos dados coletados. Já o desvio padrão é definido como sendo a raiz quadrada da média aritmética dos quadrados dos desvios e, dessa forma, é determinado pela raiz quadrada da variância. Por sua vez, a variância é determinada pela média dos quadrados dos desvios em relação à média aritmética. Por fim, o coeficiente de variação é uma medida que fornece a variação dos dados em relação à média.

Preenchendo as lacunas, temos:

"O **desvio absoluto médio** é uma boa medida de dispersão, porque dá a distância média de cada número em relação à média. Todavia, para muitos propósitos, é mais conveniente elevar ao quadrado cada **desvio** e tomar a média de todos esses quadrados. Essa grandeza é chamada **variância**". (DOWNING; CLARK, 2011).

Gabarito: B.

2. (CESPE/ANATEL/2004)

meses	fev	mar	abr	mai	jun	jul	ago	set	out	nov
N	100	70	70	60	50	100	50	50	30	20

A tabela acima mostra os números mensais de reclamações (N) feitas por usuários de telefonia fixa, registradas em uma central de atendimento, entre os meses de fevereiro a novembro de 2003. Considerando esses dados, julgue o item que se segue.

O desvio médio absoluto da sequência formada pelos números mensais de reclamações é um valor entre 25 e 35

Comentários:

Iniciaremos calculando a média das reclamações:

$$\bar{x} = \frac{100 + 70 + 70 + 60 + 50 + 100 + 50 + 50 + 30 + 20}{10}$$

$$\bar{x} = \frac{600}{10}$$

$$\bar{x} = 60$$

A partir daí, montamos uma tabela para calcular os desvios:

Reclamações (x_i)	Frequências (f_i)	Desvio em relação à média ($x_i - \bar{x}$)
100	2	$100 - 60 = 40$
70	2	$70 - 60 = 10$
60	1	$60 - 60 = 0$
50	3	$50 - 60 = -10$
30	1	$30 - 60 = -30$
20	1	$20 - 60 = -40$

O desvio médio é dado pela média dos desvios absolutos, considerando a frequência:

$$\frac{2 \times 40 + 2 \times 10 + 1 \times 0 + 3 \times 10 + 1 \times 30 + 1 \times 40}{2 + 2 + 1 + 3 + 1 + 1}$$

$$\frac{200}{10} = 20$$

Gabarito: Errado.

Questões Comentadas

Variância

1. (CESPE/TCE SC/2022) Julgue o item a seguir, considerando conceitos de estatística.

Com os seguintes dados, a variância da população é de 149,25.

36; 64; 18; 40; 35; 30; 41; 32

Comentários:

A variância é determinada pela média dos quadrados dos desvios em relação à média aritmética. A variância populacional é dada por:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}$$

em que: x_i é o valor de ordem i assumido pela variável; μ é a média populacional de x; σ^2 é a variância populacional; e n é o número de dados da população.

Vamos iniciar calculando a média do conjunto:

$$\mu = \frac{36 + 64 + 18 + 40 + 35 + 30 + 41 + 32}{8} = \frac{296}{8} = 37$$

Agora, vamos calcular os desvios em relação à média. Para isso, vamos escrever os dados em uma tabela para melhor compreensão:

x_i	$x_i - \mu$	$(x_i - \mu)^2$
36	$36 - 37 = -1$	1
64	$64 - 37 = 27$	729
18	$18 - 37 = -19$	361
40	$40 - 37 = 3$	9
35	$35 - 37 = -2$	4
30	$30 - 37 = -7$	49
41	$41 - 37 = 4$	16

32	$32 - 37 = -5$	25
Total		**1194**

Calculando a variância, temos:

$$\sigma^2 = \frac{1194}{8} = 149{,}25$$

Portanto, a questão está correta.

Gabarito: Certo.

2. (CESPE/PETROBRAS/2022) O item a seguir é apresentada uma situação hipotética seguida de uma assertiva a ser julgada a respeito de probabilidade e estatística.

Ao adicionar uma medição a mais, x_{21}, a um conjunto com inicialmente 20 medições de uma dada grandeza, $\{X_1, X_2, \cdots, X_{20}\}$, a média aritmética μ do novo conjunto não se altera. Nesse caso, a variância σ^2 do conjunto inicial relaciona-se com a variância σ_n^2 do novo conjunto na forma $\sigma_n^2 = \frac{20}{21}\sigma^2$.

Comentários:

A variância do conjunto é dada por:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}$$

Para um conjunto com 20 elementos, temos:

$$\sigma^2 = \frac{\sum_{i=1}^{20}(x_i - \mu)^2}{20}$$

$$\sum_{i=1}^{20}(x_i - \mu)^2 = 20\sigma^2$$

A questão diz que, se adicionarmos mais um termo ao conjunto, a média não se altera, logo, esse elemento x_i é igual à média:

$$x_{21} = \mu$$

Assim, a variância do novo conjunto com 21 elementos é dada por:

$$\sigma_n^2 = \frac{\sum_{i=1}^{21}(x_i - \mu)^2}{21} = \frac{\sum_{i=1}^{20}(x_i - \mu)^2 + (x_{21} - \mu)^2}{21}$$

Agora, vamos substituir o valor de $\sum_{i=1}^{20}(x_i - \mu)^2$, que calculamos anteriormente:

$$\sigma_n^2 = \frac{\sum_{i=1}^{20}(x_i - \mu)^2 + (x_{21} - \mu)^2}{21} = \frac{20\sigma^2 + (\mu - \mu)^2}{21}$$

$$\sigma_n^2 = \frac{20}{21}\sigma^2$$

Gabarito: Certo.

3. (CESPE/PETROBRAS/2022) No que diz respeito aos conceitos e cálculos utilizados em probabilidade e estatística, julgue o item a seguir.

Se, em determinada semana, as ações da PETROBRAS fecharam o pregão com as cotações, em unidades monetária, iguais a 10,0; 9,0; 11,0; 12,0 e 8,0, respectivamente de segunda à sexta-feira, então a variância dessas cotações foi igual a 2,0.

Comentários:

A variância é determinada pela média dos quadrados dos desvios em relação à média aritmética. A variância populacional é dada por:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}$$

em que: x_i é o valor de ordem i assumido pela variável; μ é a média populacional de x; σ^2 é a variância populacional; e n é o número de dados da população.

Vamos iniciar calculando a média do conjunto:

$$\mu = \frac{10 + 9 + 11 + 12 + 8}{5} = \frac{50}{5} = 10$$

Agora, vamos calcular os desvios em relação à média. Para isso, vamos escrever os dados em uma tabela para melhor compreensão:

x_i	$x_i - \mu$	$(x_i - \mu)^2$
10	$10 - 10 = 0$	0
9	$9 - 10 = -1$	1
11	$11 - 10 = 1$	1
12	$12 - 10 = 2$	4
8	$8 - 10 = -2$	4
Total		**10**

Calculando a variância, temos:

$$\sigma^2 = \frac{10}{5} = 2$$

Gabarito: Certo.

4. (CESPE/PETROBRAS/2022)

X	Frequência Relativa
0	0,23
1	0,22
2	0,50
3	0,05

Considerando que a tabela acima mostra a distribuição de frequências de uma variável x obtida com base em uma amostra aleatória simples de tamanho igual a n, julgue o item que se segue.

A variância amostral de x é inferior a 0,7.

Comentários:

A variância amostral é simbolizada pela letra s, sendo calculada a partir de uma amostra da população. Quando os valores vierem dispostos em uma tabela de frequências, a variância será calculada por meio da seguinte fórmula:

$$s^2 = \frac{\sum_{i=1}^{m}(X_i - \bar{x})^2 \times f_i}{n - 1}$$

em que $\bar{x} = \frac{\sum_{i=1}^{m} X_i \times f_i}{n}$.

Então, vamos iniciar calculando a média. Para isso, consideraremos a frequência absoluta da tabela para uma amostra com n igual a 100.

x_i	f_i	$x_i \times f_i$
0	23	$0 \times 23 = 0$
1	22	$1 \times 22 = 22$
2	50	$2 \times 50 = 100$

3	5	$3 \times 5 = 15$
	$\sum f_i = 100$	$\sum f_i \times x_i = 137$

$$\bar{x} = \frac{\sum x_i \times f_i}{\sum f_i} = \frac{137}{100} = 1{,}37$$

Em seguida, adicionaremos uma nova coluna à tabela anterior, em que calcularemos os produtos dos quadrados dos desvios por suas respectivas frequências:

x_i	f_i	$x_i \times f_i$	$(x_i - \bar{x})^2 \times f_i$
0	23	$0 \times 23 = 0$	$(0 - 1{,}37)^2 \times 23 = 43{,}24$
1	22	$1 \times 22 = 22$	$(1 - 1{,}37)^2 \times 22 = 3{,}08$
2	50	$2 \times 50 = 100$	$(2 - 1{,}37)^2 \times 50 = 20$
3	5	$3 \times 5 = 15$	$(3 - 1{,}37)^2 \times 5 = 13{,}3$
	$\sum f_i = 100$	$\sum f_i \times x_i = 137$	$\sum (x_i - \bar{x})^2 \times f_i = \mathbf{79{,}62}$

Por fim, aplicamos a fórmula da variância amostral:

$$s^2 = \frac{\sum (x_i - \bar{x})^2 \times f_i}{n - 1}$$

$$s^2 = \frac{79{,}62}{100 - 1} = \frac{79{,}62}{99} = 0{,}8$$

Portanto, a variância amostral de x é SUPERIOR a 0,7

Gabarito: Errado.

5. (FGV/CGU/2022) O ativo A está gerando grande atração de matemáticos, que conseguiram convencer a bolsa de valores a registrar preços baseados em quantidades pouco usuais, como $\sqrt{2}$ e π. Durante cinco dias foram observados os seguintes preços de dois ativos, A e B, respectivamente:

$$\left(\sqrt{2}; \frac{\pi}{2}; 1; \frac{1}{3}; \frac{1}{6}\right) e\ (100{,}30; 400{,}18; 207{,}01; 508{,}00; 912{,}11)$$

Considerando esses valores, sobre a média e a variância dos retornos durante esses cinco dias, é correto afirmar que:

a) os retornos do ativo A têm maior média e maior variância;

b) os retornos do ativo A têm maior média e menor variância;

c) os retornos do ativo A têm menor média e maior variância;

d) os retornos do ativo A têm menor média e menor variância;

e) a média e a variância dos retornos dos dois ativos são iguais.

Comentários:

Vamos analisar a média dos dois conjuntos. No conjunto A, temos:

$$\left(\sqrt{2}; \frac{\pi}{2}; 1; \frac{1}{3}; \frac{1}{6}\right) \cong (1{,}4; 1{,}6; 1; 0{,}33; 0{,}16)$$

A média desse conjunto é próxima de 0,9, pois, somando todos os elementos, obtemos um valor em torno de 4,5. Dessa forma, variância do conjunto é mínima, pois os desvios em relação à média (variabilidade do conjunto) são pequenos.

Agora, vamos analisar o conjunto B:

$$(100{,}30; 400{,}18; 207{,}01; 508{,}00; 912{,}11)$$

Agora, a soma dos elementos está próxima de 2125, portanto, a média resulta em um valor em torno de 425. Analisando a variabilidade do conjunto, percebemos que os dados estão bem dispersos entre si, assim a variância do conjunto B será consideravelmente maior que a do conjunto A.

Então, comparando os dois conjuntos, temos que:

- a média de A é menor que a média de B.
- a variância de A é menor que a variância de B.

Gabarito: D.

6. (CESPE/TJ RJ/2021) Considere que, em um estudo para avaliar a satisfação dos serviços de comunicação de dados oferecidos por uma operadora, no qual foram utilizadas duas variáveis, X e Y, observou-se que X = 6Y + 24 e que o valor da variância de Y foi igual a 1. Nesse caso, o valor da variância de X é

a) 30.

b) 60.

c) 6.

d) 24.

e) 36.

Comentários:

Temos no enunciado que $X = 6Y + 24$. Temos também que a variância de Y é igual a 1 $\sigma^2(\text{Y}) = 1$. Queremos saber quanto vale $\sigma^2(\text{X})$. Assim, temos:

$$\sigma^2(X) = \sigma^2(6Y + 24)$$

Uma das propriedades da variância diz que somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, a variância do conjunto não é alterada. Assim, podemos desconsiderar a constante 24. Então, temos:

$$\sigma^2(X) = \sigma^2(6Y + 24)$$

$$\sigma^2(X) = \sigma^2(6Y)$$

$$\sigma^2(X) = 6^2\sigma^2(Y)$$

$$\sigma^2(X) = 36\sigma^2(Y)$$

$$\sigma^2(X) = 36 \times 1 = 36$$

Gabarito: E.

7. (CESPE/MJ SP/2021) Acerca de planejamento de pesquisa estatística, julgue o item que se seguem.

A média do erro entre a média calculada e as observações reais em um conjunto de dados é conhecida como variância.

Comentários:

A variância é determinada pela média dos quadrados dos desvios/erros em relação à média aritmética. Por meio dessa medida de dispersão ou variabilidade, podemos avaliar o quanto os dados estão dispersos em relação à média aritmética. A variância populacional é simbolizada pela letra grega σ (sigma), sendo calculada usando todos os elementos da população, pela seguinte fórmula:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}$$

em que: x_i é o valor de ordem i assumido pela variável; μ é a média populacional de x; σ^2 é a variância populacional; e n é o número de dados da população.

Gabarito: Certo.

8. (CESPE/BANESE/2021) A respeito do conjunto de dados {11, 6, 28, 51, 49, 32, 33}, julgue o item a seguir.

Esse conjunto de dados possui variância amostral inferior a 300.

Comentários:

A variância amostral é simbolizada pela letra s, sendo calculada a partir de uma amostra da população, pela seguinte fórmula:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

em que: x_i é o valor de ordem i assumido pela variável; $\bar{x}$ é a média amostral de x; s^2 é a variância amostral; e n é o número de dados da amostra.

Então, vamos iniciar pelo cálculo da média:

$$\bar{x} = \frac{11 + 6 + 28 + 51 + 49 + 32 + 33}{7} = \frac{210}{7} = 30$$

Agora, calculando a variância, temos:

$$s^2 = \frac{(11-30)^2 + (6-30)^2 + (28-30)^2 + (51-30)^2 + (49-30)^2 + (32-30)^2 + (33-30)^2}{7-1}$$

$$s^2 = \frac{361 + 576 + 4 + 441 + 361 + 4 + 9}{6}$$

$$s^2 = \frac{1756}{6}$$

$$s^2 = 292{,}66$$

Gabarito: Certo.

9. (CESPE/BANESE/2021)

	X	Y
Média	5	10
Desvio padrão	2	2

Com base nas informações apresentadas na tabela precedente e considerando que a covariância entre as variáveis X e Y seja igual a 3, julgue o item que se segue.

A variância de X é igual a 4.

Comentários:

A variância corresponde ao desvio padrão elevado ao quadrado:

$(\sigma) \rightarrow$ desvio padrão

$(\sigma^2) \rightarrow$ variância

Assim, basta elevarmos o desvio padrão de X ao quadrado para encontrarmos a variância:

$$\sigma^2 = 2^2 = 4$$

Gabarito: Certo.

10. (CESPE/SEDUC AL/2021) Com base em estatística, julgue o item a seguir.

Para um conjunto de dados $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_n}$ quaisquer, a variância será sempre um número positivo.

Comentários:

A variância é sempre maior ou igual a zero, isto é, nunca terá valor negativo, mas também pode assumir valor nulo. A variância de um conjunto é zero quando todos os elementos são iguais. Se todos os elementos são iguais, a média aritmética do conjunto coincide com o valor dos elementos e todos os desvios também são iguais a zero. Logo, a variância também é zero.

Gabarito: Errado.

11. (CESPE/TCE-RJ/2021)

X	Frequência Absoluta
0	5
1	10
2	20
3	15
Total	**50**

Considerando que a tabela precedente mostra a distribuição de frequências de uma variável quantitativa X, julgue o item a seguir.

A variância amostral de X é superior a 0,89.

Comentários:

A variância amostral é simbolizada pela letra s, sendo calculada a partir de uma amostra da população, pela seguinte fórmula:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

em que: x_i é o valor de ordem i assumido pela variável; $\bar{x}$ é a média amostral de x; s^2 é a variância amostral; e n é o número de dados da amostra.

Então, vamos iniciar calculando a média, para isso precisamos ponderar cada valor de x representado na tabela:

$$\bar{x} = \frac{(0 \times 5) + (1 \times 10) + (2 \times 20) + (3 \times 15)}{5 + 10 + 20 + 15} = \frac{95}{50} = 1{,}9$$

Calculando a variância:

$$s^2 = \frac{5(0 - 1{,}9)^2 + 10(1 - 1{,}9)^2 + 20(2 - 1{,}9)^2 + 15(3 - 1{,}9)^2}{50 - 1}$$

$$s^2 = \frac{18{,}5 + 8{,}1 + 0{,}2 + 18{,}15}{49}$$

$$s^2 = \frac{44{,}95}{49}$$

$$s^2 = 0{,}91$$

Gabarito: Certo.

12. (VUNESP/EsFCEx/2021) Uma amostra aleatória de 10 elementos de uma população para a estimação da média e da variância de uma variável com distribuição normal forneceu 500 e 25 844 para a soma dos valores e dos quadrados dos valores, respectivamente. É correto afirmar que a estimativa de máxima verossimilhança para a variância é

a) 84,4.

b) 25,844.

c) 258,44.

d) 9,38.

e) 93,8.

Comentários:

A variância é dada por:

$$\sigma^2 = \overline{x^2} - (\bar{x})^2$$

em que $\overline{x^2}$ é a média dos quadrados; e $(\bar{x})^2$ é o quadrado da média.

Ora, o enunciado nos diz que a soma dos valores do conjunto é igual a 500. Se dividirmos pela quantidade de elementos, encontraremos a média do conjunto:

$$\boldsymbol{\bar{x}} = \frac{500}{10} = 50$$

O enunciado também afirma que a soma dos quadrados dos valores é igual a 25844. Se dividirmos pela quantidade de elementos, encontraremos a média dos quadrados:

$$\overline{\boldsymbol{x^2}} = \frac{25844}{10} = 2.584,4$$

Então, temos:

$$\sigma^2 = \overline{x^2} - (\bar{x})^2$$

$$\sigma^2 = 2584,4 - (50)^2$$

$$\sigma^2 = 2584,4 - 2500$$

$$\sigma^2 = 84,4$$

Gabarito: A.

13. (VUNESP/EsFCEx/2021) Alunos de uma turma realizaram cinco provas de uma disciplina. Entretanto, o professor divulgou as notas das quatro provas e a variância populacional das cinco notas. João é aluno desta turma e deseja saber qual é a sua nota na 5a prova, as notas dele foram:

1ª Prova	2ª Prova	3ª Prova	4ª Prova
4	6	8	6

Sabendo que a variância populacional das suas notas foi 1,6, a nota do João na 5a Prova é

a) 3.

b) 5.

c) 6.

d) 2.

e) 4.

Comentários:

Vamos calcular a média das notas de João:

$$\boldsymbol{\bar{x}} = \frac{4+6+8+6+\mathrm{Y}}{5} = \frac{24+\mathrm{Y}}{5}$$

Agora, calculando a média dos quadrados, temos:

$$\overline{x^2} = \frac{4^2+6^2+8^2+6^2+\mathrm{Y}^2}{5} = \frac{152+\mathrm{Y}^2}{5}$$

Como sabemos, a variância pode ser expressa por:

$$\sigma^2 = \overline{x^2} - (\bar{x})^2$$

$$1{,}6 = \overline{x^2} - (\bar{x})^2$$

Substituindo as expressões anteriores, temos:

$$\sigma^2 = \left(\frac{152 + \mathrm{Y}^2}{5}\right) - \left(\frac{24 + \mathrm{Y}}{5}\right)^2$$

$$\sigma^2 = \left(\frac{152 + \mathrm{Y}^2}{5}\right) - \left(\frac{576 + 48\mathrm{Y} + \mathrm{Y}^2}{25}\right)$$

Colocando as frações na mesma base:

$$\sigma^2 = \left(\frac{760 + 5\mathrm{Y}^2}{25}\right) - \left(\frac{576 + 48\mathrm{Y} + \mathrm{Y}^2}{25}\right)$$

$$\sigma^2 = \left(\frac{760 + 5\mathrm{Y}^2 - 576 - 48\mathrm{Y} - \mathrm{Y}^2}{25}\right)$$

$$\sigma^2 = \left(\frac{760 + 5\mathrm{Y}^2 - 576 - 48\mathrm{Y} - \mathrm{Y}^2}{25}\right)$$

$$\sigma^2 = \left(\frac{184 - 48Y + 4\mathrm{Y}^2}{25}\right)$$

$$1{,}6 = \left(\frac{184 - 48Y + 4\mathrm{Y}^2}{25}\right)$$

$$40 = 184 - 48Y + 4\mathrm{Y}^2$$

$$4\mathrm{Y}^2 - 48Y + 144 = 0$$

Agora, dividindo todos os termos por 4, obtemos:

$$\mathrm{Y}^2 - 12Y + 38 = 0$$

Resolvendo a equação quadrática acima, encontramos que:

$$\Delta = 144 - 144 = 0$$

$$\mathrm{Y} = \frac{12}{2} = 6$$

Gabarito: C.

14. (FCC/ALAP/2020) O número de empregados de uma empresa é igual a 200, sendo que 60% são homens e o restante mulheres. Nesta empresa, a média aritmética dos salários da população formada pelos salários dos homens é igual a 5 mil reais, com um coeficiente de variação igual a 30%, e a média aritmética dos salários da população formada pelos salários das mulheres também é igual a 5 mil reais, porém com um coeficiente de variação igual a 20%. Considerando

a população formada por todos os 200 empregados da empresa, obtém-se que a variância, em mil reais ao quadrado, dos respectivos salários é igual a

a) 1,69

b) 1,75

c) 1,30

d) 2,50

e) 3,25

Comentários:

Vamos iniciar calculando o coeficiente de variação para as mulheres e para os homens:

$$C_v = \frac{\sigma}{\bar{x}}$$

Em que C_v é o coeficiente de variação; σ é o desvio padrão; e $\bar{x}$ é a média.

Portanto, para os homens, temos:

$$0{,}3 = \frac{\sigma_H}{5000}$$

$$\sigma_H = 5000 \times 0{,}3$$

$$\sigma_H = 1.500$$

Para as mulheres:

$$0{,}2 = \frac{\sigma_M}{5000}$$

$$\sigma_M = 5000 \times 0{,}2$$

$$\sigma_M = 1.000$$

Para os homens, teremos a seguinte variância:

$${\sigma_H}^2 = 1.500^2$$

$${\sigma_H}^2 = 2.250.000$$

$${\sigma_H}^2 = 2{,}25\ mil$$

Para mulheres, a variância será:

$${\sigma_M}^2 = 1.000^2$$

$${\sigma_M}^2 = 1.000$$

$${\sigma_M}^2 = 1\ mil$$

Agora, somando as variâncias de forma proporcional à representatividade de cada parcela, temos:

$$\sigma_T{}^2 = \sigma_H{}^2 \times 0{,}6 + \sigma_M{}^2 \times 0{,}4$$

$$\sigma_T{}^2 = 2{,}25 \times 0{,}6 + 1 \times 0{,}4$$

$$\sigma_T{}^2 = 1{,}35 + 0{,}4$$

$$\sigma_T{}^2 = 1{,}75$$

Gabarito: B.

15. (CESGRANRIO/BB/2018) Uma amostra aleatória de tamanho 5 é retirada de uma população e observa-se que seus valores, quando postos em ordem crescente, obedecem a uma Progressão Aritmética.

Se a variância amostral não viciada vale 40, qual é o valor da razão da Progressão Aritmética?

a) 3

b) $5\sqrt{2}$

c) 4

d) $2\sqrt{5}$

e) 1

Comentários:

O enunciado nos informa que 5 números formam uma P.A. Assim, podemos considerar que a diferença entre eles é r, que é a razão da progressão aritmética:

$$x - 2r; \quad x - r; \quad x; \quad x + r; \quad x + 2r$$

Calculando a média dessa amostra, temos:

$$\bar{x} = \frac{(x - 2r) + (x - r) + x + (x + r) + (x + 2r)}{5}$$

$$\bar{x} = \frac{5x}{5}$$

$$\bar{x} = x$$

Temos que a variância é dada por:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n - 1}$$

$$40 = \frac{\big((x - 2r) - x\big)^2 + \big((x - r) - x\big)^2 + (x - x)^2 + \big((x + r) - x\big)^2 + ((x + 2r) - x)^2}{5 - 1}$$

$$40 = \frac{(-2r)^2 + (-r)^2 + 0^2 + r^2 + (2r)^2}{4}$$

$$4r^2 + r^2 + r^2 + 4r^2 = 40 \times 4$$

$$10r^2 = 160$$

$$r^2 = 16$$

$$r = \sqrt{16}$$

Assim, descobrimos que a razão da progressão aritmética é:

$$r = 4$$

Gabarito: C.

16. (FCC/TRT 14ª Região/2018) Considere uma população P_1 formada pela renda, em unidades monetárias (u.m.), dos 100 indivíduos que são sócios de um clube. Seja x_i a renda, $x_i > 0$, do sócio i.

Dados:

$\sum_{i=1}^{100} x_i^2 = 2.662.400(u.m)^2$ e Coeficiente de variação de P_1 igual a 20%.

Decide-se excluir de P_1 um total de 20 sócios que possuem renda igual à média de P_1 , formando uma nova população P_2 com tamanho 80. O módulo da diferença, em (u.m.) ^{2}, entre as variâncias de P_1 e P_2 é de

a) 144.

b) 0.

c) 64.

d) 256.

e) 400.

Comentários:

Iniciaremos o cálculo pelas medidas da população 1. Considerando os dados do problema, a variância de $\boldsymbol{P_1}$ é:

$$\sigma_1{}^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}$$

$$\sigma_1{}^2 = \frac{1}{n} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{x_i}^2 - \bar{x}^2$$

$$\sigma_1{}^2 = \frac{\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{x_i}^2}{n} - \overline{x_1}^2$$

$$\sigma_1{}^2 = \frac{2.662.400}{100} - \overline{x_1}^2$$

$$\sigma_1{}^2 = 26.624 - \overline{x_1}^2$$

$$\overline{x_1}^2 = 26.624 - {\sigma_1}^2$$

Aplicando a fórmula do coeficiente de variação, temos:

$$C_v = \frac{\sigma}{\bar{x}}$$

$$0{,}2 = \frac{\sigma_1}{\overline{x_1}}$$

Elevando todos os termos ao quadrado, temos:

$$0{,}2^2 = \frac{\sigma_1^2}{\overline{x_1}^2} \Longrightarrow 0{,}04 = \frac{{\sigma_1}^2}{\overline{x_1}^2}$$

$$0{,}04 = \frac{{\sigma_1}^2}{26.624 - {\sigma_1}^2}$$

$$\sigma_1^2 = 1.064{,}96 - 0{,}04 \times {\sigma_1}^2$$

$$\sigma_1^2 + 0{,}04 \times {\sigma_1}^2 = 1.064{,}96$$

$$1{,}04 \times {\sigma_1}^2 = 1.064{,}96$$

$${\sigma_1}^2 = \frac{1.064{,}96}{1{,}04}$$

$${\sigma_1}^2 = 1.024$$

Então:

$${\sigma_1}^2 = \sum_{i=1}^{n} \frac{(x_i - \overline{x_1})^2}{n}$$

$$1.024 = \sum_{i=1}^{100} \frac{(x_i - \overline{x_1})^2}{100}$$

$$\sum_{i=1}^{100} (x_i - \overline{x_1})^2 = 102.400$$

Excluindo 20 sócios de $\boldsymbol{P_1}$ com salários iguais à média, formamos a população $\boldsymbol{P_2}$:

$$\overline{x_2} = \frac{100\overline{x_1} - 20\overline{x_1}}{80} = \frac{80\overline{x_1}}{80}$$

$$\overline{x_2} = \overline{x_1}$$

Calculando a variância de $\boldsymbol{P_2}$:

$${\sigma_2}^2 = \sum_{i=1}^{n} \frac{(x_i - \overline{x_1})^2}{n} \rightarrow {\sigma_2}^2 = \sum_{i=21}^{80} \frac{(x_i - \overline{x_1})^2}{80}$$

$$\sigma_2{}^2 = \sum_{i=1}^{100} \frac{(x_i - \overline{x_1})^2}{80}$$

$$\sigma_2{}^2 = \frac{102.400}{80}$$

$$\sigma_2{}^2 = 1.280$$

Então, o módulo da diferença entre as variâncias, em $(\text{u.m.})^2$, será:

$$1280 - 1024 = 256$$

Gabarito: D.

17. (FGV/SEPOG-RO/2017) Considere a seguinte amostra de notas de alunos: 5,0; 6,0; 5,0; 4,0; 10,0. A variância amostral dessas notas pode ser igual a

a) 5,5.

b) 6,2.

c) 6,8.

d) 7,2.

e) 9,0.

Comentários:

Inicialmente, vamos calcular a média:

$$\bar{x} = \frac{5 + 6 + 5 + 4 + 10}{5} = \frac{30}{5}$$

$$\bar{x} = 6$$

Agora, calcularemos os desvios de cada nota em relação à média:

Notas (x_i)	Desvio em relação à média $(x_i - \bar{x})$	$(x_i - \bar{x})^2$
5	$5 - 6 = -1$	1
6	$6 - 6 = 0$	0
5	$5 - 6 = -1$	1
4	$4 - 6 = -2$	4

10	$10 - 6 = 4$	16
Total		**22**

De posse dessa informação, podemos calcular a variância. Devemos lembrar que, por se tratar de uma variância amostral, aplicamos $n - 1$.

$$s^2 = \frac{22}{5-1} = \frac{22}{4} = 5{,}5$$

Gabarito: A.

18. (FCC/DPE SP/2015) Foi realizado um censo em uma faculdade com 200 alunos e obteve-se com relação às alturas dos alunos, em centímetros (cm), um coeficiente de variação igual a 10%. Se a soma dos quadrados de todas as alturas foi igual a 5.499.450 cm 2, então a correspondente variância apresentou um valor igual a

a) 289,00 cm 2.

b) 256,00 cm 2.

c) 306,25 cm 2.

d) 324,00 cm 2.

e) 272,25 cm 2.

Comentários:

Iniciaremos calculando a média dos quadrados:

$$\overline{x^2} = \frac{5.499.450}{200} = 27.497{,}25$$

O coeficiente de variação foi informado na questão. Sabemos que ele é resultado da divisão entre o desvio padrão e a média, então:

$$\frac{10}{100} = \frac{\sigma}{\bar{x}}$$

$$\bar{x} = 10\sigma$$

A variância resulta da diferença entre a média dos quadrados e o quadrado da média. Vamos aplicar o valor da média na fórmula da variância:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

$$\sigma^2 = 27.497{,}25 - (10\sigma)^2$$

$$\sigma^2 = 27.497{,}25 - 100\sigma^2$$

$$101\sigma^2 = 27.497{,}25$$

$$\sigma^2 = \frac{27.497,25}{1,001}$$

$$\sigma^2 = 272,25$$

Gabarito: E.

19. (FCC/CNMP/2015) Em um censo realizado em um clube apurou-se a altura em centímetros (cm) de seus 200 associados. A média aritmética apresentou um valor igual a 160 cm com um coeficiente de variação igual a 18,75%. O resultado da divisão da soma de todos os valores das alturas elevados ao quadrado pelo número de associados é, em cm^2, de

a) 27.050.

b) 25.600.

c) 26.050.

d) 26.500.

e) 25.060.

Comentários:

Primeiro, vamos calcular o desvio padrão. Temos no enunciado as informações do coeficiente de variação e da média. Sabemos que o coeficiente de variação é dado pela divisão entre o desvio padrão e a média, logo, temos:

$$C_v = \frac{\sigma}{\bar{x}}$$

$$0,1875 = \frac{\sigma}{160}$$

$$\sigma = 160 \times 0,1875$$

$$\sigma = 30$$

Agora, vamos elevar o desvio padrão ao quadrado para encontrarmos a variância:

$$\sigma^2 = 30^2$$

$$\sigma^2 = 900$$

Para calcularmos o que pede a questão, precisamos saber outra forma de calcularmos a variância, que também pode ser dada pela diferença entre a média dos quadrados e o quadrado da média:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

A média dos quadrados é exatamente o resultado da divisão da soma de todos os valores das alturas elevados ao quadrado pelo número de associados. Demonstrado:

$$\overline{x^2} = \frac{\sum x^2}{n}$$

Então, temos:

$$\sigma^2 = \frac{\sum x^2}{n} - \bar{x}^2$$

$$900 = \frac{\sum x^2}{n} - 160^2$$

$$\frac{\sum x^2}{n} = 900 + 160^2$$

$$\frac{\sum x^2}{n} = 26.500$$

Gabarito: D.

20. (CESGRANRIO/EPE/2014) Uma amostra de tamanho 200, $x_1, x_2, \dots, x_{200}$, foi retirada de uma população, e seus valores foram transformados segundo a função $y_i = 4x_i - 1$ para $i = 1, 2, \dots, 200$.

Sabendo-se que a média e a variância dos dados transformados $y_1, y_2, \dots, y_{200}$ são, respectivamente, 3 e 16, os valores da média e da variância dos dados originais são, respectivamente,

a) 1 e 1

b) 1 e 4

c) 3/4 e 63

d) 11 e 64

e) 11 e 256

Comentários:

Ao somarmos ou subtrairmos uma constante de todos os valores de uma variável, a média do conjunto também aumenta ou diminui no valor da constante. A mesma regra se aplica à multiplicação e divisão, a média é multiplicada ou dividida pela constante.

Assim, temos a função dada no enunciado:

$$y_i = 4x_i - 1$$

Calculando a média dos x_i:

$$\bar{x} = \frac{\bar{y} + 1}{4}$$

$$\bar{x} = \frac{3 + 1}{4}$$

$$\bar{x} = 1$$

Para a variância a regra é outra, **somando-se ou subtraindo-se uma constante de todos os valores de uma variável**, **a variância do conjunto não se altera**. Porém, **se multiplicarmos ou dividirmos todos os valores de uma variável por uma constante, a variância fica multiplicada ou dividida pelo quadrado da constante**.

Assim temos:

$$Var(x) = Var\left(\frac{y+1}{4}\right)$$

$$Var(x) = \frac{1}{4^2} \times Var(y+1)$$

$$Var(x) = \frac{1}{4^2} \times Var(y)$$

$$Var(x) = \frac{16}{4^2}$$

$$Var(x) = \frac{16}{16}$$

$$Var(x) = 1$$

Gabarito: A.

21. (CESGRANRIO/FINEP/2014) O enunciado a seguir deve ser usado para responder à questão.

Abaixo são apresentadas estatísticas das notas brutas obtidas pelos candidatos em um concurso público:

Média aritmética: 78

Variância: 100

A nota de cada candidato foi transformada em nota padronizada, calculada considerando-se a seguinte fórmula:

$$\boldsymbol{Nota\ padronizada = 50 + 5 \times \frac{Nota\ bruta\ do\ candidato\ -\ Media\ aritmetica\ das\ notas\ brutas}{Desvio\ padrao\ das\ notas\ brutas}}$$

A variância das notas padronizadas é

a) 25

b) 50,5

c) 52,5

d) 55

e) 75

Comentários:

Pelas propriedades, sabemos que a variância de uma constante é 0 (zero). Também sabemos que o desvio padrão (σ) não é influenciado pela adição, e que corresponde à raiz da variância. Assim, temos:

$$Var_{N_{padron}} = Var(50) + Var\left(5 \times \left(\frac{N_{bruta} - \bar{x}}{\sigma}\right)\right)$$

$$Var_{N_{padron}} = 0 + Var\left(\frac{5 \times (N_{bruta} - 78)}{10}\right)$$

Agora, **se multiplicarmos ou dividirmos todos os valores de uma variável por uma constante, a variância fica multiplicada ou dividida pelo QUADRADO da constante:**

$$Var(aX + b) = a^2 \times Var(X)$$

Aplicando essa propriedade, temos:

$$Var_{N_{padron}} = Var\left(\left(\frac{5}{10}\right) \times (N_{bruta} - 78)\right)$$

$$Var_{N_{padron}} = \left(\frac{5}{10}\right)^2 \times Var(N_{bruta} - 78)$$

Somando-se ou subtraindo-se uma constante de todos os valores de uma variável, a variância do conjunto não se altera, logo, $Var(N_{bruta} - 78) = Var(N_{bruta})$.

$$Var_{N_{padron}} = \frac{25}{100} \times Var(N_{bruta})$$

$$Var_{N_{padron}} = \frac{25}{100} \times 100$$

$$Var_{N_{padron}} = 25$$

Gabarito: A.

22. (FCC/TRT 16ª Região/2014) Seja {$X_1, X_2, X_3, ..., X_{80}$} uma população constituída de 80 números estritamente positivos, sabendo-se que a média aritmética e o desvio padrão desta população são, respectivamente iguais a 20 e 15. Resolve-se excluir desta população 30 números, cuja soma de seus quadrados é igual a 12.000, formando uma nova população e o novo valor da variância passa a ter o valor de 436. O correspondente novo valor da média aritmética da nova população apresenta um valor igual a

a) 16.

b) 24.

c) 22.

d) 23.

e) 18.

Comentários:

De início, vamos calcular a variância usando a seguinte fórmula:

$$\sigma^2 = \frac{\sum_{i=1}^{n} \bar{x}^2}{n} - \bar{x}^2$$

$$15^2 = \frac{\sum_{i=1}^{n} \bar{x}^2}{80} - 20^2$$

$$\sum_{i=1}^{n} x^2 = 50.000$$

Após a retirada de 30 elementos da população, a nova média será:

$$436 = \frac{50.000 - 12.000}{80 - 30} - \bar{x}^2$$

$$436 = \frac{38.000}{50} - \bar{x}^2$$

$$\bar{x}^2 = 760 - 436$$

$$\bar{x}^2 = 18$$

Gabarito: E.

23. (FCC/TRT 13ª Região/2014) Em uma determinada carreira profissional composta por 400 trabalhadores, verifica-se que a média aritmética das alturas de todos os trabalhadores é igual a 170 cm. Sabe-se que a média aritmética das alturas dos 250 trabalhadores do sexo masculino é igual à média aritmética das alturas dos 150 trabalhadores do sexo feminino. Os desvios padrões das alturas dos trabalhadores do sexo masculino e dos trabalhadores do sexo feminino são iguais a 12 cm e 20 cm, respectivamente. A variância (em cm^2) das alturas de todos os trabalhadores desta carreira profissional é igual a

a) 232.

b) 225.

c) 228.

d) 196.

e) 240.

Comentários:

Nessa questão, a média das alturas dos trabalhadores do sexo masculino é igual à média das trabalhadoras do sexo feminino. Logo, para encontrarmos a variância de todos os trabalhadores, precisamos calcular a variância ponderada para o sexo masculino e feminino.

A variância pode ser encontrada elevando o desvio padrão ao quadrado. Assim, temos que:

$$\sigma_M^2 = 12^2 = 144$$

$$\sigma_F^2 = 20^2 = 400$$

Agora, podemos calcular a variância total:

$$\sigma^2 = \frac{\sigma_M^2 \times 250 + \sigma_F^2 \times 150}{400}$$

$$\sigma^2 = \frac{144 \times 250 + 400 \times 150}{400}$$

$$\sigma^2 = \frac{36.000 + 60.000}{400}$$

$$\sigma^2 = \frac{96.000}{400}$$

$$\sigma^2 = 240$$

Gabarito: E.

24. (FGV/AL-BA/2014) A média das idades de um grupo de 4 amigos é de 36 anos, e o desvio padrão é igual a 2. Daqui a cinco anos, a média e a variância das idades desse grupo serão iguais a:

a) 41 e 4.

b) 41 e 50.

c) 56 e 2.

d) 56 e 50.

e) 56 e 200.

Comentários:

O desvio padrão e a variância não se alteram com a adição ou subtração de uma constante c. Sabemos que a variância é igual ao desvio padrão ao quadrado. Logo:

$$\sigma^2 = 2^2$$

$$\sigma^2 = 4$$

Como a média era 36 e foram somados mais cinco anos, a nova média também será adicionada de 5 unidades:

$$\bar{x} = 36 + 5 = 41$$

Gabarito: A.

25. (CESGRANRIO/BNDES/2013) Em um departamento de uma empresa, o gerente decide dar um aumento a todos os empregados, dobrando o salário de todos eles.

Em relação às estatísticas dos novos salários, considere as afirmativas abaixo.

I - A média dobra.

II - A variância dobra.

III - A moda dobra.

É correto o que se afirma em

a) I, apenas

b) II, apenas

c) I e III, apenas

d) II e III, apenas

e) I, II e III

Comentários:

Sabemos que, ao somarmos ou subtrairmos uma constante de todos os valores de uma variável, a média do conjunto também **aumenta ou diminui no valor da constante**. A mesma regra se aplica à multiplicação e divisão, a média é **multiplicada ou dividida pela constante**.

Para a variância a regra é outra, somando-se ou subtraindo-se uma constante de todos os valores de uma variável, a variância do conjunto **não se altera**. Porém, se multiplicarmos ou dividirmos todos os valores de uma variável por uma constante, a variância fica **multiplicada ou dividida pelo quadrado da constante**.

Assim, apenas os itens I e III estão corretos.

Gabarito: C.

26. (FCC/TRF 2ª Região/2012) A soma dos quadrados dos valores dos elementos de uma população de tamanho 20 é igual a 65,6 e o respectivo desvio padrão igual a 0,2. A média aritmética dos elementos desta população é igual a

a) 0,8.

b) 1,2.

c) 1,8.

d) 2,4.

e) 3,0.

Comentários:

A questão nos forneceu a soma dos quadrados e o desvio padrão. Sabendo que o desvio padrão é igual a raiz quadrada da variância, podemos encontrar a média por meio do cálculo da variância com os dados que temos. A variância é dada pela diferença entre a média dos quadrados e o quadrado da média:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

A média dos quadrados é obtida pela divisão entre a soma dos quadrados e o número de observações:

$$\overline{x^2} = \frac{65,6}{20} = 3,28$$

E a variância equivale ao quadrado do desvio padrão:

$$\sigma^2 = 0,2^2 = 0,04$$

Aplicando a fórmula da variância, temos:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

$$0,04 = 3,28 - \bar{x}^2$$

$$\bar{x}^2 = 3,28 - 0,04$$

$$\bar{x}^2 = 3,24$$

$$\bar{x} = \sqrt{3,24}$$

$$\bar{x} = 1,8$$

Gabarito: C.

27. (FCC/TRT 6ª Região/2012) Em um censo realizado em uma empresa, verificou-se que a média aritmética dos salários de seus empregados foi igual a R$ 2.000,00 com um desvio padrão igual a R$ 125,00. Analisando, separadamente, o grupo de todos os empregados do sexo masculino e o grupo de todos os empregados do sexo feminino obteve-se as seguintes informações:

Grupo	Média Aritmética (R$)	Desvio Padrão (R$)
Homens	2.000,00	150,00
Mulheres	2.000,00	100,00

A porcentagem de empregados do sexo masculino na empresa é de

a) 40%.

b) 45%.

c) 48%.

d) 50%.

e) 52%.

Comentários:

Conforme o enunciado, a média é a mesma para homens e mulheres. Logo, para a variância total dos dois grupos, precisamos calcular a variância ponderada para os homens e para as mulheres. A variância é encontrada elevando o desvio padrão ao quadrado. Assim, temos que:

$$\sigma_H^2 = 150^2 = 22.500$$

$$\sigma_M^2 = 100^2 = 10.000$$

$$\sigma^2 = 125^2 = 15.625$$

Agora, podemos calcular a média ponderada, considerando H o percentual para homens e M o percentual para mulheres:

$$15.625 = 22.500 \times H + 10.000 \times M$$

Devemos lembrar que:

$$H + M = 100\% = 1$$

$$M = 1 - H$$

Substituindo, temos:

$$15.625 = 22.500H + 10.000 \times (1 - H)$$

$$15.625 = 22.500H + 10.000 - 10.000H$$

$$22.500H + 10.000H = 15.625 - 10.000$$

$$12.500H = 5.625$$

$$H = \frac{5.625}{12.500}$$

$$H = 0{,}45 \rightarrow 45\%$$

Gabarito: B.

28. (CESGRANRIO/BNDES/2010) Em uma pesquisa de preços de determinado produto, foram obtidos os valores, em reais, de uma amostra aleatória colhida em 6 estabelecimentos que o comercializam.

Ação da Empresa	Resultado
P	5,00
Q	8,00

R	**6,00**
S	**6,00**
T	**4,00**
U	**7,00**

A variância dessa amostra é

a) 1,50

b) 1,75

c) 2,00

d) 2,25

e) 2,50

Comentários:

Vamos iniciar calculando a média:

$$\bar{x} = \frac{5+8+6+6+4+7}{6}$$

$$\bar{x} = \frac{36}{6}$$

$$\bar{x} = 6$$

Agora, vamos calcular os desvios de cada nota em relação à média:

(x_i)	5	8	6	6	4	7	**Total**
Desvio em relação à média $(x_i - \bar{x})$	5-6=-1	8-6=2	6-6=0	6-6=0	4-6=-2	7-6=1	
$(x_i - \bar{x})^2$	1	4	0	0	4	1	**10**

Pronto, já podemos calcular a variância. Lembrando que, como se trata de variância amostral, aplicamos n-1.

$$s^2 = \frac{10}{n-1} = \frac{10}{6-1} = \frac{10}{5} = 2$$

Gabarito: C.

29. (FGV/SEFAZ-RJ/2010) A média, a mediana e a variância das idades de um grupo de vinte pessoas são, hoje, iguais, respectivamente, a 34, 35 e 24. Daqui a dez anos, os valores da média, da mediana e da variância das idades dessas pessoas serão, respectivamente:

a) 44, 35 e 34.

b) 44, 45 e 12.

c) 44, 45 e 24.

d) 34, 35 e 12.

e) 44, 45 e 124.

Comentários:

Ao somarmos uma constante a um conjunto, a média e mediana também aumentarão no valor dessa constante. Assim temos que:

$$\bar{x} = 34 + 10 = 44$$

$$M_d = 35 + 10 = 45$$

Para a variância, temos que a regra é que se somarmos uma constante ela não será alterada. Portanto:

$$\sigma^2 = 24$$

Gabarito: C.

30. (FGV/SEAD-AP/2010) Os dados a seguir são as quantidades de empregados de cinco pequenas empresas: 6, 5, 8, 5, 6.

A variância da quantidade de empregados dessas cinco empresas é igual a:

a) 0,8.

b) 1,2.

c) 1,6.

d) 2,0.

e) 2,4.

Comentários:

Queremos calcular a variância, para isso precisamos inicialmente calcular a média de empregados das 5 empresas:

$$\bar{x} = \frac{6 + 5 + 8 + 5 + 6}{5}$$

$$\bar{x} = 6$$

Agora vamos calcular a média dos quadrados:

$$\overline{x^2} = \frac{6^2 + 5^2 + 8^2 + 5^2 + 6^2}{5}$$

$$\overline{x^2} = \frac{186}{5}$$

$$\overline{x^2} = 37,2$$

Sabemos que a variância é dada pela diferença entre a média dos quadrados e o quadrado da média, logo:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

$$\sigma^2 = 37,2 - 6^2$$

$$\sigma^2 = 1,2$$

Gabarito: B.

31. (FGV/SEN/2008) A média

$$\left(\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\right)$$

e a variância amostral

$$\left(s^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2\right)$$

de um conjunto de 20 observações são, respectivamente, 5 e 1. Uma nova observação, de valor igual a 5, foi acrescentada ao conjunto inicial, passando-se a ter 21 valores. A nova variância amostral será igual a:

a) 1,10.

b) 1,05.

c) 1,00.

d) 0,95.

e) 0,90.

Comentários:

Temos que a média inicial vale 5 e que são 20 observações, então:

$$5 = \frac{\sum_{i=1}^{20} x_i}{20}$$

$$\sum_{i=1}^{20} x_i = 20 \times 5$$

$$\sum_{i=1}^{20} x_i = 100$$

Agora, vamos acrescentar mais um valor igual a 5 no conjunto e calcular a nova média:

$$\bar{x} = \frac{1}{21}\sum_{i=1}^{20} x_i$$

$$\bar{x} = \frac{100 + 5}{21}$$

$$\bar{x} = \frac{105}{21}$$

$$\bar{x} = 5$$

A nova média também vale 5.

Calculando a nova variância. Temos no enunciado que inicialmente ela vale 1:

$$1 = \frac{\sum_{i=1}^{20}(x_i - 5)^2}{20 - 1}$$

$$\sum_{i=1}^{20}(x_i - 5)^2 = (20 - 1) \times 1$$

$$\sum_{i=1}^{20}(x_i - 5)^2 = 19$$

Acrescentando a 21ª observação:

$$s^2 = \frac{\sum_{i=1}^{21}(x_i - 5)^2 + (5 - 5)^2}{21 - 1}$$

$$s^2 = \frac{19 + 0}{20}$$

$$s^2 = 0{,}95$$

Gabarito: D.

Questões Comentadas

Desvio-Padrão

1. (CESPE/PC PB/2022 - ADAPTADA)

Situação hipotética 17A4-I

Um padrão de referência possui concentração de 25 mg/mL da substância X. Um técnico, ao calibrar dois aparelhos que medem a concentração desta substância X, fez medidas durante 5 dias (amostra 1 no dia 1, amostra 2 no dia 2, e assim por diante) e encontrou os seguintes valores.

Aparelho A

Amostra	Concentração (mg/ml)
1	21
2	22
3	21
4	20
5	21

Aparelho B

Amostra	Concentração (mg/ml)
1	29
2	25
3	21
4	24
5	26

Considerando os dados obtidos na situação hipotética 17A4-I, os valores para a média e desvio-padrão dos aparelhos A e B são

a) médiaA= 21 mg/mL; desvio-padrãoA= 0,71 mg/mL; médiaB= 25 mg/mL; desvio-padrãoB= 2,91 mg/mL;

b) médiaA= 21 mg/mL; desvio-padrãoA= 2 mg/mL; médiaB= 25 mg/mL; desvio-padrãoB= 10 mg/mL;

c) médiaA= 21 mg/mL; desvio-padrãoA= 0,63 mg/mL; médiaB= 24 mg/mL; desvio-padrãoB= 3,58 mg/mL;

d) médiaA= 21 mg/mL; desvio-padrãoA= 2 mg/mL; médiaB= 24 mg/mL; desvio-padrãoB= 4 mg/mL;

e) médiaA= 21 mg/mL; desvio-padrãoA= 0,71 mg/mL; médiaB= 24 mg/mL; desvio-padrãoB= 4 mg/mL;

Comentários:

Para a amostra A, temos:

a.1) média:

$$\bar{x}_A = \frac{21 + 22 + 21 + 20 + 21}{5} = \frac{105}{5} = 21$$

b.1) desvio padrão:

$$\sigma^2{}_A = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

$$\sigma^2{}_A = \frac{(21-21)^2 + (22-21)^2 + (21-21)^2 + (20-21)^2 + (21-21)^2}{5-1}$$

$$\sigma^2{}_A = \frac{(0)^2 + (1)^2 + (0)^2 + (-1)^2 + (0)^2}{5-1}$$

$$\sigma^2{}_A = \frac{2}{4} = \frac{1}{2}$$

$$\sigma_A = \sqrt{\frac{1}{2}} = \frac{\sqrt{1}}{\sqrt{2}} = \frac{\sqrt{2}}{2} = 0{,}71$$

Para a amostra B, temos:

a.2) média:

$$\bar{x}_B = \frac{29 + 25 + 21 + 24 + 26}{5} = \frac{125}{5} = 25$$

b.2) desvio padrão:

$$\sigma^2{}_B = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

$$\sigma^2{}_B = \frac{(29-25)^2 + (25-25)^2 + (21-25)^2 + (24-25)^2 + (26-25)^2}{5-1}$$

$$\sigma^2{}_B = \frac{(4)^2 + (0)^2 + (-4)^2 + (-1)^2 + (1)^2}{5-1}$$

$$\sigma^2{}_B = \frac{34}{4} = \frac{17}{2}$$

$$\sigma_B = \sqrt{\frac{17}{2}} = 2{,}91$$

Na questão original, a banca informava que o desvio padrão de B era igual a 4, quando deveria ser de 2,91.

Gabarito: A.

2. (CESPE/COREN SE/2021) Considere que os tempos de espera X e de atendimento Y, ambos em minutos, para determinado serviço ambulatorial se relacionem como Y = 2X – 1. Se o desvio padrão de X for igual a 2 minutos, então o desvio padrão de Y, em minutos, será igual a

a) 2.

b) 5.

c) 3.

d) 4.

Comentários:

O enunciado nos informa que $Y = 2X - 1$. Também nos diz que o desvio padrão de X é igual a 2 $\sigma(\text{X}) = 2$. Queremos saber quanto vale $\sigma(\text{Y})$. Assim, temos:

$$\sigma(Y) = \sigma(2\text{X} - 1)$$

Uma das propriedades da variância diz que somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, a variância do conjunto não é alterada. Assim, podemos desconsiderar a constante (-1). Então, temos:

$$\sigma(Y) = \sigma(2\text{X} - 1)$$

$$\sigma(Y) = \sigma(2\text{X})$$

$$\sigma(Y) = 2(2)$$

$$\sigma(Y) = 2 \times 2$$

$$\sigma(Y) = 4$$

Gabarito: D.

3. (CESPE/PC SE/2021) Com base em uma amostra aleatória simples de tamanho n = 16 retirada de uma população normal com média desconhecida μ e variância $\sigma^2 = 9$, deseja-se testar a hipótese nula $H_0: \mu = 0$ contra a hipótese alternativa $H1: \mu \neq 0$ por meio da estatística $\sqrt{n}\overline{X}/\sigma$, na qual $\overline{X}$ denota a média amostral.

Com respeito a esse teste de hipóteses, julgue o item a seguir, sabendo que o valor da média amostral observado na amostra foi igual a 1 e que, relativo a esse teste, o P-valor foi igual a 0,18.

O desvio padrão da média amostral $\overline{\boldsymbol{X}}$ é igual a 0,75.

Comentários:

O desvio padrão da média amostral é expresso por:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

O enunciado nos informa que $n = 16$ e que a variância é igual a 9. Sabemos que o desvio padrão é a raiz quadrada da variância, então, substituindo na fórmula, temos:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{9}}{\sqrt{16}}$$

$$\sigma_{\bar{x}} = \frac{3}{4} = 0{,}75$$

Gabarito: Certo.

4. (CESPE/CBM AL/2021) Determinado dado tetraédrico (dado em formato de tetraedro regular), com vértices numerados de 1 a 4, foi lançado 21 vezes, de modo que o resultado do lançamento desse dado correspondia ao vértice voltado para cima. A tabela seguinte mostra a frequência com que se obteve cada resultado.

Resultado	Quantidade de lançamentos
1	2
2	5
3	5
4	9

Com base nessa situação hipotética, julgue o item a seguir.

O desvio padrão dos resultados é superior a 1.

Comentários:

Vamos iniciar calculando a média do conjunto. Para isso, precisamos multiplicar cada resultado pela quantidade de lançamentos:

$$\mu = \frac{(1 \times 2) + (2 \times 5) + (3 \times 5) + (4 \times 9)}{2 + 5 + 5 + 9} = \frac{63}{21} = 3$$

O desvio padrão é dado por:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2 \times f_i}{n}}$$

em que n é a quantidade de lançamentos.

$$\sigma = \sqrt{\frac{2(1-3)^2 + 5(2-3)^2 + 5(3-3)^2 + 9(4-3)^2}{21}}$$

$$\sigma = \sqrt{\frac{(2 \times 4) + (5 \times 1) + (5 \times 0) + (9 \times 1)}{21}}$$

$$\sigma = \sqrt{\frac{8 + 5 + 0 + 9}{21}}$$

$$\sigma = \sqrt{\frac{22}{21}}$$

$$\sigma \cong 1{,}024$$

$$\sigma > 1$$

Portando, o desvio padrão é maior que 1.

Gabarito: Certo.

5. (FGV/Pref. Paulínia/2021) Em uma turma de 10 alunos, as notas dos alunos em uma avaliação foram:

6	7	7	8	8	8	8	9	9	10

O desvio padrão dessa lista de notas é, aproximadamente,

a) 0,8.

b) 0,9.

c) 1,1.

d) 1,3.

e) 1,5.

Comentários:

Calculando a média do conjunto:

$$\bar{x} = \frac{6+7+7+8+8+8+8+9+9+10}{10} = \frac{80}{10} = 8$$

Temos que o desvio padrão é dado por:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}$$

em que n é a quantidade de alunos.

$$\sigma = \sqrt{\frac{(6-8)^2+(7-8)^2+(7-8)^2+(8-8)^2+(8-8)^2+(8-8)^2+(8-8)^2+(9-8)^2+(9-8)^2+(10-8)^2}{10}}$$

$$\sigma = \sqrt{\frac{4+1+1+0+0+0+0+1+1+4}{10}}$$

$$\sigma = \sqrt{\frac{12}{10}}$$

$$\sigma = \sqrt{1{,}2}$$

$$\sigma \cong 1{,}1$$

Gabarito: C.

6. (FGV/IMBEL/2021) Considere os números 1, 2, 3, 6, 8.

O desvio padrão dessa lista de números é, aproximadamente, igual a

a) 1,7.

b) 2,1.

c) 2,6.

d) 3,0.

e) 3,4.

Comentários:

Calculando a média do conjunto:

$$\mu = \frac{1+2+3+6+8}{5} = \frac{20}{5} = 4$$

Temos que o desvio padrão é dado por:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}}$$

em que n é a quantidade de elementos do conjunto.

$$\sigma = \sqrt{\frac{(1-4)^2 + (2-4)^2 + (3-4)^2 + (6-4)^2 + (8-4)^2}{5}}$$

$$\sigma = \sqrt{\frac{9+4+1+4+16}{5}}$$

$$\sigma = \sqrt{\frac{34}{5}}$$

$$\sigma = \sqrt{6{,}8}$$

$$\sigma \cong 2{,}6$$

Gabarito: C.

7. (VUNESP/PB Saúde/2021) Seja P_1 uma população formada pelos salários dos 20 funcionários, em R\$ 1.000,00, de uma empresa X, e seja P_2 a população formada pelos salários dos 15 funcionários, em R\$ 1.000,00, de uma outra empresa Y. As somas dos quadrados de todos os elementos de P_1 e P_2, em $(R\$\,1.000{,}00)^2$, são iguais a 323,2 e 375,6, respectivamente. O coeficiente de variação de P_1 é igual a 10%, e a média aritmética dos salários de P_1 é igual a 80% da média aritmética dos salários de P_2. Então, o desvio padrão de P_2 é igual a

a) R\$ 200,00.

b) R\$ 300,00.

c) R\$ 400,00.

d) R\$ 500,00.

e) R\$ 600,00.

Comentários:

Inicialmente, vamos calcular a média de P_1. Temos as informações do coeficiente de variação e da soma dos quadrados de P_1. Então, podemos fazer:

$$CV_{P1} = \frac{\sigma}{\mu_{P_1}} = \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i)^2}{n} - \mu_{P_1}^2}}{\mu_{P_1}}$$

$$0,1 = \frac{\sqrt{\frac{323,2}{20} - \mu_{P_1}^2}}{\mu_{P_1}}$$

$$0,1\mu_{P_1} = \sqrt{\frac{323,2}{20} - \mu_{P_1}^2}$$

$$0,01\mu_{P_1}^2 = \frac{323,2}{20} - \mu_{P_1}^2$$

$$1,01\mu_{P_1}^2 = \frac{323,2}{20}$$

$$\mu_{P_1}^2 = \frac{323,2}{20 \times 1,01}$$

$$\mu_{P_1}^2 = \frac{323,2}{20 \times 1,01}$$

$$\mu_{P_1}^2 = \frac{323,2}{20,02}$$

$$\mu_{P_1}^2 \cong 16$$

$$\mu_{P_1} = \sqrt{16}$$

$$\mu_{P_1} = 4$$

Assim, a média de P_1 é 4 mil reais.

Agora, o enunciado afirma que a população P_2 possui 15 funcionários. Além disso, também diz que a soma dos quadrados é 375,6 e que a média de P_1 é igual a 80% da média de P_2. Queremos saber o desvio padrão de P_2. Aplicando a fórmula, temos:

$$\sigma_{P_2} = \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \mu_{P_2}\right)^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n}(x_i)^2}{n} - \mu_{P_2}^2}$$

$$\sigma_{P_2} = \sqrt{\frac{375,6}{15} - \left(\frac{\mu_{P_1}}{0,8}\right)^2}$$

$$\sigma_{P_2} = \sqrt{\frac{375,6}{15} - \left(\frac{4}{0,8}\right)^2}$$

$$\sigma_{P_2} = \sqrt{25,04 - 25}$$

$$\sigma_{P_2} = \sqrt{0,04}$$

$$\sigma_{P_2} = 0{,}2$$

Portanto, o desvio padrão de P_2 é 0,2 mil reais, ou seja, o desvio padrão para P_2 é de 200 reais.

Gabarito: A.

8. (VUNESP/EsFCEx/2020) A produção de n unidades de certa peça foi dividida em lotes iguais, com o mesmo número de unidades por lote. Para uma análise sobre peças defeituosas, o Controle de Qualidade da empresa examinou os 40 primeiros lotes produzidos, obtendo o seguinte número de peças defeituosas por lote:

Peças defeituosas	Frequência
0 |— 2	4
2 |— 4	10
4 |— 6	12
6 |— 8	8
8 |— 10	6

O desvio padrão dessa distribuição de frequência é igual a

a) 2,80 peças.

b) 3,20 peças.

c) 3,40 peças.

d) 2,40 peças.

e) 2,20 peças.

Comentários:

Para identificarmos a média, precisamos saber os pontos médios das classes e multiplicá-los pelas respectivas frequências, dividindo o somatório pelo número de observações:

$$\bar{x} = \frac{1 \times 4 + 3 \times 10 + 5 \times 12 + 7 \times 8 + 9 \times 6}{40}$$

$$\bar{x} = \frac{204}{40}$$

$$\bar{x} = 5{,}1$$

Agora, calculamos a variância:

$$\sigma^2 = \frac{4 \times (1 - 5{,}1)^2 + 10 \times (3 - 5{,}1)^2 + 12 \times (5 - 5{,}1)^2 + 8 \times (7 - 5{,}1)^2 + 6 \times (9 - 5{,}1)^2}{40}$$

$$\sigma^2 \cong 5{,}79$$

$$\sigma = \sqrt{5{,}79} \cong 2{,}4 \rightarrow Desvio\ padrão$$

Gabarito: D.

9. (VUNESP/EBSERH/2020) Um aluno tirou as seguintes notas ao longo do semestre: 4, 8, 6, 1 e 6. A média, a mediana e o desvio padrão foram, respectivamente:

a) 5; 6 e 5,6.

b) 5; 5 e 5,6.

c) 5; 6 e 2,4.

d) 5; 6 e 31,6.

e) 6; 6 e 5,6.

Comentários:

Vamos inicialmente colocar os dados em ordem crescente para identificar a mediana:

$$1, 4, 6, 6, 8$$

A mediana é o termo central da amostra. Logo:

$$M_d = 6$$

Calculando a média:

$$\bar{x} = \frac{4+8+6+1+6}{5} = \frac{25}{5}$$

$$\bar{x} = 5$$

Agora, calculamos o desvio padrão:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$\sigma = \sqrt{\frac{(4-5)^2 + (8-5)^2 + (6-5)^2 + (1-5)^2 + (6-5)^2}{5}}$$

$$\sigma = \sqrt{\frac{1+9+1+16+1}{5}}$$

$$\sigma = \sqrt{\frac{28}{5}}$$

$$\sigma \cong 2{,}4$$

Gabarito: C.

10. (CESPE/Pref. São Cristóvão/2019) A tabela seguinte mostra a distribuição das idades dos 30 alunos da turma A do quinto ano de uma escola de ensino fundamental.

Idade (em anos)	9	10	11	12	13	14
Quantidade de estudantes	6	22	0	1	0	1

A partir dessa tabela, julgue o item.

O desvio padrão das idades é inferior a 1 ano.

Comentários:

Calculando a média:

$$\bar{x} = \frac{9 \times 6 + 10 \times 22 + 11 \times 0 + 12 \times 1 + 13 \times 0 + 14 \times 1}{30}$$

$$\bar{x} = \frac{300}{30}$$

$$\bar{x} = 10$$

Agora, calcularemos os desvios de cada idade em relação à média:

Idades (x_i)	Quantidade de Estudantes (f_i)	Desvio em relação à média ($x_i - \bar{x}$)	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2 \times f_i$
9	6	$9 - 10 = -1$	1	6
10	22	$10 - 10 = 0$	0	0
11	0	$11 - 10 = 1$	1	0
12	1	$12 - 10 = 2$	4	4
13	0	$13 - 10 = 3$	9	0
14	1	$14 - 10 = 4$	16	16

O desvio padrão é calculado por meio da seguinte fórmula:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 \times f_i}{n}}$$

Aplicando a fórmula, temos:

$$\sigma = \sqrt{\frac{1 \times 6 + 10 \times 22 + 1 \times 0 + 4 \times 1 + 9 \times 0 + 16 \times 1}{30}}$$

Portanto, o desvio padrão é menor que 1:

$$\sigma = \sqrt{\frac{26}{30}} < 1.$$

Gabarito: Certo.

11. (FCC/BANRISUL/2019) Uma população é formada por 4 elementos, ou seja, {4, 5, 5, 8}. O coeficiente de variação, definido como o resultado da divisão do respectivo desvio padrão pela média aritmética da população, é igual a

a) $3/11$.

b) $9/22$.

c) $3/22$.

d) $9/11$.

e) $1/5$.

Comentários:

Primeiro, vamos calcular a média:

$$\bar{x} = \frac{4 + 5 + 5 + 8}{4}$$

$$\bar{x} = \frac{22}{4}$$

$$\bar{x} = 5{,}5$$

Devemos observar que as alternativas estão em forma de frações, então tomaremos para a média a simplificação de $\frac{22}{4}$. Logo:

$$\bar{x} = \frac{11}{2}$$

O desvio padrão da população é dado pela seguinte fórmula:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}$$

Assim, temos:

$$\sigma = \sqrt{\frac{(4-5{,}5)^2 + (5-5{,}5)^2 + (5-5{,}5)^2 + (8-5{,}5)^2}{4}} =$$

$$\sigma = \sqrt{\frac{9}{4}} = \frac{3}{2}.$$

O coeficiente de variação é dado por:

$$C_v = \frac{\sigma}{\bar{x}}$$

Então:

$$C_v = \frac{\frac{2}{3}}{\frac{11}{2}}$$

$$C_v = \frac{3}{2} \times \frac{2}{11} = \frac{6}{22}$$

$$C_v = \frac{3}{11}$$

Gabarito: A.

12. (FCC/Pref. Recife/2019) Uma população é formada pelos salários dos empregados de uma empresa. Decide-se dar um aumento de 10% sobre todos os salários mais um adicional fixo de R$ 500,00 para todos os salários. Com relação às medidas de tendência central e de dispersão é correto afirmar que a nova população formada terá

a) um desvio padrão igual ao desvio padrão da população anterior multiplicado por 1,10 acrescido de R$ 500,00.

b) uma variância igual à variância da população anterior multiplicada por 1,21 acrescida de 250.000 (R$)2.

c) uma média aritmética igual à média aritmética da população anterior acrescida de R$ 500,00.

d) uma mediana igual à mediana da população anterior acrescida de R$ 500,00.

e) um desvio padrão igual ao desvio padrão da população anterior multiplicado por 1,10 e uma variância igual à variância da população anterior multiplicada por 1,21.

Comentários:

Inicialmente, temos uma população (P_1) formada pelos salários antes do aumento. Depois, temos uma população que chamaremos de (P_2) para os salários depois do aumento. Assim, temos:

$$P_2 = P_1 \times 1{,}1 + 500$$

Logo, a população P_2 é formada pelos salários da população P_1 multiplicados por 1,1, que corresponde ao aumento de 10%, mais um valor fixo de 500 reais.

Analisando as alternativas temos:

- A alternativa A fala de desvio padrão. Sabemos que o desvio padrão não sofre influências de somas ou subtrações, logo, a alternativa está errada ao apontar a soma de R$ 500.
- A alternativa B fala da variância. Para acharmos a variância, precisamos elevar ao quadrado os dois lados da equação. Assim como o desvio padrão, a variância também não sofre influências de somas ou subtrações, logo, a multiplicação seria de fato por 1,21, porém sem o acréscimo de 250.000.
- A alternativa C fala da média, teríamos a seguinte equação: $\overline{P_2} = \overline{P_1} \times 1{,}1 + 500$. Assim, a alternativa erra ao omitir a multiplicação por 1,1.
- A alternativa D fala de mediana. Portanto, temos o mesmo raciocínio da média e o mesmo erro de omissão da multiplicação por 1,1 na alternativa.
- A alternativa E demonstra de forma correta o desvio padrão e a variância, conforme já visto nas alternativas anteriores.

Gabarito: E.

13. (FCC/Pref. Recife/2019) Considere uma população P formada por números estritamente positivos. Com relação às medidas de tendência central e de dispersão é correto afirmar que

a) multiplicando todos os elementos de P por 16, o desvio padrão da nova população é igual ao desvio padrão de P multiplicado por 4.

b) dividindo todos os elementos de P por 2, a variância da nova população é igual a variância de P multiplicada por 0,25.

c) adicionando uma constante K > 0 a todos os elementos de P, a média aritmética e a variância da nova população formada são iguais a média aritmética e desvio padrão de P, respectivamente.

d) a variância e o desvio padrão de P são iguais somente no caso em que todos os elementos de P são iguais.

e) subtraindo uma constante K > 0 de todos os elementos de P, o desvio padrão e a média aritmética da nova população são iguais ao desvio padrão e média aritmética de P subtraídos de K, respectivamente.

Comentários:

Analisando as alternativas temos:

- A alternativa A fala de desvio padrão. Assim, se todos os elementos de P fossem multiplicados por 16 o desvio padrão também seria multiplicado por 16. Alternativa errada.
- A alternativa B fala da variância. Sabemos que para a variância temos um efeito ao quadrado, logo, se dividirmos todos os elementos por 2, a variância será dividida por $2^2 = 4$. Dividir algo por 4 é o mesmo que multiplicar por 0,25. Logo, a alternativa está correta.

- A alternativa C fala da média e da variância. Para a média, temos que somada a uma constante c, a nova média também será somada dessa constante. Já para a variância e para o desvio padrão, a soma da constante não tem influência, portanto, a variância será sempre o quadrado do desvio padrão. Alternativa errada.
- A alternativa D erra ao afirmar que somente em um determinado caso a variância e o desvio padrão serão iguais. Além do caso mencionado na alternativa, temos o caso em que a variância vale 1 e desvio padrão vale $\sqrt{1} = 1$. Alternativa errada.
- A alternativa E erra ao afirmar que o desvio padrão sofrerá influência de subtração. Como vimos em itens anteriores, o desvio padrão não sofre influência de somas e subtrações. Alternativa errada.

Gabarito: B.

14. (CESPE/IFF/2018)

Foram feitas dez medidas do comprimento da caneta mostrada na figura. Os valores dessas medidas estão expressos na tabela a seguir.

medida	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
comprimento (mm)	136	135	135	137	134	135	136	135	136	135

Com base nessas informações, é correto afirmar que o valor do desvio padrão, em mm, desse experimento é igual a

a) 0,00.

b) 0,64.

c) 0,71.

d) 0,80.

e) 0,84.

Comentários:

Calculando a média:

$$\bar{x} = \frac{134 + 5 \times 135 + 3 \times 136 + 137}{10}$$

$$\bar{x} = \frac{1354}{10}$$

$$\bar{x} = 135{,}4$$

O desvio padrão amostral é dado pela fórmula apresentada a seguir. Repare que usamos $(n-1)$ no numerador pois se trata de uma amostra e não de toda a população:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 \times f_i}{n-1}}$$

Aplicando a fórmula, temos:

$$s = \sqrt{\frac{\sum_{i=1}^{10}(x_i - \bar{x})^2 \times f_i}{9}}$$

$$s = \sqrt{\frac{(134 - 135{,}4)^2 + 5 \times (135 - 135{,}4)^2 + 3 \times (136 - 135{,}4)^2 + (137 - 135{,}4)^2}{9}}$$

$$s = \sqrt{\frac{6{,}4}{9}}$$

$$s = 0{,}84$$

Gabarito: E.

15. (FCC/ARTESP/2017) O departamento de operações de uma autarquia do Estado fez um levantamento do número de acidentes em um determinado trecho de rodovia no ano de 2016, conforme tabela a seguir.

Mês	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Nº de Acidentes	36	28	12	5	3	2	2	4	9	11	22	38

Os números indicam que há uma dispersão significativa, portanto, o desvio padrão para esta amostra é representado por

a) 13,30.

b) 14,33.

c) 12,74.

d) 10,40.

e) 11,50.

Comentários:

Vamos iniciar calculando a média:

$$\bar{x} = \frac{36 + 28 + 12 + 5 + 3 + 2 + 2 + 4 + 9 + 11 + 22 + 38}{12}$$

$$\bar{x} = \frac{172}{12}$$

$$\bar{x} = \frac{43}{3} = 14{,}33$$

Agora, vamos montar uma tabela para simplificar o cálculo da média dos quadrados:

Valor (x)	X^2
36	1.296
28	784
12	144
5	25
3	9
2	4
2	4
4	16
9	81
11	121
22	484
38	1.444
Total	4.412

Portanto, a média dos quadrados é:

$$\overline{x^2} = \frac{4.412}{12} = 367{,}67$$

A variância populacional é dada pela diferença entre a média dos quadrados e o quadrado da média:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

$$\sigma^2 = 367{,}67 - (14{,}33)^2$$

$$\sigma^2 = 367{,}67 - 205{,}35$$

$$\sigma^2 = 162{,}32$$

Se multiplicarmos a variância populacional por $\frac{n}{n-1}$, encontraremos a variância amostral:

$$s^2 = 162{,}32 \times \frac{12}{11}$$

$$s^2 = 162{,}32 \times 1{,}09$$

$$s^2 = 177{,}07$$

$$s = \sqrt{177{,}07}$$

$$s = 13{,}30$$

Gabarito: A.

16. (FCC/TRT 20ª Região/2016) Em uma associação de determinada carreira profissional é realizado um censo em que foram apurados os salários de todos os seus 320 associados em número de salários mínimos (S.M.). O coeficiente de variação correspondente foi de 16% e a soma dos quadrados de todos os salários, em (S.M.)2, foi de 8.204,80. O desvio padrão dos salários destes associados é, em S.M., de

a) 0,80

b) 0,64

c) 0,96

d) 0,40

e) 1,60

Comentários:

O coeficiente de variação foi informado na questão. Sabemos que ele é resultado da divisão entre o desvio padrão e a média, então:

$$\frac{16}{100} = \frac{\sigma}{\bar{x}}$$

$$\bar{x} = \frac{100\sigma}{16}$$

A variância resulta da diferença entre a média dos quadrados e o quadrado da média. Vamos aplicar o valor da média na fórmula da variância:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

$$\sigma^2 = \frac{8.204{,}80}{320} - (\frac{100\sigma}{16})^2$$

$$\sigma^2 = \frac{8.204{,}80}{320} - \frac{10.000\sigma^2}{256}$$

$$\sigma^2 = \frac{32.819{,}2 - 50.000\sigma^2}{1280}$$

$$1280\sigma^2 = 32.819{,}2 - 50.000\sigma^2$$

$$51280\sigma^2 = 32.819{,}2$$

$$\sigma^2 = \frac{32.819,2}{51280}$$

$$\sigma^2 = 0,64$$

$$\sigma = \sqrt{0,64}$$

$$\sigma = 0,8$$

Gabarito: A.

17. (FGV/IBGE/2016) A partir de uma amostra de tamanho 2n+1, sendo n um número inteiro, elaborou-se a distribuição de frequência de tal forma que apenas os dados grupados ficaram disponíveis. Apesar disso, é possível determinar com certeza a classe à qual pertence o valor exato:

a) Da moda;

b) Da mediana;

c) Da média;

d) Dos quartis;

e) Do desvio padrão.

Comentários:

O enunciado diz que a amostra tem tamanho $2n + 1$. Sendo assim, a quantidade de termos será sempre ímpar. Nessas condições, a mediana obrigatoriamente ocupará o termo central da amostra, ficando posicionada em uma das classes de dados, e não entre duas classes limítrofes.

Gabarito: B.

18. (FCC/CNMP/2015) Em uma empresa, 55% dos empregados são do sexo masculino e a média aritmética dos salários de todos os empregados da empresa é igual a R$ 3.000,00. Sabe-se que a média aritmética dos salários dos empregados do sexo masculino é igual a média aritmética dos salários dos empregados do sexo feminino, sendo que os coeficientes de variação são iguais a 10% e 15%, respectivamente. O desvio padrão dos salários de todos os empregados da empresa é, em R$, de

a) 360,00.

b) 375,00.

c) 367,50.

d) 390,00.

e) 420,00.

Comentários:

Temos que a média para os dois conjuntos masculino e feminino é dada pela média ponderada dos conjuntos individuais. Como a média geral é igual a 3000, temos:

$$\bar{x} = 0{,}55 \times \bar{x}_M + 0{,}45 \times \bar{x}_F$$

$$\bar{x} = 3.000$$

Agora, queremos saber qual é o desvio padrão de cada grupo. Sabemos que o coeficiente de variação é dado pela divisão entre o desvio padrão e a média. O enunciado já nos forneceu o coeficiente de variação para os homens e para as mulheres. Assim, temos:

$$C_v = \frac{\sigma}{\bar{x}}$$

Para o sexo masculino:

$$C_{vM} = \frac{\sigma_M}{\bar{x}_M}$$

$$0{,}1 = \frac{\sigma_M}{3000}$$

$$\sigma_M = 3000 \times 0{,}1$$

$$\sigma_M = 300$$

Para o sexo feminino:

$$C_{vF} = \frac{\sigma_F}{\bar{x}_F}$$

$$0{,}15 = \frac{\sigma_F}{3000}$$

$$\sigma_F = 3000 \times 0{,}15$$

$$\sigma_F = 450$$

Para a variância total de todos os empregados, precisamos calcular a variância ponderada para os sexos masculino e feminino:

$$\sigma^2 = \sigma_M^2 \times 0{,}55 + \sigma_F^2 \times 0{,}45$$

$$\sigma^2 = 300^2 \times 0{,}55 + 450^2 \times 0{,}45$$

$$\sigma^2 = 49.500 + 91.125$$

$$\sigma^2 = 140.625$$

Para o desvio padrão de todos os empregados, basta tirarmos a raiz quadrada:

$$\sigma = \sqrt{140.625}$$

$$\sigma = 375$$

Gabarito: B.

19. (FCC/MPE AM/2013) Seja x_1, x_2,..., x_9 uma amostra de 9 observações da variável X. Sabe-se que:

$$\sum_1^9 X_i = 72 \sum_1^9 X_i^2 = 576,5$$

Nessas condições, o desvio padrão dessa amostra é

a) 0,50

b) 0,40

c) 0,15

d) 0,80

e) 0,25

Comentários:

Vamos iniciar com os dados da questão:

- a amostra tem tamanho n = 9;
- o somatório dos elementos é $\sum_1^9 \boldsymbol{X_i} = \mathbf{72}$;
- o somatório dos quadrados é $\sum_1^9 \boldsymbol{X_i^2} = \mathbf{576,5.}$

Com esses dados, podemos calcular a variância populacional, que é dada pela diferença entre a média dos quadrados e o quadrado da média:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

Perceba que $\overline{x^2}$ será igual a $\frac{576,5}{9} = 64{,}0555$ e que $\bar{x}^2$ será igual a $\left(\frac{72}{9}\right)^2 = 8^2 = 64$. Assim, temos:

$$\sigma^2 = 64{,}0555 - 64$$

$$\sigma^2 = 0{,}0555$$

Se multiplicarmos a variância populacional por $\frac{n}{n-1}$, encontraremos a variância amostral:

$$s^2 = 0{,}0555 \times \frac{9}{9-1}$$

$$s^2 = 0{,}0555 \times 1{,}125$$

$$s^2 \cong 0{,}0625$$

Para acharmos o desvio padrão amostral, basta calcularmos a raiz quadrada:

$$s = \sqrt{0{,}0625}$$

$$s = 0{,}25$$

Gabarito: E.

20. (FCC/SERGAS/2013) A tabela abaixo apresenta a distribuição de frequências relativas dos salários, em número de salários mínimos (S.M.), dos 100 funcionários de uma empresa

Classes de salários (em S.M)	Frequências relativas
1 |— 3	0,3
3 |— 5	0,4
5 |— 7	0,3

O valor do desvio padrão desses 100 funcionários, considerado como desvio padrão populacional e obtido por meio dessa tabela, calculado como se todos os valores de cada classe de salários coincidissem com o ponto médio da referida classe, em número de S.M., é

a) $\sqrt{1{,}2}$

b) $\sqrt{2{,}2}$

c) $\sqrt{2}$

d) $\sqrt{1{,}8}$

e) $\sqrt{2{,}4}$

Comentários:

Vamos inicialmente reescrever a tabela acrescentando as informações dos pontos médios e multiplicações pelas respectivas frequências:

Classes	Pontos médios (x)	(x^2)	Frequência (f)	$x \times f$	$x^2 \times f$
1 |— 3	2	4	0,3	0,6	1,2
3 |— 5	4	16	0,4	1,6	6,4
5 |— 7	6	36	0,3	1,8	10,8
Total				4	18,4

Sendo assim, encontramos os valores da média e da média dos quadrados:

$$\bar{x} = 4$$

$$\overline{x^2} = 18{,}4$$

Agora, já podemos calcular a variância, que é determinada por:

$$\sigma^2 = 18{,}4 - 4^2$$

$$\sigma^2 = 18{,}4 - 16$$

$$\sigma^2 = 2{,}4$$

Para acharmos o desvio padrão, basta calcularmos a raiz quadrada:

$$\sigma = \sqrt{2{,}4}$$

Gabarito: E.

21. (FGV/TJ-AM/2013) Em relação à medida de desvio padrão, analise as afirmativas a seguir.

I. Tal medida apresenta a propriedade adimensional.

II. Tal medida mostra a dispersão dos dados em relação à média.

III. Tal medida nunca é negativa.

Assinale:

a) Se apenas a afirmativa II estiver correta.

b) Se apenas as afirmativas I e III estiverem corretas.

c) Se apenas as afirmativas II e III estiverem corretas.

d) Se apenas as afirmativas I e II estiverem corretas.

e) Se todas as afirmativas estiverem corretas.

Comentários:

Vamos analisar cada um dos itens:

- item I: falso, pois o desvio padrão tem a mesma unidade de medida que a variável apresentada;
- item II: verdadeiro. O desvio padrão mede a dispersão dos dados, tomando a média como referência;
- item III: verdadeiro. É a raiz quadrada da variância e, portanto, sempre é positiva.

Gabarito: C.

22. (FGV/SEFAZ-RJ/2011) O desvio-padrão da população {2; 4; 2; 4; 2; 4; 2; 4} é

a) 1,5 .

b) 1,0 .

c) 2,5 .

d) 2,0 .

e) 3,0 .

Comentários:

Vamos inicialmente calcular a média:

$$\bar{x} = \frac{2+4+2+4+2+4+2+4}{8}$$

$$\bar{x} = \frac{24}{8}$$

$$\bar{x} = 3$$

Agora, vamos calcular a média dos quadrados:

$$\overline{x^2} = \frac{2^2+4^2+2^2+4^2+2^2+4^2+2^2+4^2}{8}$$

$$\overline{x^2} = \frac{80}{8}$$

$$\overline{x^2} = 10$$

Sabemos que a variância é dada pela diferença entre a média dos quadrados e o quadrado da média, logo:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

$$\sigma^2 = 10 - 3^2 = 1$$

O desvio padrão é dado pela raiz quadrada da variância:

$$\sigma = \sqrt{1}$$

$$\sigma = 1$$

Gabarito: B.

23. (CESGRANRIO/BACEN/2009) Considere as informações a seguir para responder à questão.

A viabilidade financeira do projeto de uma microempresa leva em consideração dados históricos de 100 projetos semelhantes.

A tabela abaixo mostra a distribuição de frequências do VPL - Valor Presente Líquido (valores em milhões de reais) de um conjunto de microempresas similares.

VPL	Frequência relativa
$-10 < x \leq 0$	10%
$0 < x \leq 10$	80%

$\mathbf{10 < x \leq 20}$	10%

Segundo os dados históricos, o valor, em milhões de reais, que mais se aproxima do desvio padrão do VPL da microempresa é

a) 1

b) 2

c) 2,5

d) 4

e) 4,5

Comentários:

Queremos calcular o desvio padrão. Vamos reescrever a tabela colocando os pontos médios e multiplicado as frequências:

VPL	Pontos médios (x_i)	x_i^2	Frequência (f)	$x_i^2 \times f$
$\mathbf{-10 < x \leq 0}$	-5	25	0,1	2,5
$\mathbf{0 < x \leq 10}$	5	25	0,8	20
$\mathbf{10 < x \leq 20}$	15	225	0,1	22,5
	Total		1	45

Vamos iniciar a questão calculando a média. Na tabela percebemos uma distribuição simétrica, portanto a média será o ponto médio da classe central. Assim temos que:

$$\bar{x} = 5$$

$$\overline{x^2} = \frac{45}{1} = 45$$

A variância é dada por:

$$Var = \overline{x^2} - \bar{x}^2$$

$$Var = 45 - 5^2$$

$$Var = 45 - 25$$

$$Var = 20$$

O desvio-padrão será:

$$\sigma = \sqrt{Var}$$

$$\sigma = \sqrt{20}$$

$$\sigma \cong 4{,}47$$

Gabarito: E.

24. (CESGRANRIO/BNDES/2007) O enunciado abaixo refere-se à questão.

Um grupo é formado por 10 pessoas, cujas idades são:

17 19 19 20 20 20 20 21 22 22

Seja μ a média aritmética das idades e σ seu desvio padrão. O número de pessoas desse grupo cujas idades pertencem ao intervalo $[\mu - \sigma, \mu + \sigma]$ é

(Considere $\sqrt{2} = 1{,}4$)

a) 9

b) 8

c) 7

d) 6

e) 5

Comentários:

Vamos iniciar calculando a média:

$$\bar{x} = \frac{17 + 19 + 19 + 20 + 20 + 20 + 20 + 21 + 22 + 22}{10}$$

$$\bar{x} = \frac{200}{10}$$

$$\bar{x} = 20$$

Agora, vamos calcular os desvios de cada nota em relação à média:

(x_i)	17	19	19	20	20	20	20	21	22	22	**Total**
$(x_i - \bar{x})$	-3	-1	-1	0	0	0	0	1	2	2	
$(x_i - \bar{x})^2$	9	1	1	0	0	0	0	1	4	4	**20**

Pronto, já podemos calcular a variância:

$$\sigma^2 = \frac{20}{10} = 2$$

Calculando o desvio padrão:

$$\sigma = \sqrt{2} = 1{,}4$$

Feito isso, podemos aplicar o intervalo dado no enunciado:

$$\bar{x} - \sigma = 20 - 1{,}4 = 18{,}6$$

$$\bar{x} + \sigma = 20 + 1{,}4 = 21{,}4$$

Assim, temos nesse intervalo:

19 19 20 20 20 20 21

Gabarito: C.

QUESTÕES COMENTADAS

Coeficiente de Variação (ou Dispersão Relativa)

1. (CESPE/TELEBRAS/2022) Com respeito ao conjunto de dados {0, 0, 1, 1, 1, 3}, julgue o item que se segue.

O coeficiente de variação é igual ou superior a 1,2.

Comentários:

O coeficiente de variação é uma medida que fornece a variação dos dados em relação à média, sendo calculado por meio da expressão:

$$CV = \frac{\sigma}{\mu} \times 100\ (\%)$$

Portanto, vamos calcular a média do conjunto:

$$\mu = \frac{0+0+1+1+1+3}{6} = \frac{6}{6} = 1$$

Sabemos que o desvio padrão da população é dado pela seguinte fórmula:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}}$$

$$\sigma = \sqrt{\frac{(0-1)^2 + (0-1)^2 + (1-1)^2 + (1-1)^2 + (1-1)^2 + (3-1)^2}{6}}$$

$$\sigma = \sqrt{\frac{1+1+0+0+0+4}{6}} = \sqrt{\frac{6}{6}} = \sqrt{1} = 1$$

Substituindo, podemos calcular o coeficiente de variação:

$$CV = \frac{\sigma}{\mu} \times 100\ (\%)$$

$$CV = \frac{1}{1} = 1$$

Gabarito: Errado.

2. (CESPE/TELEBRAS/2022 - ADAPTADA) Com respeito ao conjunto de dados {5a, 2a, 2a}, em que a representa uma constante não nula, julgue o próximo item.

O coeficiente de variação independe do valor da constante a.

Comentários:

Para a média, temos:

$$\mu = \frac{5a + 2a + 2a}{3} = \frac{9a}{3} = 3a$$

Para o desvio padrão, temos:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}$$

$$\sigma^2 = \frac{(5a - 3a)^2 + (2a - 3a)^2 + (2a - 3a)^2}{3}$$

$$\sigma^2 = \frac{(2a)^2 + (-a)^2 + (-a)^2}{3}$$

$$\sigma^2 = \frac{4a^2 + a^2 + a^2}{3} = \frac{6a^2}{3} = 2a^2$$

$$\sigma = \sqrt{2}|a|$$

Substituindo, podemos calcular o coeficiente de variação:

$$CV = \frac{\sigma}{\mu} \times 100\ (\%)$$

Agora, devemos ficar atentos ao fato de que se $a < 1$, teremos:

$$CV = \frac{\sqrt{2}|a|}{-3a} = -\frac{\sqrt{2}}{3}$$

Por outro lado, se $a > 1$, teremos:

$$CV = \frac{\sqrt{2}|a|}{3a} = \frac{\sqrt{2}}{3}$$

A questão original deu o gabarito como correto.

Gabarito: Errado.

3. (FGV/PC-AM/2022) Suponha que um pesquisador tenha as seguintes informações de uma amostra de dados:

Média = 5

Variância = 25

Soma dos desvios absolutos em relação à média = 10

Tamanho da amostra = 5

Assim, o coeficiente de variação dessa amostra em termos decimais será igual a

a) 1.

b) 2.

c) $\sqrt{5}$.

d) 5.

e) 10.

Comentários:

O coeficiente de variação é dado pela razão entre o desvio padrão e a média. Sabemos que o desvio padrão é a raiz quadrada da variância. Portanto, temos:

$$C_v = \frac{\sigma}{\bar{x}}$$

$$C_v = \frac{\sqrt{25}}{5} = \frac{5}{5} = 1$$

Gabarito: A.

4. (CESGRANRIO/BB/2021) Um pesquisador recebeu os dados de uma amostra de tamanho 100 de uma população e calculou a média amostral μ, o desvio padrão amostral σ e o coeficiente de variação amostral $CV = \frac{\sigma}{\mu}$. Antes de iniciar a análise, ele foi informado de que os dados dessa amostra estavam todos errados, mas que podiam ser corrigidos somando-se 3 a cada um dos dados que recebeu.

Após fazer tal correção, o valor do coeficiente de variação amostral passou a ser

a) $\frac{3\sigma}{\mu+3}$

b) $\frac{300\sigma}{\mu+300}$

c) $\frac{\sigma}{\mu+3}$

d) $\frac{\sigma}{\mu+300}$

e) $\frac{\sigma}{\mu+0,03}$

Comentários:

A média, ao ser somada uma constante k, a nova média também será somada dessa constante, portanto, $\mu + 3$. Já para a variância e desvio padrão, a soma da constante não tem influência. Assim, o valor do coeficiente de variação amostral passará a ser:

$$CV = \frac{\sigma}{\mu + 3}$$

Gabarito: C.

5. (CESPE/SERPRO/2021) Considerando que o número X de erros registrados em determinado tipo de código computacional siga uma distribuição binomial com média igual a 4 e variância igual a 3, julgue o item a seguir.

O coeficiente de variação da distribuição de erros X é igual a 3.

Comentários:

O coeficiente de variação é dado pela razão entre o desvio padrão e a média. Sabemos que o desvio padrão é a raiz quadrada da variância, portanto, temos:

$$C_v = \frac{\sigma}{\bar{x}}$$

$$C_v = \frac{\sqrt{3}}{4} \cong 0{,}433$$

Gabarito: Errado.

6. (VUNESP/EsFCEx/2020) Sejam P_1 e P_2 duas populações independentes formadas por números estritamente positivos com tamanhos 20 e 25, respectivamente. O coeficiente de variação de P_1 é igual a 50% com a soma dos quadrados de seus elementos igual a 2.500. Sabe-se que a soma dos quadrados dos elementos de P_2 é igual a 2.900 e a média aritmética é igual a média aritmética de P_1. O coeficiente de variação de P_2 é igual a

a) 0,40.

b) 0,20.

c) 0,60.

d) 0,64.

e) 0,80.

Comentários:

Iniciaremos o cálculo pela população P_1. Considerando os dados do problema, o desvio padrão de P_1 é:

$$\sigma_1 = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}$$

$$\sigma_1 = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}{n}}$$

$$\sigma_1 = \sqrt{\frac{\sum_{i=1}^{20} x_i^2 - 20\bar{x}_1^2}{20}}$$

$$\sigma_1 = \sqrt{\frac{2500 - 20\bar{x}_1^2}{20}}$$

$$\sigma_1 = \sqrt{125 - \bar{x}_1^2}$$

Aplicando a fórmula do coeficiente de variação para P_1, sendo $C_{v_1} = 50\%$:

$$C_{v_1} = \frac{\sigma_1}{\overline{x_1}}$$

$$0{,}5 = \frac{\sqrt{125 - \bar{x}_1^2}}{\overline{x_1}}$$

$$0{,}5\overline{x_1} = \sqrt{125 - \bar{x}_1^2}$$

$$0{,}5^2\bar{x}_1^2 = 125 - \bar{x}_1^2$$

$$0{,}25\bar{x}_1^2 = 125 - \bar{x}_1^2$$

$$1{,}25\bar{x}_1^2 = 125$$

$$\bar{x}_1^2 = \frac{125}{1{,}25}$$

$$\bar{x}_1^2 = 100 \rightarrow \bar{x}_1 = \sqrt{100}$$

$$\bar{x}_1 = 10$$

Na população P_2, temos que a soma dos quadrados é 2900 e a média é igual a de P_1, então:

$$\sigma_2 = \sqrt{\frac{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}{n}}$$

$$\sigma_2 = \sqrt{\frac{\sum_{i=1}^{25} x_i^2 - 25\bar{x}_2^2}{25}}$$

$$\sigma_2 = \sqrt{\frac{2900 - 25\bar{x}_2^2}{25}}$$

$$\sigma_2 = \sqrt{\frac{2900 - 25 \times 10^2}{25}}$$

$$\sigma_2 = \sqrt{116 - 100}$$

$$\sigma_2 = \sqrt{16}$$

$$\sigma_2 = 4$$

Agora, podemos calcular o coeficiente de variação de P_1:

$$C_{v_2} = \frac{\sigma_2}{\overline{x_2}}$$

$$C_{v_2} = \frac{4}{10}$$

$$C_{v_2} = 0{,}4$$

Gabarito: A.

7. (CESGRANRIO/BB/2018) Há dez anos a média das idades, em anos completos, de um grupo de 526 pessoas era de 30 anos, com desvio padrão de 8 anos.

Considerando-se que todas as pessoas desse grupo estão vivas, o quociente entre o desvio padrão e a média das idades, em anos completos, hoje, é

a) 0,45

b) 0,42

c) 0,20

d) 0,27

e) 0,34

Comentários:

Temos no enunciado que a média há 10 anos era 30 anos para um grupo de 526 pessoas. Então, vamos calcular a soma das idades:

$$30 = \frac{x_1 + x_2 + \cdots + x_{526}}{526}$$

$$x_1 + x_2 + \cdots + x_{526} = 526 \times 30$$

$$x_1 + x_2 + \cdots + x_{526} = 15.780$$

Agora, vamos calcular a média atual das idades, considerando que cada pessoa envelheceu 10 anos:

$$\bar{x} = \frac{15.780 + (526 \times 10)}{526}$$

$$\bar{x} = \frac{15.780 + 5.260}{526}$$

$$\bar{x} = \frac{21.040}{526}$$

$$\bar{x} = 40$$

Utilizando as propriedades do desvio padrão, sabemos que essa medida não sofre alteração com o acréscimo de uma constante, como é o caso. Assim, não houve alteração do desvio padrão em relação à media anterior. Desta forma, o quociente entre o desvio padrão e a média das idades atual é:

$$\frac{8}{40} = 0{,}2$$

Gabarito: C.

8. (CESGRANRIO/LIQUIGÁS/2018) Uma amostra $(x_1, x_2, ..., x_{10})$ de tamanho 10 de uma população nos forneceu os seguintes valores

$$\sum_{i=1}^{10} x_i = 20 \; e \; \sum_{i=1}^{10} x_i^2 = 140.$$

Qual o valor do coeficiente de variação amostral?

a) $\frac{7}{20}$

b) $\frac{5}{3}$

c) $\frac{1}{7}$

d) $\frac{1}{6}$

e) 6

Comentários:

O coeficiente de variação é determinado pela divisão entre o desvio padrão e a média aritmética. Por sua vez, a variância populacional (σ^2) é dada pela diferença entre a média dos quadrados e o quadrado da média. Assim, vamos iniciar a resolução calculando o quadrado da média:

$$\bar{x} = \frac{\sum_{i=1}^{10} x_i}{10}$$

$$\bar{x} = \frac{20}{10}$$

$$\bar{x} = 2$$

$$\bar{x}^2 = 2^2 = 4$$

Vamos calcular a média dos quadrados:

$$\overline{x^2} = \frac{\sum_{i=1}^{10} x_i^2}{10}$$

$$\overline{x^2} = \frac{140}{10}$$

$$\overline{x^2} = 14$$

Agora, já podemos calcular a variância populacional:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

$$\sigma^2 = 14 - 4$$

$$\sigma^2 = 10$$

Temos a variância populacional, mas precisamos calcular a variância amostral, logo:

$$s^2 = \sigma^2 \times \frac{n}{n-1}$$

$$s^2 = 10 \times \frac{10}{10-1}$$

$$s^2 = 10 \times \frac{10}{9}$$

$$s^2 = \frac{100}{9}$$

$$s = \sqrt{\frac{100}{9}}$$

Assim, já podemos calcular o coeficiente de variação amostral:

$$CV = \frac{s}{\bar{x}} = \frac{\sqrt{\frac{100}{9}}}{2} = \frac{\frac{10}{3}}{2} = \frac{5}{3}$$

Gabarito: B.

9. (CESPE/IPHAN/2018) Cinco municípios de um estado brasileiro possuem as seguintes quantidades de patrimônios históricos: {2, 3, 5, 3, 2}. Admitindo que a média e o desvio-padrão desse conjunto de valores sejam iguais a 3 e 1,2, respectivamente, julgue o item seguinte.

O coeficiente de variação é superior a 0,3 e inferior a 0,5.

Comentários:

O coeficiente de variação é dado pela razão entre o desvio padrão e a média. Portanto, temos:

$$C_v = \frac{\sigma}{\bar{x}}$$

$$C_v = \frac{12}{3} = 0{,}4$$

Gabarito: Certo.

10. (FCC/TCE-RS/2018) Uma população é formada por 100 números estritamente positivos x_i com $1 \leq i \leq 100$, ou seja, $\{x_1, x_2, x_3, \ldots, x_{100}\}$, em que x_i representa a renda familiar anual da família i, em milhares de reais.

Dados:

$\sum_{i=1}^{100} x_i = 6.400$ **mil reais e**

$\sum_{i=1}^{100} x_i^2 = 467.200$ **(mil reais)²**

O coeficiente de variação desta população é igual a

a) 37,5%

b) 18,0%

c) 32,5%

d) 24,0%

e) 27,5%

Comentários:

Vamos iniciar a resolução calculando a média. Temos que o total das observações é 6400 e temos 100 observações. Assim, para encontrarmos a média, basta dividirmos 6400 por 100:

$$\bar{x} = \frac{6400}{100} = 64$$

A questão também nos forneceu o valor total de x_i^2, podemos calcular a média dos quadrados:

$$\overline{x^2} = \frac{467200}{100} = 4672$$

Agora, podemos encontrar o valor da variância, que é dada pela diferença entre a média dos quadrados e o quadrado da média:

$$s^2 = \overline{x^2} - \bar{x}^2$$

$$s^2 = 4672 - 64^2$$

$$s^2 = 4672 - 4096$$

$$s^2 = 576$$

O desvio padrão é a raiz quadrada da variância, logo:

$$s = \sqrt{576}$$

$$s = 24$$

Pronto, já podemos encontrar o coeficiente de variação, que é dado pela divisão entre o desvio padrão e a média:

$$C_v = \frac{s}{\bar{x}}$$

$$C_v = \frac{24}{64} = 0{,}375 \rightarrow 37{,}5\%$$

Gabarito: A.

11. (FGV/IBGE/2017) Alguns economistas estão discutindo sobre a volatilidade dos preços em duas economias, relativamente parecidas, tendo como moedas peras (A) e maçãs (B). Sabe-se que as médias dos preços são 100 peras e 120 maçãs, respectivamente. É fornecido, ainda, o desvio-padrão dos preços em A, igual a 25 peras, e a variância em B, igual a 400 maçãs ao quadrado.

Considerando as principais medidas estatísticas de dispersão como medidas de volatilidade, é correto afirmar que:

a) O desvio padrão dos preços em A é inferior ao de B;

b) A taxa de conversão da moeda A para B é de 1,2;

c) A taxa de inflação em A deve ser menor do que em B;

d) Os preços em B são, em média, mais caros do que em A;

e) A medida adimensional de dispersão de A é superior à de B.

Comentários:

Vamos analisar cada uma das alternativas:

- letra A: o desvio padrão é igual à raiz quadrada da variância, logo:

$$s = \sqrt{s^2} \rightarrow \sqrt{400} = 20$$

- letra B: não podemos afirmar qual a taxa de conversão das moedas, pois o problema não traz informações suficientes para isso;
- letra C: não podemos afirmar nada sobre a taxa de inflação, pois o problema não traz informações suficientes para isso;
- letra D: não podemos afirmar nada sobre os preços das moedas, pois o problema não traz informações suficientes para isso; e
- letra E: a medida adimensional se refere ao coeficiente de variação $C_v = \frac{\sigma}{\bar{x}}$. Logo:

$$C_{v_A} = \frac{25}{100} = 0{,}25$$

$$C_{v_B} = \frac{20}{120} \cong 0{,}16$$

Gabarito: E.

12. (FGV/IBGE/2016) As principais medidas de dispersão utilizadas na estatística são a amplitude (A), a variância (Var), o desvio padrão (DP), o coeficiente de variação (CV) e o desvio-interquartílico (DI). Sobre o tema, é correto afirmar que:

a) As medidas acima listadas têm seus valores dependentes, na íntegra, dos valores da distribuição amostral;

b) A variância apresenta a vantagem de ser diretamente comparável com os valores da distribuição;

c) É possível afirmar que var(x) ≥ dp(x);

d) O desvio-interquartílico é sempre superior ou no mínimo igual à amplitude;

e) O coeficiente de variação é uma medida invariante às mudanças de escala.

Comentários:

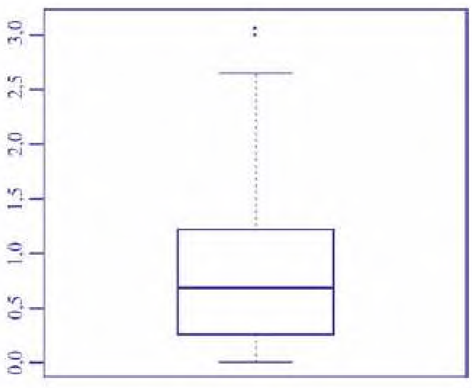
Vamos analisar cada uma das alternativas:

- letra A: a alternativa está incorreta, pois as medidas de dispersão podem ser calculadas tanto para amostras quanto para populações;
- letra B: o desvio padrão tem a mesma unidade de medida que a variável observada, enquanto a variância corresponde ao desvio padrão ao quadrado;
- letra C: a variância será menor que o desvio padrão for um valor entre 0 e 1, isto é, $0 < \sigma < 1$;
- letra D: a amplitude é a diferença entre o valor máximo e o mínimo, enquanto o desvio interquartílico é dado pela diferença entre o terceiro e o primeiro quartis; e
- letra E: o coeficiente de variação é dado pela divisão entre o desvio padrão e a média. Assim, se multiplicarmos ou dividirmos o numerador e denominador por uma mesma constante o coeficiente de variação não se altera.

Com essas informações, concluímos que a alternativa correta é a E.

Gabarito: E.

13. (CESPE/TCE-PA/2016)



média amostral	0,80
desvio padrão amostral	0,70
primeiro quartil	0,25
mediana	0,70
terceiro quartil	1,20
mínimo	0
máximo	3,10

Um indicador de desempenho X permite avaliar a qualidade dos processos de governança de instituições públicas. A figura mostra, esquematicamente, a sua distribuição, obtida mediante estudo amostral feito por determinada agência de pesquisa. A tabela apresenta estatísticas descritivas referentes a essa distribuição.

Com base nessas informações, julgue o item a seguir.

O coeficiente de variação da distribuição de X é inferior a 0,8.

Comentários:

O coeficiente de variação é dado pela razão entre o desvio padrão e a média. Portanto, temos:

$$C_v = \frac{\sigma}{\bar{x}}$$

$$C_v = \frac{0,7}{0,8} = 0,875$$

Gabarito: Errado.

14. (FCC/TRE RR/2015) Em uma escola é realizado um censo apurando-se as alturas de todos os 180 estudantes em centímetros (cm). A média aritmética das alturas dos 100 estudantes do sexo masculino foi igual a dos 80 estudantes do sexo feminino. Se X_i representa a altura do i-ésimo estudante do sexo masculino e Y_j a altura do j-ésimo estudante do sexo feminino, obteve-se

$$\sum_{i=1}^{100} x_1^2 = 2.570.000\ cm^2 \ \ e \ \sum_{j=1}^{80} y_j^2 = 2.084.080\ cm^2$$

Se o desvio padrão das alturas dos estudantes do sexo masculino foi igual a 10 cm, o coeficiente de variação considerando todos os estudantes desta escola é, em %, de

a) 18.

b) 12.

c) 10.

d) 16.

e) 15.

Comentários:

A questão já nos forneceu o desvio padrão das alturas dos homens e vale 10. Sabemos que a variância corresponde ao quadrado do desvio padrão, logo a variância dos homens X_i vale:

$$\sigma_X^2 = 100$$

Agora, vamos calcular a média dos quadrados. Temos um total de 100 homens e temos que $\boldsymbol{x_1^2}$ vale 2.570.000, então a média dos quadrados fica:

$$\overline{x^2} = \frac{2.570.000}{100} = 25.700$$

Também queremos saber qual é a média para os homens. Ora, sabemos que a variância é dada pela diferença entre a média dos quadrados e o quadrado da média, assim temos que:

$$\sigma_X^2 = \overline{x^2} - \bar{x}^2$$

substituindo os valores temos:

$$100 = 25.700 - \bar{x}^2$$

$$\bar{x}^2 = 25.700 - 100$$

$$\bar{x}^2 = 25.600$$

$$\bar{x} = \sqrt{25600}$$

$$\bar{x} = 160$$

Portanto, a média para os homens vale 160.

Agora, vamos aos valores para as mulheres. Iniciando com a média dos quadrados, temos um total de 80 mulheres e temos que $\boldsymbol{y_1^2}$ vale 2.084.080, então, a média dos quadrados fica:

$$\overline{y^2} = \frac{2.084.080}{80} = 26.051$$

O enunciado disse que a média para as mulheres é igual a média para os homens, logo $\bar{y} = 160$.

Dessa forma, calculando a variância, temos:

$$\sigma_Y^2 = \overline{y^2} - \bar{y}^2$$

$$\sigma_Y^2 = 26051 - 160^2$$

$$\sigma_Y^2 = 26051 - 25600$$

$$\sigma_Y^2 = 451$$

Temos que a média dos dois conjuntos são iguais, então a variância da união dos dois conjuntos será dada pela média ponderada das variâncias dos conjuntos individuais:

$$\sigma^2 = \frac{\sigma_X^2 \times 100 + \sigma_Y^2 \times 80}{100 + 80}$$

$$\sigma^2 = \frac{100 \times 100 + 451 \times 80}{180}$$

$$\sigma^2 = \frac{10.000 + 36.080}{180}$$

$$\sigma^2 = \frac{46.080}{180}$$

$$\sigma^2 = 256$$

Para o desvio padrão, basta tirarmos a raiz quadrada:

$$\sigma = \sqrt{256}$$

$$\sigma = 16$$

Agora podemos calcular o coeficiente de variação, que é dado pela divisão entre o desvio padrão e a média:

$$C_v = \frac{\sigma}{\bar{x}}$$

Temos que a média para os dois conjuntos masculino e feminino é dada pela média ponderada dos conjuntos individuais, como a média é igual para os públicos masculino e feminino, a média geral também será igual, ou seja, vale 160. Assim, temos:

$$C_v = \frac{16}{160}$$

$$C_v = 0{,}1 \rightarrow 10\%$$

Gabarito: C.

15. (VUNESP/TJ-SP/2015) Leia o texto a seguir para responder à questão.

Na tabela a seguir, são apresentados os dados colhidos de uma amostra de 9 vendedores de um grande magazine, para comparação entre o tempo T de experiência do vendedor (em anos) e seu movimento mensal V de vendas (em mil). Na tabela, estão registrados também a média mensal de tempo de experiência, a média mensal de vendas, o desvio padrão (DP) de cada variável, além do coeficiente de correlação r da relação entre as variáveis T e V.

FUNC	T (Anos)	V (vendas)
A	4,2	12,5
B	6,1	15,5
C	4	12,9
D	5,4	14,2
E	1,2	8,2
F	1,5	9,1
G	3,2	9,9
H	1	6,4
I	5	12,1
MÉDIA	**3,51**	**11,20**
DP	**1,90**	**2,98**
r	**0,96**	

O valor mais próximo do coeficiente de variação da variável T é

a) 0,28.

b) 0,34.

c) 0,48.

d) 0,54.

e) 0,96

Comentários:

O coeficiente de variação é dado pela divisão do desvio padrão e a média:

$$C_v = \frac{\sigma}{\bar{x}}$$

$$C_v = \frac{1,90}{3,51} \cong 0,54$$

Gabarito: D.

16. (VUNESP/Pref. SP/2015) O indicador da variabilidade de uma determinada distribuição de um conjunto de dados é denominado

a) Mediana.

b) Moda.

c) Coeficiente de variação.

d) Quociente de correlação.

e) Coeficiente de determinação.

Comentários:

As medidas de tendência central estudam o centro da amostra. São dadas pela média, mediana e moda.

As medidas de dispersão têm a finalidade de identificar o quanto os dados estão dispersos em torno da média de uma amostra. Também são indicativos de variabilidade da amostra. São coeficientes de variação, desvio padrão, amplitude e variância.

Gabarito: C.

17. (CESGRANRIO/EPE/2014) Um pesquisador está interessado em comparar a variabilidade de duas variáveis com médias diferentes e desvios padrões diferentes, presentes num dado estudo estatístico.

A medida estatística adequada a ser usada nesse contexto é a(o)

a) covariância

b) diferença entre a maior e a menor variância

c) razão entre o maior e o menor desvios padrões

d) coeficiente de variação

e) coeficiente de correlação

Comentários:

Das medidas estatísticas, a que mais se adequa aos critérios do enunciado é o **coeficiente de variação**, definido pela razão entre o **desvio padrão** e a **média** da amostra, sendo uma medida adimensional. O coeficiente de variação é usado para expressar o grau de variabilidade dos dados estatísticos, desconsiderando a influência da ordem de grandeza da variável.

Gabarito: D.

18. (CESGRANRIO/IBGE/2014) Os dados a seguir foram obtidos de empregados de uma empresa com três fábricas: I, II e III. A variável de interesse é salário.

Empresa	Média	Desvio padrão
Fábrica I	1185	630,49
Fábrica II	600	355,97
Fábrica III	2150	1106,16

Comparando-se a variabilidade de salários em relação ao salário médio das três fábricas, através de seus coeficientes de variação, conclui-se que a variabilidade da fábrica

a) I é menor apenas do que a da fábrica III.

b) II é menor apenas do que a da fábrica I.

c) II é menor apenas do que a da fábrica III.

d) II é menor do que as das outras duas fábricas.

e) III é menor do que as das outras duas fábricas.

Comentários:

Sabemos que o coeficiente de variação é dado pela divisão entre o desvio padrão e a média. Assim, podemos calcular o coeficiente de variação das três empresas:

Fabrica I: $C_v = \frac{630,49}{1185} = 0,53$

Fabrica II: $C_v = \frac{355,97}{600} = 0,59$

Fabrica III: $C_v = \frac{1106,16}{2150} = 0,51$

Logo, o menor coeficiente de variação é da fábrica III.

Gabarito: E.

19. (FCC/TRT 13ª Região/2014) Seja X uma população $\{X_1, X_2, X_3, \dots, X_{100}\}$ formada por 100 números estritamente positivos com um desvio padrão igual a 4 e com a soma dos quadrados de todos estes 100 números igual a 41.600. Seja Y uma outra população $\{Y_1, Y_2, Y_3, \dots, Y_{50}\}$ formada por 50 números também estritamente positivos com uma média igual a da população anterior e com a soma dos quadrados de todos estes 50 números igual a 20.200. Os coeficientes de variação de X e de Y

a) são, ambos, iguais a 20%.

b) são iguais a 20% e 5%, respectivamente.

c) são, ambos, superiores a 15%.

d) apresentam uma diferença de valor absoluto igual a 10%.

e) apresentam um produto igual a 4%.

Comentários:

Conforme o enunciado, temos que a média de x é igual a média de y. Sabemos que $\bar{x} = \frac{\sum x}{n}$, então, segue que:

$$\bar{x} = \frac{\sum_1^{100} x}{100}$$

$$\bar{y} = \frac{\sum_1^{50} y}{50}$$

$$\bar{x} = \bar{y}$$

Para encontrarmos a variância, usamos:

$$\sigma^2 = \frac{\sum x^2}{n} - \bar{x}^2$$

Sabemos que a variância é igual ao desvio padrão ao quadrado, então, podemos encontrar a média. Fazendo para a população x, temos:

$$\sigma^2 = \frac{\sum_1^{100} x^2}{100} - \bar{x}^2$$

$$(4)^2 = \frac{41.600}{100} - \bar{x}^2$$

$$16 = 416 - \bar{x}^2$$

$$\bar{x}^2 = 416 - 16$$

$$\bar{x}^2 = 400$$

$$\bar{x} = \sqrt{400}$$

$$\bar{x} = 20$$

Agora, vamos calcular o coeficiente de variação de x, que é dado pela divisão entre o desvio padrão e a média. Logo, temos:

$$C_{vX} = \frac{\sigma_X}{\bar{x}}$$

$$C_{vX} = \frac{4}{20}$$

$$C_{vX} = 0{,}2 \rightarrow 20\%$$

Agora, vamos calcular a variância de Y:

$$\sigma^2 = \frac{\sum_1^{50} y^2}{50} - \bar{y}^2$$

$$\sigma^2 = \frac{20.200}{50} - (20)^2$$

$$\sigma^2 = 404 - 400$$

$$\sigma^2 = 4$$

Pronto, podemos calcular o coeficiente de variação de y. Sabemos que o desvio padrão pode ser encontrado tirando a raiz da variância. Logo, o desvio padrão de y é 2:

$$C_{vY} = \frac{\sigma_Y}{\bar{y}}$$

$$C_{vY} = \frac{2}{20}$$

$$C_{vY} = 0{,}1 \rightarrow 10\%$$

Assim a diferença entre C_{vX} e C_{vY} é:

$$C_{vX} - C_{vY} = 20 - 10$$

$$C_{vX} - C_{vY} = 10\%$$

Gabarito: D.

20. (FCC/TRT 16ª Região/2014) A média aritmética dos salários, em março de 2014, dos empregados em uma empresa é igual a R$ 2.500,00 com um coeficiente de variação igual a 9,6%. Decide-se aumentar os salários de todos os empregados, tendo que escolher uma entre as duas opções abaixo:

Opção I: Reajuste de todos os salários, em março de 2014, em 10% mais um abono fixo de R$ 250,00 para todos os salários.

Opção II: Reajuste de todos os salários, em março de 2014, em x% mais um abono fixo de R$ 200,00 para todos os salários.

Existe um valor para x tal que se for escolhida a opção II, a nova média aritmética passa a ser igual à nova média aritmética caso fosse escolhida a opção I. Nesta situação, o novo coeficiente de variação com a escolha da opção II passa a ser de

a) 8,00%.

b) 8,80%.

c) 9,00%.

d) 8,96%.

e) 9,60%.

Comentários:

Primeiro, devemos calcular o desvio padrão. Temos no enunciado as informações do coeficiente de variação e da média. Sabemos que o coeficiente de variação é dado pela divisão entre o desvio padrão e a média, logo, temos:

$$C_v = \frac{\sigma}{\bar{x}}$$

$$0{,}096 = \frac{\sigma}{2500}$$

$$\sigma = 2500 \times 0{,}096$$

$$\sigma = 240$$

Em seguida, vamos considerar a opção I, para calcularmos a nova média:

$$\overline{x_I} = \bar{x} \times 1{,}1 + 250$$

$$\overline{x_I} = 2500 \times 1{,}1 + 250$$

$$\overline{x_I} = 3000$$

Agora, vamos considerar a opção II, para encontrarmos o percentual (j) desta opção. Para tanto, precisamos igualar as médias:

$$\overline{x_{II}} = \bar{x} \times j + 200$$

$$3000 = 2500 \times j + 200$$

$$2500 \times j = 3000 - 200$$

$$2500 \times j = 2800$$

$$j = \frac{2800}{2500}$$

$$j = 1{,}12$$

Então, o percentual para o qual a nova média da opção II é igual à média da opção I é de 12%. Agora, vamos calcular o novo desvio padrão, considerando a opção II:

$$\sigma_{II} = \sigma \times 1{,}12$$

$$\sigma_{II} = 240 \times 1{,}12$$

$$\sigma_{II} = 268{,}8$$

Agora, basta aplicarmos a fórmula do coeficiente de variação:

$$C_{vII} = \frac{\sigma_{II}}{\bar{x}_{II}}$$

$$C_{vII} = \frac{268{,}8}{3000}$$

$$C_{vII} = 0{,}0896 \rightarrow 8{,}96\%$$

Gabarito: D.

21. (FGV/AL-BA/2014) A tabela a seguir mostra média e desvio padrão das notas dos alunos em um exame nacional em cinco estados diferentes:

	Média	Desvio padrão
Estado I	500	100
Estado II	600	120
Estado III	500	140
Estado IV	450	120
Estado V	600	100

Assinale a opção que indica o Estado que apresentou o menor coeficiente de variação das notas.

a) I

b) II

c) III

d) IV

e) V

Comentários:

O coeficiente de variação é dado pela divisão entre o desvio padrão e a média:

$$C_v = \frac{\sigma}{\bar{x}}$$

Aplicando aos cinco estados, temos:

	Média	Desvio padrão	CV
Estado I	500	100	0,2
Estado II	600	120	0,2
Estado III	500	140	0,28
Estado IV	450	120	0,26
Estado V	600	100	0,16

Gabarito: E.

22. (CESGRANRIO/IBGE/2013) Seja X uma variável aleatória com distribuição normal cuja média é μ e o desvio padrão é σ.

Se Y = 2X - 1 tem distribuição normal com média 5 e variância 20, o coeficiente de variação populacional $\frac{\sigma}{\mu}$ vale

a) $\frac{\sqrt{42}}{6}$

b) $\frac{\sqrt{21}}{6}$

c) $\frac{\sqrt{5}}{3}$

d) $\frac{\sqrt{39}}{9}$

e) $\frac{4\sqrt{5}}{9}$

Comentários:

Pelas propriedades da variância, sabemos que, se somarmos ou subtrairmos uma constante de todos os valores de uma variável, a variância do conjunto não se altera. Porém, se multiplicarmos ou dividirmos todos os valores de uma variável por uma constante, a variância fica multiplicada ou dividida pelo quadrado da constante. Assim, temos que:

$$Var(Y) = Var(2X - 1)$$

$$Var(Y) = 2^2 Var(X) + Var(1)$$

A variância de uma constante é sempre zero, então, $Var(1) = 0$. Dessa forma, temos:

$$20 = 4\sigma^2 + 0$$

$$\sigma^2 = 5$$

$$\sigma = \sqrt{5}$$

Pelas propriedades da média, sabemos que, se somarmos ou subtrairmos uma constante de todos os valores de uma variável, a média do conjunto também **aumenta ou diminui no valor da constante**. A mesma regra se aplica à multiplicação e divisão, a média é **multiplicada ou dividida pela constante**. Assim, temos:

$$\bar{Y} = 2 \times \bar{X} - 1$$

$$5 = 2\mu - 1$$

$$2\mu = 6$$

$$\mu = 3$$

Calculando o coeficiente de variação:

$$C_v = \frac{\sigma}{\mu}$$

$$C_v = \frac{\sqrt{5}}{3}$$

Gabarito: C.

23. (CESGRANRIO/BB/2013) A variância de um conjunto de dados é 4 m².

Para o mesmo conjunto de dados foram tomadas mais duas medidas de variabilidade:

a diferença entre o terceiro e o primeiro quartil e o coeficiente de variação.

Esses dois valores caracterizam-se, respectivamente, por

a) possuírem unidades de medida m² e m.

b) possuírem unidades de medida m e m².

c) ser adimensional e possuir unidade de medida m².

d) possuir unidade de medida m e ser adimensional.

e) possuir unidade de medida m² e ser adimensional.

Comentários:

Sabemos que a variância é representada no quadrado da unidade de dados original, logo os dados originais da questão estão em metros.

A amplitude interquartílica é dada por $Q_3 - Q_1$. Assim como os dados originais, estão expressos em metros.

O coeficiente de variação é dado por $C_v = \frac{\sigma}{\bar{x}}$, sendo assim, adimensional.

Gabarito: D.

24. (CESGRANRIO/BNDES/2013) Quatro variáveis são utilizadas em um modelo de previsão da quantidade produzida de uma determinada commodity agrícola. São elas:

• temperatura, em graus Celsius

• quantidade de fertilizante, em toneladas

• variação dos preços praticados no mercado internacional, em %

• quantidade produzida de um produto similar, em toneladas

Para determinar qual dessas variáveis apresenta a maior variabilidade, deve-se utilizar

a) apenas a média de cada uma das variáveis

b) apenas a variância de cada uma das variáveis

c) apenas o desvio padrão de cada uma das variáveis

d) a relação $\frac{desvio\ padrão}{média}$ de cada uma das variáveis

e) a relação $\frac{variância}{média}$ de cada uma das variáveis

Comentários:

Vamos analisar as alternativas:

Alternativa A – **Errada**. A média não representa uma medida de variabilidade e sim de tendência central.

Alternativa B – **Errada**. A variância é uma medida de dispersão absoluta, para comparar variáveis diferentes costumamos usar uma medida de dispersão relativa.

Alternativa C - **Errada**. O desvio padrão é também uma medida de dispersão absoluta.

Alternativa D – **Correta**. Conforme já dito, o coeficiente de variação é uma medida de dispersão relativa dada por:

$$Cv = \frac{desvio\ padrão}{média}$$

Alternativa E – **Errada**. O item está errado porque a média no denominador não está elevada ao quadrado.

Gabarito: D.

25. (FCC/TRT 5ª Região/2013) Seja Xi um elemento de uma população de tamanho 20, com $1 \leq i \leq 20$. Sabe-se que

$\sum_{i=1}^{20} X_i = 300 \ e \ \sum_{i=1}^{20} X_i^2 = 4.731,2$

O coeficiente de variação desta população apresenta um valor c, tal que

a) c < 21,0%.

b) 21,0% ≤ c < 21,5%.

c) 21,5% ≤ c < 22,0%.

d) 22,0% ≤ c < 22,5%.

e) c ≥ 22,5%.

Comentários:

Vamos iniciar com os dados da questão:

- a amostra tem tamanho n=20;
- o somatório dos elementos é $\sum_{i=1}^{20} X_i = 300$;
- o somatório dos quadrados é $\sum_{i=1}^{20} X_i^2 = 4.731{,}2$.

Sabemos que o coeficiente de variação é dado pela divisão entre o desvio padrão e a média, logo, temos:

$$C_v = \frac{\sigma}{\bar{x}}$$

Calculando a média, temos:

$$\bar{x} = \frac{300}{20} = 15$$

Para calcularmos o desvio padrão, devemos saber que a variância populacional é dada pela diferença entre a média dos quadrados e o quadrado da média:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

Ora, se o desvio padrão é a raiz quadrada da variância, podemos calcular da seguinte forma:

$$\overline{x^2} = \frac{4.731{,}2}{20} = 236{,}56$$

$$\bar{x}^2 = 15^2 = 225$$

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

$$\sigma^2 = 236{,}56 - 225$$

$$\sigma^2 = 11{,}56$$

Pronto, vamos encontrar o desvio padrão:

$$\sigma = \sqrt{11{,}56}$$

$$\sigma = 3{,}4$$

Agora, podemos calcular o coeficiente de variação:

$$C_v = \frac{3,4}{15}$$

$$C_v = 0,226 \rightarrow 22,6\%$$

Gabarito: E.

26. (FCC/TRT 6ª Região/2012) Um levantamento realizado em uma indústria revelou que o diâmetro médio de todas as 40 peças, marca Alpha, em estoque, é igual a 10 cm. Sabendo-se que a soma dos quadrados das medidas dos diâmetros de todas estas 40 peças apresenta o valor de 4.078,40 cm^2, então o coeficiente de variação correspondente é igual a

a) 10,0%.

b) 12,0%.

c) 12,5%.

d) 14,0%.

e) 15,0%.

Comentários:

Queremos encontrar o coeficiente de variação, sabemos que é dado pela divisão do desvio padrão pela média:

$$C_v = \frac{\sigma}{\bar{x}}$$

Ora, já temos a média e vale 10:

$$\bar{x} = 10$$

A questão nos deu o a soma dos quadrados das medidas das 40 peças $\sum_1^{40} \boldsymbol{X_i^2} = 4078{,}40$. Sabendo que o desvio padrão é igual a raiz quadrada da variância, podemos usar o cálculo da variância para descobrirmos quem é o desvio padrão.

A variância que é dada pela diferença entre a média dos quadrados e o quadrado da média:

$$\sigma^2 = \overline{x^2} - \bar{x}^2$$

A média dos quadrados é obtida pela divisão entre a soma dos quadrados pelo número de observações:

$$\overline{x^2} = \frac{4078,40}{40} = 101,96$$

Aplicando, temos:

$$\sigma^2 = 101,96 - 10^2$$

$$\sigma^2 = 101,96 - 100$$

$$\sigma^2 = 1,96$$

Agora, encontrando o desvio padrão:

$$\sigma = \sqrt{1,96}$$

$$\sigma = 1,4$$

Pronto, com todos os dados disponíveis, podemos calcular o coeficiente de variação:

$$C_v = \frac{\sigma}{\bar{x}}$$

$$C_v = \frac{1,4}{10}$$

$$C_v = 0,14 \rightarrow 14\%$$

Gabarito: D.

27. (CESGRANRIO/BACEN/2009) Considere as informações a seguir para responder à questão.

A viabilidade financeira do projeto de uma microempresa leva em consideração dados históricos de 100 projetos semelhantes.

A tabela abaixo mostra a distribuição de frequências do VPL - Valor Presente Líquido (valores em milhões de reais) de um conjunto de microempresas similares.

VPL	Frequência relativa
$-10 < x \leq 0$	10%
$0 < x \leq 10$	80%
$10 < x \leq 20$	10%

Um projeto alternativo para o investidor apresenta um VPL esperado, em reais, de 6 milhões e um risco (desvio padrão) de 2 milhões. Pela ótica do risco relativo, qual o melhor investimento, a microempresa ou o projeto alternativo?

a) A microempresa, pois apresenta um Coeficiente de Variação maior.

b) A microempresa, pois apresenta um Coeficiente de Variação menor.

c) O projeto alternativo, pois apresenta um Coeficiente de Variação maior.

d) O projeto alternativo, pois apresenta um Coeficiente de Variação menor.

e) É indiferente, pois os investimentos apresentam Coeficientes de Variação iguais.

Comentários:

Queremos calcular o coeficiente de variação. Vamos reescrever a tabela colocando os pontos médios e multiplicado as frequências:

VPL	Pontos médios (x_i)	x_i^2	Frequência (f)	$x_i^2 \times f$
$-10 < x \leq 0$	-5	25	0,1	2,5
$0 < x \leq 10$	5	25	0,8	20
$10 < x \leq 20$	15	225	0,1	22,5
	Total		1	45

Vamos iniciar a questão calculando a média. Na tabela, percebemos uma distribuição simétrica. Portanto, a média será o ponto médio da classe central. Assim, temos que:

$$\bar{x} = 5$$

$$\overline{x^2} = \frac{45}{1} = 45$$

A variância é dada por:

$$Var = \overline{x^2} - \bar{x}^2$$

$$Var = 45 - 5^2$$

$$Var = 45 - 25$$

$$Var = 20$$

O desvio-padrão será:

$$\sigma = \sqrt{Var}$$

$$\sigma = \sqrt{20}$$

$$\sigma \cong 4{,}47$$

O coeficiente de variação é dado por:

$$C_v = \frac{\sigma}{\bar{x}}$$

$$C_v = \frac{4{,}47}{5}$$

$$C_v = 0{,}89$$

Considerando o projeto alternativo:

$$C_v = \frac{\sigma}{\bar{x}}$$

$$C_v = \frac{2}{6}$$

$$C_v = 0{,}33$$

Logo, o projeto alternativo tem um coeficiente de variação menor e, consequentemente, menor risco.

Gabarito: D.

28. (FGV/SEN/2008) O coeficiente de variação amostral (em porcentagem) de um conjunto de salários é 110%.

Se os salários desse conjunto forem reajustados em 20%, o novo coeficiente de variação amostral será:

a) 110%.

b) 112,2%.

c) 114,2%.

d) 122%.

e) 130%.

Comentários:

Ao multiplicarmos o coeficiente de variação por uma constante o valor do coeficiente não será alterado. Portanto, permanecerá em 110%.

Gabarito: A.

29. (FGV/SEFAZ-RJ/2008) Uma companhia utiliza um sistema de avaliação de desempenho de seus funcionários por meio de dois indicadores de performance: Qualidade das tarefas e a Tempestividade com que as tarefas são realizadas.

Os funcionários receberam, na última avaliação, as medidas indicadas na tabela a seguir:

Medidas	Indicador	
	Qualidade	Tempestividade
Média	50	25
Desvio-Padrão	10,0	6,0
Coeficiente de Variação %	20	24

Com base na tabela, é correto afirmar que:

a) a média aritmética não é uma boa medida para representar a performance dos funcionários em face do elevado nível de dispersão das avaliações.

b) as avaliações da Qualidade foram mais dispersas do que as avaliações da Tempestividade.

c) as avaliações da Qualidade foram mais homogêneas do que as da Tempestividade.

d) os funcionários demoram mais para realizar as tarefas, mas a qualidade das tarefas é melhor.

e) nada se pode afirmar sem o conhecimento do tamanho da amostra.

Comentários:

Analisando as alternativas:

Letra A: **Errada**. Considera-se que para valores inferiores a 50%, a média será tanto mais representativa do fato quanto menor for o valor de seu coeficiente de variação. Logo, as duas médias da tabela seriam sim representativas já que os coeficientes de variação são inferiores de 50%.

Letra B: **Errada**. Para essa análise consideramos o coeficiente de variação, a análise é que a dispersão relativa das avaliações de qualidade é menor (menos dispersas) que a das avaliações de tempestividade.

Letra C: **Certa**. O coeficiente de variação para a qualidade foi menor, logo as avaliações de qualidade são mais homogêneas, já que a homogeneidade está relacionada à pequena dispersão.

Letra D: **Errada**. A questão não traz informações suficientes para essa afirmação.

Letra E: **Errada**. No enunciado da questão temos a informação de que todos os funcionários foram avaliados, portanto, o quadro apresenta informações da população toda e não apenas de uma amostra.

Gabarito: C.

30. (FGV/SEFAZ RJ/2007) Uma amostra de 100 servidores de uma repartição apresentou média salarial de R$ 1.700,00 com uma dispersão de R$ 240,00. Pode-se afirmar que:

a) a média aritmética não é uma boa medida para representar a amostra em função do elevado valor do desvio-padrão.

b) a melhor medida para representar a amostra é a remu-neração por unidade de desvio-padrão.

c) o salário mediano representaria melhor a amostra de-vido ao alto nível de heterogeneidade dos salários na amostra.

d) a amostra não é suficientemente grande para analisar-mos o valor encontrado para a média dos salários.

e) a média aritmética pode perfeitamente representar os salários da amostra pelo fato de esta apresentar uma dispersão relativa inferior a 20%.

Comentários:

Vamos focar na análise das alternativas A e E.

Letra A: **Errada**. Considera-se que para valores inferiores a 50%, a média será tanto mais representativa do fato quanto menor for o valor de seu coeficiente de variação. Vamos calcular o coeficiente de variação:

$$C_v = \frac{\sigma}{\bar{x}}$$

$$C_v = \frac{240}{1700} =$$

$$C_v = 14{,}11\%$$

Logo, a média seria sim representativas, pois o coeficiente de variação de variação é inferior a 50%.

Com essa análise percebemos que a alternativa E é a correta.

Gabarito: E.

QUESTÕES COMENTADAS

Variância Relativa

1. (FCC/SEFAZ BA/2019) O coeficiente de variação de Pearson correspondente a uma população P_1 com média aritmética igual a 20 e tamanho 20 é igual a 30%. Decide-se excluir de P_1, em um determinado momento, dois elementos iguais a 11 cada um, formando uma nova população P_2. A variância relativa de P_2 é igual a

a) 10/147.

b) 4/49.

c) 16/147.

d) 8/49.

e) 4/441.

Comentários:

Vamos iniciar a resolução do problema com a população 1. Calculamos o somatório de P_1:

$$\overline{x_1} = \frac{\sum x_1}{n} \rightarrow 20 = \frac{\sum x_1}{20} \rightarrow \sum x_1 = 400$$

Calculando o desvio padrão e a variância de P_1:

$$C_v = \frac{\sigma}{\overline{x_1}}$$

$$0{,}3 = \frac{\sigma}{20} \Rightarrow \sigma = 20 \times 0{,}3 = 6 \rightarrow desvio\ padrão$$

$$\sigma^2 = 6^2 = 36 \rightarrow variância$$

A variância para 20 elementos:

$${\sigma_1}^2 = \frac{\sum {x_1}^2}{n} - \overline{x_1}^2$$

$$36 = \frac{\sum {x_1}^2}{20} - 20^2$$

$$36 + (20)^2 = \frac{\sum {x_1}^2}{20}$$

$$\sum {x_1}^2 = 20 \times (36 + 400)$$

$$\sum {x_1}^2 = 20 \times 436$$

$$\sum {x_1}^2 = 8.720$$

Para a população 2, o enunciado diz que foram excluídos dois termos iguais a 11. Assim, o somatório fica:

$$\sum x_2 = \sum x_1 - 11 - 11$$

$$\sum x_2 = \sum x_1 - 22$$

$$\sum x_2 = 400 - 22$$

$$\sum x_2 = 378$$

Para a média de P_2, temos:

$$\overline{x_2} = \frac{\sum x_2}{18} \rightarrow \frac{378}{18} = 21$$

Fazendo o somatório dos quadrados:

$$\sum {x_2}^2 = \sum {x_1}^2 - 11^2 - 11^2$$

$$\sum {x_2}^2 = 8.720 - 121 - 121$$

$$\sum {x_2}^2 = 8.478$$

Temos que a variância de P_2 é:

$$\sigma_2{}^2 = \frac{\sum {x_2}^2}{n} - \overline{x_2}^2$$

$$\sigma_2{}^2 = \frac{8.478}{18} - 21^2$$

$$\sigma_2{}^2 = 471 - 441$$

$$\sigma_2{}^2 = 30$$

Calculando a variância reativa de P_2:

$$C_v{}^2 = \frac{\sigma_2{}^2}{\overline{x_2}^2} \rightarrow \frac{30}{21^2}$$

Simplificando, temos:

$$C_v{}^2 = \frac{10}{7 \times 21}$$

$$C_v{}^2 = \frac{10}{147}$$

Gabarito: A.

Lista de Questões

Medidas de Dispersão

1. (FCC/TJ SC/2021) O número de processamento de documentos da Divisão de Pesquisa e Informação do Tribunal de Justiça requer uma análise refinada de variáveis quantitativas. Dentre os procedimentos analíticos, indicam-se as medidas de dispersão, representadas por

a) Média harmônica e Média geométrica.

b) Amplitude total e Desvio médio absoluto.

c) Moda e Mediana.

d) Média aritmética simples e Média aritmética ponderada.

e) Mediana para dados agrupados em classes e Mediana para dados não agrupados.

2. (VUNESP/MPE-SP/2016) 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.

Gabarito

Medidas de Dispersão

1. LETRA B
2. LETRA D

LISTA DE QUESTÕES

Amplitude Total

1. (VUNESP/IPSM SJC/2018) Considere as informações a seguir para construir uma distribuição de frequência sem intervalo de classe e responder a questão.

Um dado foi lançado 50 vezes e foram registrados os seguintes resultados:

5 4 6 1 2 5 3 1 3 3

4 4 1 5 5 6 1 2 5 1

3 4 5 1 1 6 6 2 1 1

4 4 4 3 4 3 2 2 2 3

6 6 3 2 4 2 6 6 2 1

A amplitude total é

a) 50.

b) 6.

c) 7.

d) 10.

e) 5.

2. (CESPE/SEDF/2017) Um levantamento estatístico, feito em determinada região do país, mostrou que jovens com idades entre 4 e 17 anos assistem à televisão, em média, durante 6 horas por dia. A tabela a seguir apresenta outras estatísticas produzidas por esse levantamento.

	Distribuição dos tempos gastos assistindo televisão (*T, em horas*)
1° quartil	2
2° quartil	4
3° quartil	8
1° decil	1
9° decil	10

Tendo como referência essas informações, julgue o seguinte item.

A amplitude total dos tempos T é igual ou superior a 9 horas.

3. (CESGRANRIO/IBGE/2016) Suponha que, em uma pesquisa on-line sobre as idades dos habitantes de um condomínio, um respondente de 30 anos digite erroneamente sua idade como sendo 300 anos. Considere que esse erro passe despercebido e que não haja outros erros na base de dados.

Nessas condições, a única conclusão que NÃO pode ser formulada é:

a) A média de idades calculada a partir dos dados da base será maior do que a média de idades reais dos respondentes.

b) A mediana de idades calculada a partir dos dados da base será maior do que a mediana de idades reais dos respondentes.

c) A amplitude de idades calculada a partir dos dados da base será maior do que a amplitude de idades reais dos respondentes.

d) O valor máximo das idades calculado a partir dos dados da base será maior do que a idade real do respondente mais velho.

e) A diferença entre as duas maiores idades dos dados da base será maior do que a diferença das idades reais dos dois respondentes mais velhos.

4. (CESGRANRIO/IBGE/2016) Uma pesquisa em determinado município coletou, dentre outros dados, o número de filhos em cada família. Algumas estatísticas são apresentadas na Tabela abaixo.

Número de filhos	
Média	2
Mediana	1
Moda	0
Desvio-padrão	3
Amplitude	5

Segundo essas estatísticas,

a) metade das famílias tem mais do que 2 filhos.

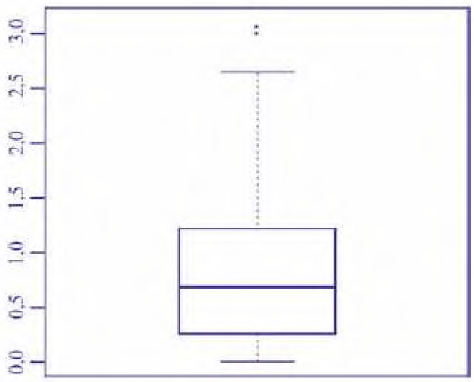
b) o mais comum é que famílias tenham 2 filhos.

c) mais da metade das famílias não têm filhos.

d) uma família padrão tem em média 3 filhos.

e) de todas as famílias entrevistadas, nenhuma tem 6 filhos.

5. (CESPE/TCE-PA/2016)



média amostral	0,80
desvio padrão amostral	0,70
primeiro quartil	0,25
mediana	0,70
terceiro quartil	1,20
mínimo	0
máximo	3,10

Um indicador de desempenho X permite avaliar a qualidade dos processos de governança de instituições públicas. A figura mostra, esquematicamente, a sua distribuição, obtida mediante estudo amostral feito por determinada agência de pesquisa. A tabela apresenta estatísticas descritivas referentes a essa distribuição.

Com base nessas informações, julgue o item a seguir.

A amplitude total da amostra é inferior a 3.

6. (CESPE/TCE-PA/2016)

Número diário de denúncias registradas (X)	Frequência Relativa
0	0,3
1	0,1
2	0,2
3	0,1
4	0,3
Total	**1,0**

A tabela precedente apresenta a distribuição de frequências relativas da variável X, que representa o número diário de denúncias registradas na ouvidoria de determinada instituição pública. A partir das informações dessa tabela, julgue o item seguinte.

A amplitude total da amostra é igual ou superior a 5.

7. (FGV/DPE-RJ/2014) Dentre as informações coletadas dos cidadãos através do 1º atendimento da Defensoria Pública estão as variáveis idade, renda e o número de dependentes. Cada uma é classificada em três diferentes níveis A, B e C, com valores de referência conforme a tabela:

Variáveis	A	B	C	D
Renda Média (R$ 100)	9	11	11	17
Idade Média (anos)	20	32	36	48
Dependentes (pessoas)	2	3	3	2

Portanto, as unidades de medida são distintas (R$, anos e pessoas). Mesmo assim, através de uma estatística de amplitude, escolhida convenientemente, aqui representada por VB, é possível comparar as dispersões. Logo, renda, idade e número de dependentes seguem a ordenação

a) VB (renda) < VB (idade) < VB (dependentes).

b) VB (renda) < VB (dependentes) < VB (idade).

c) VB (idade) < VB (dependentes) < VB (renda).

d) VB (idade) < VB (renda) < VB (dependentes).

e) VB (dependentes) < VB (renda) < VB (idade).

GABARITO

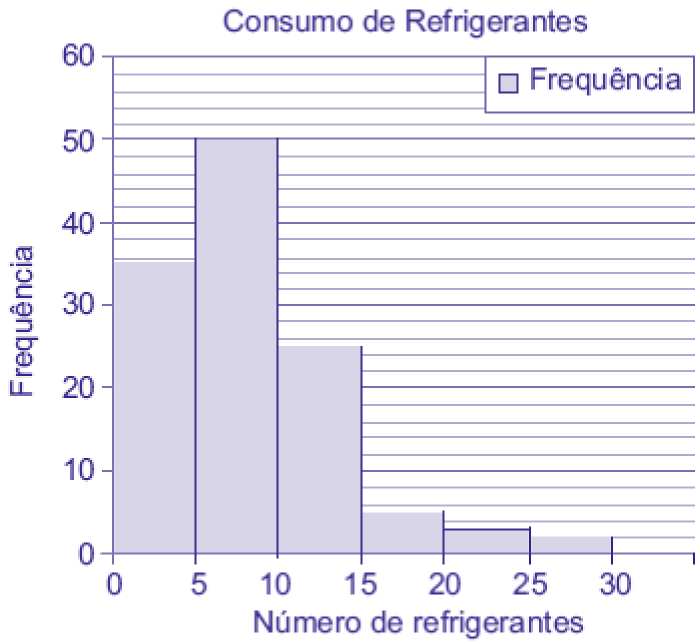
Amplitude Total

1. LETRA E
2. CERTO
3. LETRA B
4. LETRA E
5. CERTO
6. ERRADO
7. LETRA E

LISTA DE QUESTÕES

Amplitude Interquartílica

1. (CESGRANRIO/BB/2018) Uma escola de Ensino Médio decide pesquisar o comportamento de seus estudantes quanto ao número de refrigerantes consumidos semanalmente por eles. Para isso, uma amostra aleatória de 120 estudantes foi selecionada, e os dados foram sintetizados no histograma abaixo, em classes do tipo [0, 5), [5, 10), [10, 15), [15, 20), [20, 25) e [25, 30].



Qual o valor da amplitude interquartílica, obtido por meio do método de interpolação linear dos dados agrupados em classes?

a) 15

b) $\frac{15}{2}$

c) $\frac{29}{5}$

d) $\frac{47}{7}$

e) 10

2. (CESGRANRIO/PETROBRAS/2018) Você dispõe de um montante para investir em ações e precisa decidir em que empresa(s) vai alocar esse montante. Três empresas lhe parecem interessantes, e você resolve consultar o desempenho delas nos últimos sessenta meses para minimizar possíveis riscos da sazonalidade no movimento da Bolsa de Valores. Os dados revelaram a seguinte distribuição, em %, das rentabilidades mensais das ações:

Medidas Estatísticas	Empresa A	Empresa B	Empresa C
Rentabilidade média mensal	0,50	0,60	0,40
Desvio padrão	1,00	1,20	0,80
Rentabilidade mínima	-1,80	-2,20	-1,20
Rentabilidade máxima	2,20	2,30	1,80
1º quartil	-0,20	-0,30	-0,10
3º quartil	0,80	0,90	0,70

A alocação dos recursos vai ser feita de acordo com a atitude conservadora de não investir em empresa com rentabilidade considerada outlier, entendendo como tal aquela que apresentar valor além de 1,5 desvio quartílico abaixo ou acima dos quartis 1 e 3.

Com base nesse critério, a escolha do investimento deve recair sobre a(s)

a) empresa A, apenas

b) empresa B, apenas

c) empresa C, apenas

d) empresas A e C, apenas

e) três empresas

3. (CESPE/SEE-DF/2017) Um levantamento estatístico, feito em determinada região do país, mostrou que jovens com idades entre 4 e 17 anos assistem à televisão, em média, durante 6 horas por dia. A tabela a seguir apresenta outras estatísticas produzidas por esse levantamento.

	Distribuição dos tempos gastos assistindo televisão (*T, em horas*)
1° quartil	2
2° quartil	4
3° quartil	8
1° decil	1

9° decil	10

Tendo como referência essas informações, julgue o seguinte item.

O desvio quartílico dos tempos T foi igual a 3.

4. (FGV/IBGE/2017) A população de um estudo é dividida em quatro estratos, sendo o menor com 10% dos indivíduos e os demais com tamanhos acrescidos de dez pontos percentuais, progressivamente. Os estratos se distinguem por classes de renda com amplitude constante, sendo maiores quanto menor a renda. Sobre os estratos sabe-se que:

$$Rd_{Estrato1} = 65\ Rd_{Estrato2} = 45\ e\ Rd_{Estrato4} = 5$$

Onde os valores acima representam os limites inferiores da renda dos extratos, inclusive.

Portanto, é correto afirmar que:

a) Tomando os pontos médios das classes como representativos, a renda média é igual a Md(Rd) = 38;

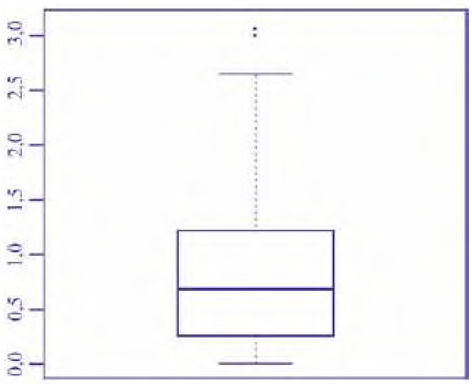
b) A mediana da distribuição de renda, Me(Rd), é menor que 45 e maior do que ou igual a 25;

c) Tomando os pontos médios das classes como representativos, a moda da renda é igual a Mo(Rd) = 35;

d) O valor máximo atingido pela renda nessa distribuição é igual a Mx(Rd) = 85;

e) O valor do desvio-interquartílico da distribuição de renda deverá ser superior a 50.

5. (CESPE/TCE-PA/2016)



média amostral	0,80
desvio padrão amostral	0,70
primeiro quartil	0,25
mediana	0,70
terceiro quartil	1,20
mínimo	0
máximo	3,10

Um indicador de desempenho X permite avaliar a qualidade dos processos de governança de instituições públicas. A figura mostra, esquematicamente, a sua distribuição, obtida mediante estudo amostral feito por determinada agência de pesquisa. A tabela apresenta estatísticas descritivas referentes a essa distribuição.

Com base nessas informações, julgue o item a seguir.

O intervalo interquartílico da distribuição do indicador X é superior a 1,4.

6. (CESGRANRIO/EPE/2014) A Tabela a seguir apresenta a vazão média em cada mês para um determinado rio.

Mês	Vazão média mensal (m^3/s)
Janeiro	97
Fevereiro	60
Março	50
Abril	60
Maio	70
Junho	85
Julho	60
Agosto	50
Setembro	68
Outubro	117
Novembro	80
Dezembro	43

De acordo com os dados da Tabela, a mediana e a amplitude interquartílica das vazões valem, respectivamente,

a) 70 e 25

b) 64 e 25

C) 64 e 27,5

d) 68 e 27,5

e) 70 e 27,5

7. (FCC/TCE-PR/2011) Atenção: Considere as informações a seguir para responder à questão.

A distribuição dos salários dos 1000 funcionários da companhia A, em número de salários mínimos, está apresentada na tabela abaixo:

Faixa salarial (em número de salários mínimos)	Frequência absoluta
$1 \vdash 3$	200
$3 \vdash 5$	400
$5 \vdash 7$	200
$7 \vdash 9$	200

A distância interquartil desses salários, definida por $Q_3 - Q_1$, onde Q_3 e Q_1 são, respectivamente, os quartis de ordem 3 e 1, calculados pelo método da interpolação linear, em número de salários mínimos, é

a) 2,75.

b) 3,00.

c) 3,25.

d) 3,50.

e) 4,00.

8. (FCC/BACEN/2006) Considere a distribuição de frequências a seguir para resolver a questão.

Salários dos empregados das empresas XYZ em dezembro de 2005

Salários	Frequências simples absolutas
$1.000,00 \vdash 2.000,00$	2
$2.000,00 \vdash 3.000,00$	8
$3.000,00 \vdash 4.000,00$	16
$4.000,00 \vdash 5.000,00$	10
$5.000,00 \vdash 6.000,00$	4

A amplitude do intervalo entre o primeiro decil e o terceiro quartil, encontrados pelo método da interpolação linear, é

a) R$ 2 500,00

b) R$ 2 400,00

c) R$ 2 150,00

d) R$ 2 000,00

e) R$ 1 400,00

GABARITO

Amplitude Interquartílica

1. LETRA D
2. LETRA C
3. CERTO
4. LETRA B
5. ERRADO
6. LETRA C
7. LETRA C
8. LETRA C

LISTA DE QUESTÕES

Desvios em Relação à Média Aritmética e Mediana

1. (VUNESP/Pref. Campinas/2019) Considere a tabela-1 e o enunciado seguintes para responder à questão.

x_i	f_i	$x_i \times f_i$	$(x_i - \bar{x})^2$
1	2	2	A
3	4	12	4
5	6	30	B
7	4	28	4
9	2	18	C
Totais	**18**	**90**	**40**

Tabela-1

A tabela-1 de distribuição de frequência mostra a organização e síntese de 18 dados x_i colhidos como amostra para um estudo estatístico, onde a coluna f_i é a que registra os valores das frequências, enquanto a coluna $(x_i - \overline{x})^2$ contém os valores dos quadrados dos desvios.

Os valores substituídos pelas letras A, B e C na tabela são, respectivamente:

a) 0, 16, 0.

b) 4, 4, 4.

c) 16, 0, 16.

d) 0, 4, 0.

e) 16, 4, 16.

2. (VUNESP/TJ SP/2015) Leia o texto a seguir para responder à questão.

Uma pequena empresa que emprega apenas cinco funcionários paga os seguintes salários mensais (em mil reais):

0,9	1,2	1,4	1,5	2,0

Considerando-se a média dos salários, o valor do desvio do salário de quem ganha R$ 1.400,00 mensais é

a) –1.000.

b) –400.

c) 0.

d) 200.

e) 400.

3. (VUNESP/TJ SP/2015) Leia o texto a seguir para responder à questão.

Uma pequena empresa que emprega apenas cinco funcionários paga os seguintes salários mensais (em mil reais):

0,9	1,2	1,4	1,5	2,0

Somando-se os valores absolutos dos desvios individuais dos salários tomados em relação à média, encontra-se o valor de

a) 1.400,00.

b) 1.200,00.

c) 1.000,00.

d) 800,00.

e) 0.

4. (FGV/SEFAZ MS/2006) Analise as afirmativas a seguir, a respeito da média aritmética:

I. A soma dos resíduos em relação à média aritmética é sempre igual a zero.

II. É em relação à média aritmética que a soma dos valores absolutos dos resíduos é mínima.

III. É em relação à média aritmética que a soma dos quadrados dos resíduos é mínima.

Assinale:

a) se somente a afirmativa II estiver correta.

b) se somente as afirmativas I e II estiverem corretas.

c) se somente as afirmativas I e III estiverem corretas.

d) se somente as afirmativas II e III estiverem corretas.

e) se todas as afirmativas estiverem corretas.

5. (FGV/SEFAZ-MS/2006) Analise as afirmativas a seguir, a respeito da mediana:

I. A soma dos resíduos em relação à mediana é sempre igual a zero.

II. É em relação à mediana que a soma dos valores absolutos dos resíduos é mínima.

III. É em relação à mediana que a soma dos quadrados dos resíduos é mínima.

Assinale:

a) se somente a afirmativa II estiver correta.

b) se somente as afirmativas I e II estiverem corretas.

c) se somente as afirmativas I e III estiverem corretas.

d) se somente as afirmativas II e III estiverem corretas.

e) se todas as afirmativas estiverem corretas.

6. (CESPE/ANATEL/2004)

meses	fev	mar	abr	mai	jun	jul	ago	set	out	nov
N	100	70	70	60	50	100	50	50	30	20

A tabela acima mostra os números mensais de reclamações (N) feitas por usuários de telefonia fixa, registradas em uma central de atendimento, entre os meses de fevereiro a novembro de 2003. Considerando esses dados, julgue o item que se segue.

O maior desvio absoluto dos números mensais de reclamações registradas é superior a 45.

Gabarito

Desvios em Relação à Média Aritmética e Mediana

1. LETRA C
2. LETRA C
3. LETRA A
4. LETRA C
5. LETRA A
6. ERRADO

LISTA DE QUESTÕES

Desvio Absoluto Médio

1. (FUNDATEC/SEPOG RS/2022) "O ________________ é uma boa medida de dispersão, porque dá a distância média de cada número em relação à média. Todavia, para muitos propósitos, é mais conveniente elevar ao quadrado cada _________ e tomar a média de todos esses quadrados. Essa grandeza é chamada _______________". (DOWNING; CLARK, 2011).

Assinale a alternativa que preenche, correta e respectivamente, as lacunas do trecho acima.

a) desvio médio absoluto – desvio – coeficiente de variação

b) desvio médio absoluto – desvio – variância

c) desvio médio absoluto – média – variância

d) desvio padrão – desvio – coeficiente de variação

e) desvio padrão – média – variância

2. (CESPE/ANATEL/2004)

meses	fev	mar	abr	mai	jun	jul	ago	set	out	nov
N	100	70	70	60	50	100	50	50	30	20

A tabela acima mostra os números mensais de reclamações (N) feitas por usuários de telefonia fixa, registradas em uma central de atendimento, entre os meses de fevereiro a novembro de 2003. Considerando esses dados, julgue o item que se segue.

O desvio médio absoluto da sequência formada pelos números mensais de reclamações é um valor entre 25 e 35

Gabarito

Desvio Absoluto Médio

1. LETRA B
2. ERRADO

LISTA DE QUESTÕES

Variância

1. (CESPE/TCE SC/2022) Julgue o item a seguir, considerando conceitos de estatística.

Com os seguintes dados, a variância da população é de 149,25.

36; 64; 18; 40; 35; 30; 41; 32

2. (CESPE/PETROBRAS/2022) O item a seguir é apresentada uma situação hipotética seguida de uma assertiva a ser julgada a respeito de probabilidade e estatística.

Ao adicionar uma medição a mais, x_{21}, a um conjunto com inicialmente 20 medições de uma dada grandeza, $\{X_1, X_2, \cdots, X_{20}\}$, a média aritmética μ do novo conjunto não se altera. Nesse caso, a variância σ^2 do conjunto inicial relaciona-se com a variância σ_n^2 do novo conjunto na forma $\sigma_n^2 = \frac{20}{21}\sigma^2$.

3. (CESPE/PETROBRAS/2022) No que diz respeito aos conceitos e cálculos utilizados em probabilidade e estatística, julgue o item a seguir.

Se, em determinada semana, as ações da PETROBRAS fecharam o pregão com as cotações, em unidades monetária, iguais a 10,0; 9,0; 11,0; 12,0 e 8,0, respectivamente de segunda à sexta-feira, então a variância dessas cotações foi igual a 2,0.

4. (CESPE/PETROBRAS/2022)

X	Frequência Relativa
0	0,23
1	0,22
2	0,50
3	0,05

Considerando que a tabela acima mostra a distribuição de frequências de uma variável x obtida com base em uma amostra aleatória simples de tamanho igual a n, julgue o item que se segue.

A variância amostral de x é inferior a 0,7.

5. (FGV/CGU/2022) O ativo A está gerando grande atração de matemáticos, que conseguiram convencer a bolsa de valores a registrar preços baseados em quantidades pouco usuais, como $\sqrt{2}$ e π. Durante cinco dias foram observados os seguintes preços de dois ativos, A e B, respectivamente:

$$\left(\sqrt{2}; \frac{\pi}{2}; 1; \frac{1}{3}; \frac{1}{6}\right) e\ (100,30; 400,18; 207,01; 508,00; 912,11)$$

Considerando esses valores, sobre a média e a variância dos retornos durante esses cinco dias, é correto afirmar que:

a) os retornos do ativo A têm maior média e maior variância;

b) os retornos do ativo A têm maior média e menor variância;

c) os retornos do ativo A têm menor média e maior variância;

d) os retornos do ativo A têm menor média e menor variância;

e) a média e a variância dos retornos dos dois ativos são iguais.

6. (CESPE/TJ RJ/2021) Considere que, em um estudo para avaliar a satisfação dos serviços de comunicação de dados oferecidos por uma operadora, no qual foram utilizadas duas variáveis, X e Y, observou-se que X = 6Y + 24 e que o valor da variância de Y foi igual a 1. Nesse caso, o valor da variância de X é

a) 30.

b) 60.

c) 6.

d) 24.

e) 36.

7. (CESPE/MJ SP/2021) Acerca de planejamento de pesquisa estatística, julgue o item que se seguem.

A média do erro entre a média calculada e as observações reais em um conjunto de dados é conhecida como variância.

8. (CESPE/BANESE/2021) A respeito do conjunto de dados {11, 6, 28, 51, 49, 32, 33}, julgue o item a seguir.

Esse conjunto de dados possui variância amostral inferior a 300.

9. (CESPE/BANESE/2021)

	X	Y
Média	5	10
Desvio padrão	2	2

Com base nas informações apresentadas na tabela precedente e considerando que a covariância entre as variáveis X e Y seja igual a 3, julgue o item que se segue.

A variância de X é igual a 4.

10. (CESPE/SEDUC AL/2021) Com base em estatística, julgue o item a seguir.

Para um conjunto de dados $x_1, x_2, \ldots, x_n$ quaisquer, a variância será sempre um número positivo.

11. (CESPE/TCE-RJ/2021)

X	Frequência Absoluta
0	5
1	10
2	20
3	15
Total	**50**

Considerando que a tabela precedente mostra a distribuição de frequências de uma variável quantitativa X, julgue o item a seguir.

A variância amostral de X é superior a 0,89.

12. (VUNESP/EsFCEx/2021) Uma amostra aleatória de 10 elementos de uma população para a estimação da média e da variância de uma variável com distribuição normal forneceu 500 e 25 844 para a soma dos valores e dos quadrados dos valores, respectivamente. É correto afirmar que a estimativa de máxima verossimilhança para a variância é

a) 84,4.

b) 25,844.

c) 258,44.

d) 9,38.

e) 93,8.

13. (VUNESP/EsFCEx/2021) Alunos de uma turma realizaram cinco provas de uma disciplina. Entretanto, o professor divulgou as notas das quatro provas e a variância populacional das cinco notas. João é aluno desta turma e deseja saber qual é a sua nota na 5a prova, as notas dele foram:

1ª Prova	2ª Prova	3ª Prova	4ª Prova
4	6	8	6

Sabendo que a variância populacional das suas notas foi 1,6, a nota do João na 5a Prova é

a) 3.

b) 5.

c) 6.

d) 2.

e) 4.

14. (FCC/ALAP/2020) O número de empregados de uma empresa é igual a 200, sendo que 60% são homens e o restante mulheres. Nesta empresa, a média aritmética dos salários da população formada pelos salários dos homens é igual a 5 mil reais, com um coeficiente de variação igual a 30%, e a média aritmética dos salários da população formada pelos salários das mulheres também é igual a 5 mil reais, porém com um coeficiente de variação igual a 20%. Considerando a população formada por todos os 200 empregados da empresa, obtém-se que a variância, em mil reais ao quadrado, dos respectivos salários é igual a

a) 1,69

b) 1,75

c) 1,30

d) 2,50

e) 3,25

15. (CESGRANRIO/BB/2018) Uma amostra aleatória de tamanho 5 é retirada de uma população e observa-se que seus valores, quando postos em ordem crescente, obedecem a uma Progressão Aritmética.

Se a variância amostral não viciada vale 40, qual é o valor da razão da Progressão Aritmética?

a) 3

b) $5\sqrt{2}$

c) 4

d) $2\sqrt{5}$

e) 1

16. (FCC/TRT 14ª Região/2018) Considere uma população P_1formada pela renda, em unidades monetárias (u.m.), dos 100 indivíduos que são sócios de um clube. Seja x_ia renda, $x_i > 0$, do sócio i.

Dados:

$\sum_{i=1}^{100} x_i^2 = 2.662.400(u.m)^2$ e Coeficiente de variação de P_1 igual a 20%.

Decide-se excluir de P_1 um total de 20 sócios que possuem renda igual à média de P_1 , formando uma nova população P_2 com tamanho 80. O módulo da diferença, em (u.m.) ^{2}, entre as variâncias de P_1 e P_2 é de

a) 144.

b) 0.

c) 64.

d) 256.

e) 400.

17. (FGV/SEPOG-RO/2017) Considere a seguinte amostra de notas de alunos: 5,0; 6,0; 5,0; 4,0; 10,0. A variância amostral dessas notas pode ser igual a

a) 5,5.

b) 6,2.

c) 6,8.

d) 7,2.

e) 9,0.

18. (FCC/DPE SP/2015) Foi realizado um censo em uma faculdade com 200 alunos e obteve-se com relação às alturas dos alunos, em centímetros (cm), um coeficiente de variação igual a 10%. Se a soma dos quadrados de todas as alturas foi igual a 5.499.450 cm 2, então a correspondente variância apresentou um valor igual a

a) 289,00 cm 2.

b) 256,00 cm 2.

c) 306,25 cm 2.

d) 324,00 cm 2.

e) 272,25 cm 2.

19. (FCC/CNMP/2015) Em um censo realizado em um clube apurou-se a altura em centímetros (cm) de seus 200 associados. A média aritmética apresentou um valor igual a 160 cm com um coeficiente de variação igual a 18,75%. O resultado da divisão da soma de todos os valores das alturas elevados ao quadrado pelo número de associados é, em cm^2, de

a) 27.050.

b) 25.600.

c) 26.050.

d) 26.500.

e) 25.060.

20. (CESGRANRIO/EPE/2014) Uma amostra de tamanho 200, $x_1, x_2, ..., x_{200}$, foi retirada de uma população, e seus valores foram transformados segundo a função $y_i = 4x_i - 1$ para $i = 1, 2, ..., 200$.

Sabendo-se que a média e a variância dos dados transformados $y_1, y_2, ..., y_{200}$ são, respectivamente, 3 e 16, os valores da média e da variância dos dados originais são, respectivamente,

a) 1 e 1

b) 1 e 4

c) 3/4 e 63

d) 11 e 64

e) 11 e 256

21. (CESGRANRIO/FINEP/2014) O enunciado a seguir deve ser usado para responder à questão.

Abaixo são apresentadas estatísticas das notas brutas obtidas pelos candidatos em um concurso público:

Média aritmética: 78

Variância: 100

A nota de cada candidato foi transformada em nota padronizada, calculada considerando-se a seguinte fórmula:

$$\boldsymbol{Nota\ padronizada = 50 + 5 \times \frac{Nota\ bruta\ do\ candidato\ -\ Media\ aritmetica\ das\ notas\ brutas}{Desvio\ padrao\ das\ notas\ brutas}}$$

A variância das notas padronizadas é

a) 25

b) 50,5

c) 52,5

d) 55

e) 75

22. (FCC/TRT 16ª Região/2014) Seja $\{X_1, X_2, X_3, ..., X_{80}\}$ uma população constituída de 80 números estritamente positivos, sabendo-se que a média aritmética e o desvio padrão desta população são, respectivamente iguais a 20 e 15. Resolve-se excluir desta população 30 números, cuja soma de seus quadrados é igual a 12.000, formando uma nova população e o novo valor da variância passa a ter o valor de 436. O correspondente novo valor da média aritmética da nova população apresenta um valor igual a

a) 16.

b) 24.

c) 22.

d) 23.

e) 18.

23. (FCC/TRT 13ª Região/2014) Em uma determinada carreira profissional composta por 400 trabalhadores, verifica-se que a média aritmética das alturas de todos os trabalhadores é igual a 170 cm. Sabe-se que a média aritmética das alturas dos 250 trabalhadores do sexo masculino é igual à média aritmética das alturas dos 150 trabalhadores do sexo feminino. Os desvios padrões das alturas dos trabalhadores do sexo masculino e dos trabalhadores do sexo feminino são iguais a 12 cm e 20 cm, respectivamente. A variância (em cm^2) das alturas de todos os trabalhadores desta carreira profissional é igual a

a) 232.

b) 225.

c) 228.

d) 196.

e) 240.

24. (FGV/AL-BA/2014) A média das idades de um grupo de 4 amigos é de 36 anos, e o desvio padrão é igual a 2. Daqui a cinco anos, a média e a variância das idades desse grupo serão iguais a:

a) 41 e 4.

b) 41 e 50.

c) 56 e 2.

d) 56 e 50.

e) 56 e 200.

25. (CESGRANRIO/BNDES/2013) Em um departamento de uma empresa, o gerente decide dar um aumento a todos os empregados, dobrando o salário de todos eles.

Em relação às estatísticas dos novos salários, considere as afirmativas abaixo.

I - A média dobra.

II - A variância dobra.

III - A moda dobra.

É correto o que se afirma em

a) I, apenas

b) II, apenas

c) I e III, apenas

d) II e III, apenas

e) I, II e III

26. (FCC/TRF 2ª Região/2012) A soma dos quadrados dos valores dos elementos de uma população de tamanho 20 é igual a 65,6 e o respectivo desvio padrão igual a 0,2. A média aritmética dos elementos desta população é igual a

a) 0,8.

b) 1,2.

c) 1,8.

d) 2,4.

e) 3,0.

27. (FCC/TRT 6ª Região/2012) Em um censo realizado em uma empresa, verificou-se que a média aritmética dos salários de seus empregados foi igual a R$ 2.000,00 com um desvio padrão igual a R$ 125,00. Analisando, separadamente, o grupo de todos os empregados do sexo masculino e o grupo de todos os empregados do sexo feminino obteve-se as seguintes informações:

Grupo	Média Aritmética (R$)	Desvio Padrão (R$)
Homens	2.000,00	150,00
Mulheres	2.000,00	100,00

A porcentagem de empregados do sexo masculino na empresa é de

a) 40%.

b) 45%.

c) 48%.

d) 50%.

e) 52%.

28. (CESGRANRIO/BNDES/2010) Em uma pesquisa de preços de determinado produto, foram obtidos os valores, em reais, de uma amostra aleatória colhida em 6 estabelecimentos que o comercializam.

Ação da Empresa	Resultado
P	5,00
Q	8,00
R	6,00
S	6,00
T	4,00
U	7,00

A variância dessa amostra é

a) 1,50

b) 1,75

c) 2,00

d) 2,25

e) 2,50

29. (FGV/SEFAZ-RJ/2010) A média, a mediana e a variância das idades de um grupo de vinte pessoas são, hoje, iguais, respectivamente, a 34, 35 e 24. Daqui a dez anos, os valores da média, da mediana e da variância das idades dessas pessoas serão, respectivamente:

a) 44, 35 e 34.

b) 44, 45 e 12.

c) 44, 45 e 24.

d) 34, 35 e 12.

e) 44, 45 e 124.

30. (FGV/SEAD-AP/2010) Os dados a seguir são as quantidades de empregados de cinco pequenas empresas: 6, 5, 8, 5, 6.

A variância da quantidade de empregados dessas cinco empresas é igual a:

a) 0,8.

b) 1,2.

c) 1,6.

d) 2,0.

e) 2,4.

31. (FGV/SEN/2008) A média

$$\left(\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i\right)$$

e a variância amostral

$$\left(s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)$$

de um conjunto de 20 observações são, respectivamente, 5 e 1. Uma nova observação, de valor igual a 5, foi acrescentada ao conjunto inicial, passando-se a ter 21 valores. A nova variância amostral será igual a:

a) 1,10.

b) 1,05.

c) 1,00.

d) 0,95.

e) 0,90.

Gabarito

Variância

1. CERTO
2. CERTO
3. CERTO
4. ERRADO
5. LETRA D
6. LETRA E
7. CERTO
8. CERTO
9. CERTO
10. ERRADO
11. CERTO
12. LETRA A
13. LETRA C
14. LETRA B
15. LETRA C
16. LETRA D
17. LETRA A
18. LETRA E
19. LETRA D
20. LETRA A
21. LETRA A
22. LETRA E
23. LETRA E
24. LETRA A
25. LETRA C
26. LETRA C
27. LETRA B
28. LETRA C
29. LETRA C
30. LETRA B
31. LETRA D

LISTA DE QUESTÕES

Desvio-Padrão

1. (CESPE/PC PB/2022 - ADAPTADA)

Situação hipotética 17A4-I

Um padrão de referência possui concentração de 25 mg/mL da substância X. Um técnico, ao calibrar dois aparelhos que medem a concentração desta substância X, fez medidas durante 5 dias (amostra 1 no dia 1, amostra 2 no dia 2, e assim por diante) e encontrou os seguintes valores.

Aparelho A

Amostra	Concentração (mg/ml)
1	21
2	22
3	21
4	20
5	21

Aparelho B

Amostra	Concentração (mg/ml)
1	29
2	25
3	21
4	24
5	26

Considerando os dados obtidos na situação hipotética 17A4-I, os valores para a média e desvio-padrão dos aparelhos A e B são

a) médiaA= 21 mg/mL; desvio-padrãoA= 0,71 mg/mL; médiaB= 25 mg/mL; desvio-padrãoB= 2,91 mg/mL;

b) médiaA= 21 mg/mL; desvio-padrãoA= 2 mg/mL; médiaB= 25 mg/mL; desvio-padrãoB= 10 mg/mL;

c) médiaA= 21 mg/mL; desvio-padrãoA= 0,63 mg/mL; médiaB= 24 mg/mL; desvio-padrãoB= 3,58 mg/mL;

d) médiaA= 21 mg/mL; desvio-padrãoA= 2 mg/mL; médiaB= 24 mg/mL; desvio-padrãoB= 4 mg/mL;

e) médiaA= 21 mg/mL; desvio-padrãoA= 0,71 mg/mL; médiaB= 24 mg/mL; desvio-padrãoB= 4 mg/mL;

2. (CESPE/COREN SE/2021) Considere que os tempos de espera X e de atendimento Y, ambos em minutos, para determinado serviço ambulatorial se relacionem como Y = 2X – 1. Se o desvio padrão de X for igual a 2 minutos, então o desvio padrão de Y, em minutos, será igual a

a) 2.

b) 5.

c) 3.

d) 4.

3. (CESPE/PC SE/2021) Com base em uma amostra aleatória simples de tamanho n = 16 retirada de uma população normal com média desconhecida μ e variância $\sigma^2 = 9$, deseja-se testar a hipótese nula $H_0: \mu = 0$ contra a hipótese alternativa $H1: \mu \neq 0$ por meio da estatística $\sqrt{n}\bar{X}/\sigma$, na qual $\bar{X}$ denota a média amostral.

Com respeito a esse teste de hipóteses, julgue o item a seguir, sabendo que o valor da média amostral observado na amostra foi igual a 1 e que, relativo a esse teste, o P-valor foi igual a 0,18.

O desvio padrão da média amostral $\bar{X}$ é igual a 0,75.

4. (CESPE/CBM AL/2021) Determinado dado tetraédrico (dado em formato de tetraedro regular), com vértices numerados de 1 a 4, foi lançado 21 vezes, de modo que o resultado do lançamento desse dado correspondia ao vértice voltado para cima. A tabela seguinte mostra a frequência com que se obteve cada resultado.

Resultado	Quantidade de lançamentos
1	2

2	5
3	5
4	9

Com base nessa situação hipotética, julgue o item a seguir.

O desvio padrão dos resultados é superior a 1.

5. (FGV/Pref. Paulínia/2021) Em uma turma de 10 alunos, as notas dos alunos em uma avaliação foram:

6	7	7	8	8	8	8	9	9	10

O desvio padrão dessa lista de notas é, aproximadamente,

a) 0,8.

b) 0,9.

c) 1,1.

d) 1,3.

e) 1,5.

6. (FGV/IMBEL/2021) Considere os números 1, 2, 3, 6, 8.

O desvio padrão dessa lista de números é, aproximadamente, igual a

a) 1,7.

b) 2,1.

c) 2,6.

d) 3,0.

e) 3,4.

7. (VUNESP/PB Saúde/2021) Seja P_1 uma população formada pelos salários dos 20 funcionários, em R\$ 1.000,00, de uma empresa X, e seja P_2 a população formada pelos salários dos 15 funcionários, em R\$ 1.000,00, de uma outra empresa Y. As somas dos quadrados de todos os elementos de P_1 e P_2, em $(R\$\, 1.000,00)^2$, são iguais a 323,2 e 375,6, respectivamente. O coeficiente de variação de P_1 é igual a 10%, e a média aritmética dos salários de P_1 é igual a 80% da média aritmética dos salários de P_2. Então, o desvio padrão de P_2 é igual a

a) R$ 200,00.

b) R$ 300,00.

c) R$ 400,00.

d) R$ 500,00.

e) R$ 600,00.

8. (VUNESP/EsFCEx/2020) A produção de n unidades de certa peça foi dividida em lotes iguais, com o mesmo número de unidades por lote. Para uma análise sobre peças defeituosas, o Controle de Qualidade da empresa examinou os 40 primeiros lotes produzidos, obtendo o seguinte número de peças defeituosas por lote:

Peças defeituosas	Frequência
0 |— 2	4
2 |— 4	10
4 |— 6	12
6 |— 8	8
8 |— 10	6

O desvio padrão dessa distribuição de frequência é igual a

a) 2,80 peças.

b) 3,20 peças.

c) 3,40 peças.

d) 2,40 peças.

e) 2,20 peças.

9. (VUNESP/EBSERH/2020) Um aluno tirou as seguintes notas ao longo do semestre: 4, 8, 6, 1 e 6. A média, a mediana e o desvio padrão foram, respectivamente:

a) 5; 6 e 5,6.

b) 5; 5 e 5,6.

c) 5; 6 e 2,4.

d) 5; 6 e 31,6.

e) 6; 6 e 5,6.

10. (CESPE/Pref. São Cristóvão/2019) A tabela seguinte mostra a distribuição das idades dos 30 alunos da turma A do quinto ano de uma escola de ensino fundamental.

Idade (em anos)	9	10	11	12	13	14
Quantidade de estudantes	6	22	0	1	0	1

A partir dessa tabela, julgue o item.

O desvio padrão das idades é inferior a 1 ano.

11. (FCC/BANRISUL/2019) Uma população é formada por 4 elementos, ou seja, {4, 5, 5, 8}. O coeficiente de variação, definido como o resultado da divisão do respectivo desvio padrão pela média aritmética da população, é igual a

a) 3/11.

b) 9/22.

c) 3/22.

d) 9/11.

e) 1/5.

12. (FCC/Pref. Recife/2019) Uma população é formada pelos salários dos empregados de uma empresa. Decide-se dar um aumento de 10% sobre todos os salários mais um adicional fixo de R$ 500,00 para todos os salários. Com relação às medidas de tendência central e de dispersão é correto afirmar que a nova população formada terá

a) um desvio padrão igual ao desvio padrão da população anterior multiplicado por 1,10 acrescido de R$ 500,00.

b) uma variância igual à variância da população anterior multiplicada por 1,21 acrescida de 250.000 (R$)2.

c) uma média aritmética igual à média aritmética da população anterior acrescida de R$ 500,00.

d) uma mediana igual à mediana da população anterior acrescida de R$ 500,00.

e) um desvio padrão igual ao desvio padrão da população anterior multiplicado por 1,10 e uma variância igual à variância da população anterior multiplicada por 1,21.

13. (FCC/Pref. Recife/2019) Considere uma população P formada por números estritamente positivos. Com relação às medidas de tendência central e de dispersão é correto afirmar que

a) multiplicando todos os elementos de P por 16, o desvio padrão da nova população é igual ao desvio padrão de P multiplicado por 4.

b) dividindo todos os elementos de P por 2, a variância da nova população é igual a variância de P multiplicada por 0,25.

c) adicionando uma constante K > 0 a todos os elementos de P, a média aritmética e a variância da nova população formada são iguais a média aritmética e desvio padrão de P, respectivamente.

d) a variância e o desvio padrão de P são iguais somente no caso em que todos os elementos de P são iguais.

e) subtraindo uma constante K > 0 de todos os elementos de P, o desvio padrão e a média aritmética da nova população são iguais ao desvio padrão e média aritmética de P subtraídos de K, respectivamente.

14. (CESPE/IFF/2018)

Foram feitas dez medidas do comprimento da caneta mostrada na figura. Os valores dessas medidas estão expressos na tabela a seguir.

medida	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
comprimento (mm)	136	135	135	137	134	135	136	135	136	135

Com base nessas informações, é correto afirmar que o valor do desvio padrão, em mm, desse experimento é igual a

a) 0,00.

b) 0,64.

c) 0,71.

d) 0,80.

e) 0,84.

15. (FCC/ARTESP/2017) O departamento de operações de uma autarquia do Estado fez um levantamento do número de acidentes em um determinado trecho de rodovia no ano de 2016, conforme tabela a seguir.

Mês	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Nº de Acidentes	36	28	12	5	3	2	2	4	9	11	22	38

Os números indicam que há uma dispersão significativa, portanto, o desvio padrão para esta amostra é representado por

a) 13,30.

b) 14,33.

c) 12,74.

d) 10,40.

e) 11,50.

16. (FCC/TRT 20ª Região/2016) Em uma associação de determinada carreira profissional é realizado um censo em que foram apurados os salários de todos os seus 320 associados em número de salários mínimos (S.M.). O coeficiente de variação correspondente foi de 16% e a soma dos quadrados de todos os salários, em (S.M.)2, foi de 8.204,80. O desvio padrão dos salários destes associados é, em S.M., de

a) 0,80

b) 0,64

c) 0,96

d) 0,40

e) 1,60

17. (FGV/IBGE/2016) A partir de uma amostra de tamanho 2n+1, sendo n um número inteiro, elaborou-se a distribuição de frequência de tal forma que apenas os dados grupados ficaram disponíveis. Apesar disso, é possível determinar com certeza a classe à qual pertence o valor exato:

a) Da moda;

b) Da mediana;

c) Da média;

d) Dos quartis;

e) Do desvio padrão.

18. (FCC/CNMP/2015) Em uma empresa, 55% dos empregados são do sexo masculino e a média aritmética dos salários de todos os empregados da empresa é igual a R$ 3.000,00. Sabe-se que a média aritmética dos salários dos empregados do sexo masculino é igual a média aritmética dos salários dos empregados do sexo feminino, sendo que os coeficientes de variação são iguais a 10% e 15%, respectivamente. O desvio padrão dos salários de todos os empregados da empresa é, em R$, de

a) 360,00.

b) 375,00.

c) 367,50.

d) 390,00.

e) 420,00.

19. (FCC/MPE AM/2013) Seja x_1, x_2,..., x_9 uma amostra de 9 observações da variável X. Sabe-se que:

$$\sum_{1}^{9} X_i = 72 \quad \sum_{1}^{9} X_i^2 = 576,5$$

Nessas condições, o desvio padrão dessa amostra é

a) 0,50

b) 0,40

c) 0,15

d) 0,80

e) 0,25

20. (FCC/SERGAS/2013) A tabela abaixo apresenta a distribuição de frequências relativas dos salários, em número de salários mínimos (S.M.), dos 100 funcionários de uma empresa

Classes de salários (em S.M)	Frequências relativas
1 |— 3	0,3
3 |— 5	0,4
5 |— 7	0,3

O valor do desvio padrão desses 100 funcionários, considerado como desvio padrão populacional e obtido por meio dessa tabela, calculado como se todos os valores de cada classe de salários coincidissem com o ponto médio da referida classe, em número de S.M., é

a) $\sqrt{1,2}$

b) $\sqrt{2,2}$

c) $\sqrt{2}$

d) $\sqrt{1,8}$

e) $\sqrt{2,4}$

21. (FGV/TJ-AM/2013) Em relação à medida de desvio padrão, analise as afirmativas a seguir.

I. Tal medida apresenta a propriedade adimensional.

II. Tal medida mostra a dispersão dos dados em relação à média.

III. Tal medida nunca é negativa.

Assinale:

a) Se apenas a afirmativa II estiver correta.

b) Se apenas as afirmativas I e III estiverem corretas.

c) Se apenas as afirmativas II e III estiverem corretas.

d) Se apenas as afirmativas I e II estiverem corretas.

e) Se todas as afirmativas estiverem corretas.

22. (FGV/SEFAZ-RJ/2011) O desvio-padrão da população {2; 4; 2; 4; 2; 4; 2; 4} é

a) 1,5 .

b) 1,0 .

c) 2,5 .

d) 2,0 .

e) 3,0 .

23. (CESGRANRIO/BACEN/2009) Considere as informações a seguir para responder à questão.

A viabilidade financeira do projeto de uma microempresa leva em consideração dados históricos de 100 projetos semelhantes.

A tabela abaixo mostra a distribuição de frequências do VPL - Valor Presente Líquido (valores em milhões de reais) de um conjunto de microempresas similares.

VPL	Frequência relativa
$-10 < x \leq 0$	10%
$0 < x \leq 10$	80%
$10 < x \leq 20$	10%

Segundo os dados históricos, o valor, em milhões de reais, que mais se aproxima do desvio padrão do VPL da microempresa é

a) 1

b) 2

c) 2,5

d) 4

e) 4,5

24. (CESGRANRIO/BNDES/2007) O enunciado abaixo refere-se à questão.

Um grupo é formado por 10 pessoas, cujas idades são:

17 19 19 20 20 20 20 21 22 22

Seja μ a média aritmética das idades e σ seu desvio padrão. O número de pessoas desse grupo cujas idades pertencem ao intervalo $[\mu - \sigma, \mu + \sigma]$ é

(Considere $\sqrt{2} = 1,4$)

a) 9

b) 8

c) 7

d) 6

e) 5

Gabarito

Desvio-Padrão

1. LETRA A
2. LETRA D
3. CERTO
4. CERTO
5. LETRA C
6. LETRA C
7. LETRA A
8. LETRA D
9. LETRA C
10. CERTO
11. LETRA A
12. LETRA E
13. LETRA B
14. LETRA E
15. LETRA A
16. LETRA A
17. LETRA B
18. LETRA B
19. LETRA E
20. LETRA E
21. LETRA C
22. LETRA B
23. LETRA E
24. LETRA C

LISTA DE QUESTÕES

Coeficiente de Variação (ou Dispersão Relativa)

1. (CESPE/TELEBRAS/2022) Com respeito ao conjunto de dados {0, 0, 1, 1, 1, 3}, julgue o item que se segue.

O coeficiente de variação é igual ou superior a 1,2.

2. (CESPE/TELEBRAS/2022 - ADAPTADA) Com respeito ao conjunto de dados $\{5a, 2a, 2a\}$, em que a representa uma constante não nula, julgue o próximo item.

O coeficiente de variação independe do valor da constante a.

3. (FGV/PC-AM/2022) Suponha que um pesquisador tenha as seguintes informações de uma amostra de dados:

Média = 5

Variância = 25

Soma dos desvios absolutos em relação à média = 10

Tamanho da amostra = 5

Assim, o coeficiente de variação dessa amostra em termos decimais será igual a

a) 1.

b) 2.

c) $\sqrt{5}$.

d) 5.

e) 10.

4. (CESGRANRIO/BB/2021) Um pesquisador recebeu os dados de uma amostra de tamanho 100 de uma população e calculou a média amostral μ, o desvio padrão amostral σ e o coeficiente de variação amostral $CV = \frac{\sigma}{\mu}$. Antes de iniciar a análise, ele foi informado de que os dados dessa amostra estavam todos errados, mas que podiam ser corrigidos somando-se 3 a cada um dos dados que recebeu.

Após fazer tal correção, o valor do coeficiente de variação amostral passou a ser

a) $\frac{3\sigma}{\mu+3}$

b) $\frac{300\sigma}{\mu+300}$

c) $\frac{\sigma}{\mu+3}$

d) $\frac{\sigma}{\mu+300}$

e) $\frac{\sigma}{\mu+0{,}03}$

5. (CESPE/SERPRO/2021) Considerando que o número X de erros registrados em determinado tipo de código computacional siga uma distribuição binomial com média igual a 4 e variância igual a 3, julgue o item a seguir.

O coeficiente de variação da distribuição de erros X é igual a 3.

6. (VUNESP/EsFCEx/2020) Sejam P_1 e P_2 duas populações independentes formadas por números estritamente positivos com tamanhos 20 e 25, respectivamente. O coeficiente de variação de P_1 é igual a 50% com a soma dos quadrados de seus elementos igual a 2.500. Sabe-se que a soma dos quadrados dos elementos de P_2 é igual a 2.900 e a média aritmética é igual a média aritmética de P_1. O coeficiente de variação de P_2 é igual a

a) 0,40.

b) 0,20.

c) 0,60.

d) 0,64.

e) 0,80.

7. (CESGRANRIO/BB/2018) Há dez anos a média das idades, em anos completos, de um grupo de 526 pessoas era de 30 anos, com desvio padrão de 8 anos.

Considerando-se que todas as pessoas desse grupo estão vivas, o quociente entre o desvio padrão e a média das idades, em anos completos, hoje, é

a) 0,45

b) 0,42

c) 0,20

d) 0,27

e) 0,34

8. (CESGRANRIO/LIQUIGÁS/2018) Uma amostra $(x_1, x_2, ..., x_{10})$ de tamanho 10 de uma população nos forneceu os seguintes valores

$$\sum_{i=1}^{10} x_i = 20 \ e \ \sum_{i=1}^{10} x_i^2 = 140.$$

Qual o valor do coeficiente de variação amostral?

a) $\frac{7}{20}$

b) $\frac{5}{3}$

c) $\frac{1}{7}$

d) $\frac{1}{6}$

e) 6

9. (CESPE/IPHAN/2018) Cinco municípios de um estado brasileiro possuem as seguintes quantidades de patrimônios históricos: {2, 3, 5, 3, 2}. Admitindo que a média e o desvio-padrão desse conjunto de valores sejam iguais a 3 e 1,2, respectivamente, julgue o item seguinte.

O coeficiente de variação é superior a 0,3 e inferior a 0,5.

10. (FCC/TCE-RS/2018) Uma população é formada por 100 números estritamente positivos x_i com $1 \leq i \leq 100$, ou seja, $\{x_1, x_2, x_3, \ldots, x_{100}\}$, em que x_i representa a renda familiar anual da família i, em milhares de reais.

Dados:

$\sum_{i=1}^{100} x_i = 6.400$ **mil reais e**

$\sum_{i=1}^{100} x_i^2 = 467.200$ **(mil reais)^2**

O coeficiente de variação desta população é igual a

a) 37,5%

b) 18,0%

c) 32,5%

d) 24,0%

e) 27,5%

11. (FGV/IBGE/2017) Alguns economistas estão discutindo sobre a volatilidade dos preços em duas economias, relativamente parecidas, tendo como moedas peras (A) e maçãs (B). Sabe-se que as médias dos preços são 100 peras e 120 maçãs, respectivamente. É fornecido, ainda, o desvio-padrão dos preços em A, igual a 25 peras, e a variância em B, igual a 400 maçãs ao quadrado.

Considerando as principais medidas estatísticas de dispersão como medidas de volatilidade, é correto afirmar que:

a) O desvio padrão dos preços em A é inferior ao de B;

b) A taxa de conversão da moeda A para B é de 1,2;

c) A taxa de inflação em A deve ser menor do que em B;

d) Os preços em B são, em média, mais caros do que em A;

e) A medida adimensional de dispersão de A é superior à de B.

12. (FGV/IBGE/2016) As principais medidas de dispersão utilizadas na estatística são a amplitude (A), a variância (Var), o desvio padrão (DP), o coeficiente de variação (CV) e o desvio-interquartílico (DI). Sobre o tema, é correto afirmar que:

a) As medidas acima listadas têm seus valores dependentes, na íntegra, dos valores da distribuição amostral;

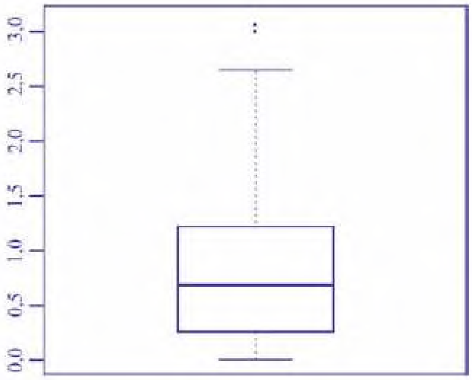
b) A variância apresenta a vantagem de ser diretamente comparável com os valores da distribuição;

c) É possível afirmar que var(x) ≥ dp(x);

d) O desvio-interquartílico é sempre superior ou no mínimo igual à amplitude;

e) O coeficiente de variação é uma medida invariante às mudanças de escala.

13. (CESPE/TCE-PA/2016)



média amostral	0,80
desvio padrão amostral	0,70
primeiro quartil	0,25
mediana	0,70
terceiro quartil	1,20
mínimo	0
máximo	3,10

Um indicador de desempenho X permite avaliar a qualidade dos processos de governança de instituições públicas. A figura mostra, esquematicamente, a sua distribuição, obtida mediante estudo amostral feito por determinada agência de pesquisa. A tabela apresenta estatísticas descritivas referentes a essa distribuição.

Com base nessas informações, julgue o item a seguir.

O coeficiente de variação da distribuição de X é inferior a 0,8.

14. (FCC/TRE RR/2015) Em uma escola é realizado um censo apurando-se as alturas de todos os 180 estudantes em centímetros (cm). A média aritmética das alturas dos 100 estudantes do sexo masculino foi igual a dos 80 estudantes do sexo feminino. Se X_i representa a altura do i-ésimo estudante do sexo masculino e Y_j a altura do j-ésimo estudante do sexo feminino, obteve-se

$$\sum_{i=1}^{100} x_1^2 = 2.570.000\ cm^2 \ e\ \sum_{j=1}^{80} y_j^2 = 2.084.080\ cm^2$$

Se o desvio padrão das alturas dos estudantes do sexo masculino foi igual a 10 cm, o coeficiente de variação considerando todos os estudantes desta escola é, em %, de

a) 18.

b) 12.

c) 10.

d) 16.

e) 15.

15. (VUNESP/TJ-SP/2015) Leia o texto a seguir para responder à questão.

Na tabela a seguir, são apresentados os dados colhidos de uma amostra de 9 vendedores de um grande magazine, para comparação entre o tempo T de experiência do vendedor (em anos) e seu movimento mensal V de vendas (em mil). Na tabela, estão registrados também a média mensal de tempo de experiência, a média mensal de vendas, o desvio padrão (DP) de cada variável, além do coeficiente de correlação r da relação entre as variáveis T e V.

FUNC	T (Anos)	V (vendas)
A	4,2	12,5
B	6,1	15,5
C	4	12,9
D	5,4	14,2
E	1,2	8,2
F	1,5	9,1
G	3,2	9,9
H	1	6,4
I	5	12,1
MÉDIA	**3,51**	**11,20**
DP	**1,90**	**2,98**
r	**0,96**	

O valor mais próximo do coeficiente de variação da variável T é

a) 0,28.

b) 0,34.

c) 0,48.

d) 0,54.

e) 0,96

16. (VUNESP/Pref. SP/2015) O indicador da variabilidade de uma determinada distribuição de um conjunto de dados é denominado

a) Mediana.

b) Moda.

c) Coeficiente de variação.

d) Quociente de correlação.

e) Coeficiente de determinação.

17. (CESGRANRIO/EPE/2014) Um pesquisador está interessado em comparar a variabilidade de duas variáveis com médias diferentes e desvios padrões diferentes, presentes num dado estudo estatístico.

A medida estatística adequada a ser usada nesse contexto é a(o)

a) covariância

b) diferença entre a maior e a menor variância

c) razão entre o maior e o menor desvios padrões

d) coeficiente de variação

e) coeficiente de correlação

18. (CESGRANRIO/IBGE/2014) Os dados a seguir foram obtidos de empregados de uma empresa com três fábricas: I, II e III. A variável de interesse é salário.

Empresa	Média	Desvio padrão
Fábrica I	1185	630,49
Fábrica II	600	355,97
Fábrica III	2150	1106,16

Comparando-se a variabilidade de salários em relação ao salário médio das três fábricas, através de seus coeficientes de variação, conclui-se que a variabilidade da fábrica

a) I é menor apenas do que a da fábrica III.

b) II é menor apenas do que a da fábrica I.

c) II é menor apenas do que a da fábrica III.

d) II é menor do que as das outras duas fábricas.

e) III é menor do que as das outras duas fábricas.

19. (FCC/TRT 13ª Região/2014) Seja X uma população $\{X_1, X_2, X_3, \ldots, X_{100}\}$ formada por 100 números estritamente positivos com um desvio padrão igual a 4 e com a soma dos quadrados de todos estes 100 números igual a 41.600. Seja Y uma outra população $\{Y_1, Y_2, Y_3, \ldots, Y_{50}\}$ formada por 50 números também estritamente positivos com uma média igual a da população anterior e com a soma dos quadrados de todos estes 50 números igual a 20.200. Os coeficientes de variação de X e de Y

a) são, ambos, iguais a 20%.

b) são iguais a 20% e 5%, respectivamente.

c) são, ambos, superiores a 15%.

d) apresentam uma diferença de valor absoluto igual a 10%.

e) apresentam um produto igual a 4%.

20. (FCC/TRT 16ª Região/2014) A média aritmética dos salários, em março de 2014, dos empregados em uma empresa é igual a R$ 2.500,00 com um coeficiente de variação igual a 9,6%. Decide-se aumentar os salários de todos os empregados, tendo que escolher uma entre as duas opções abaixo:

Opção I: Reajuste de todos os salários, em março de 2014, em 10% mais um abono fixo de R$ 250,00 para todos os salários.

Opção II: Reajuste de todos os salários, em março de 2014, em x% mais um abono fixo de R$ 200,00 para todos os salários.

Existe um valor para x tal que se for escolhida a opção II, a nova média aritmética passa a ser igual à nova média aritmética caso fosse escolhida a opção I. Nesta situação, o novo coeficiente de variação com a escolha da opção II passa a ser de

a) 8,00%.

b) 8,80%.

c) 9,00%.

d) 8,96%.

e) 9,60%.

21. (FGV/AL-BA/2014) A tabela a seguir mostra média e desvio padrão das notas dos alunos em um exame nacional em cinco estados diferentes:

	Média	**Desvio padrão**
Estado I	500	100
Estado II	600	120
Estado III	500	140
Estado IV	450	120
Estado V	600	100

Assinale a opção que indica o Estado que apresentou o menor coeficiente de variação das notas.

a) I

b) II

c) III

d) IV

e) V

22. (CESGRANRIO/IBGE/2013) Seja X uma variável aleatória com distribuição normal cuja média é μ e o desvio padrão é σ.

Se Y = 2X - 1 tem distribuição normal com média 5 e variância 20, o coeficiente de variação populacional $\frac{\sigma}{\mu}$ vale

a) $\frac{\sqrt{42}}{6}$

b) $\frac{\sqrt{21}}{6}$

c) $\frac{\sqrt{5}}{3}$

d) $\frac{\sqrt{39}}{9}$

e) $\frac{4\sqrt{5}}{9}$

23. (CESGRANRIO/BB/2013) A variância de um conjunto de dados é 4 m^2.

Para o mesmo conjunto de dados foram tomadas mais duas medidas de variabilidade:

a diferença entre o terceiro e o primeiro quartil e o coeficiente de variação.

Esses dois valores caracterizam-se, respectivamente, por

a) possuírem unidades de medida m^2 e m.

b) possuírem unidades de medida m e m^2.

c) ser adimensional e possuir unidade de medida m^2.

d) possuir unidade de medida m e ser adimensional.

e) possuir unidade de medida m^2 e ser adimensional.

24. (CESGRANRIO/BNDES/2013) Quatro variáveis são utilizadas em um modelo de previsão da quantidade produzida de uma determinada commodity agrícola. São elas:

• temperatura, em graus Celsius

• quantidade de fertilizante, em toneladas

• variação dos preços praticados no mercado internacional, em %

• quantidade produzida de um produto similar, em toneladas

Para determinar qual dessas variáveis apresenta a maior variabilidade, deve-se utilizar

a) apenas a média de cada uma das variáveis

b) apenas a variância de cada uma das variáveis

c) apenas o desvio padrão de cada uma das variáveis

d) a relação $\frac{desvio\ padrão}{média}$ de cada uma das variáveis

e) a relação $\frac{variância}{média}$ de cada uma das variáveis

25. (FCC/TRT 5ª Região/2013) Seja Xi um elemento de uma população de tamanho 20, com $1 \leq i \leq 20$. Sabe-se que

$$\sum_{i=1}^{20} X_i = 300 \ e \ \sum_{i=1}^{20} X_i^2 = 4.731,2$$

O coeficiente de variação desta população apresenta um valor c, tal que

a) c < 21,0%.

b) 21,0% ≤ c < 21,5%.

c) 21,5% ≤ c < 22,0%.

d) 22,0% ≤ c < 22,5%.

e) c ≥ 22,5%.

26. (FCC/TRT 6ª Região/2012) Um levantamento realizado em uma indústria revelou que o diâmetro médio de todas as 40 peças, marca Alpha, em estoque, é igual a 10 cm. Sabendo-se que

a soma dos quadrados das medidas dos diâmetros de todas estas 40 peças apresenta o valor de 4.078,40 cm^2, então o coeficiente de variação correspondente é igual a

a) 10,0%.

b) 12,0%.

c) 12,5%.

d) 14,0%.

e) 15,0%.

27. (CESGRANRIO/BACEN/2009) Considere as informações a seguir para responder à questão.

A viabilidade financeira do projeto de uma microempresa leva em consideração dados históricos de 100 projetos semelhantes.

A tabela abaixo mostra a distribuição de frequências do VPL - Valor Presente Líquido (valores em milhões de reais) de um conjunto de microempresas similares.

VPL	Frequência relativa
$-10 < x \leq 0$	10%
$0 < x \leq 10$	80%
$10 < x \leq 20$	10%

Um projeto alternativo para o investidor apresenta um VPL esperado, em reais, de 6 milhões e um risco (desvio padrão) de 2 milhões. Pela ótica do risco relativo, qual o melhor investimento, a microempresa ou o projeto alternativo?

a) A microempresa, pois apresenta um Coeficiente de Variação maior.

b) A microempresa, pois apresenta um Coeficiente de Variação menor.

c) O projeto alternativo, pois apresenta um Coeficiente de Variação maior.

d) O projeto alternativo, pois apresenta um Coeficiente de Variação menor.

e) É indiferente, pois os investimentos apresentam Coeficientes de Variação iguais.

28. (FGV/SEN/2008) O coeficiente de variação amostral (em porcentagem) de um conjunto de salários é 110%.

Se os salários desse conjunto forem reajustados em 20%, o novo coeficiente de variação amostral será:

a) 110%.

b) 112,2%.

c) 114,2%.

d) 122%.

e) 130%.

29. (FGV/SEFAZ-RJ/2008) Uma companhia utiliza um sistema de avaliação de desempenho de seus funcionários por meio de dois indicadores de performance: Qualidade das tarefas e a Tempestividade com que as tarefas são realizadas.

Os funcionários receberam, na última avaliação, as medidas indicadas na tabela a seguir:

Medidas	Indicador	
	Qualidade	Tempestividade
Média	50	25
Desvio-Padrão	10,0	6,0
Coeficiente de Variação %	20	24

Com base na tabela, é correto afirmar que:

a) a média aritmética não é uma boa medida para representar a performance dos funcionários em face do elevado nível de dispersão das avaliações.

b) as avaliações da Qualidade foram mais dispersas do que as avaliações da Tempestividade.

c) as avaliações da Qualidade foram mais homogêneas do que as da Tempestividade.

d) os funcionários demoram mais para realizar as tarefas, mas a qualidade das tarefas é melhor.

e) nada se pode afirmar sem o conhecimento do tamanho da amostra.

30. (FGV/SEFAZ RJ/2007) Uma amostra de 100 servidores de uma repartição apresentou média salarial de R$ 1.700,00 com uma dispersão de R$ 240,00. Pode-se afirmar que:

a) a média aritmética não é uma boa medida para representar a amostra em função do elevado valor do desvio-padrão.

b) a melhor medida para representar a amostra é a remu-neração por unidade de desvio-padrão.

c) o salário mediano representaria melhor a amostra de-vido ao alto nível de heterogeneidade dos salários na amostra.

d) a amostra não é suficientemente grande para analisar-mos o valor encontrado para a média dos salários.

e) a média aritmética pode perfeitamente representar os salários da amostra pelo fato de esta apresentar uma dispersão relativa inferior a 20%.

Gabarito

Coeficiente de Variação (ou Dispersão Relativa)

1. ERRADO
2. ERRADO
3. LETRA A
4. LETRA C
5. ERRADO
6. LETRA A
7. LETRA C
8. LETRA B
9. CERTO
10. LETRA A
11. LETRA E
12. LETRA E
13. ERRADO
14. LETRA C
15. LETRA D
16. LETRA C
17. LETRA D
18. LETRA E
19. LETRA D
20. LETRA D
21. LETRA E
22. LETRA C
23. LETRA D
24. LETRA D
25. LETRA E
26. LETRA D
27. LETRA D
28. LETRA A
29. LETRA C
30. LETRA E

LISTA DE QUESTÕES

Variância Relativa

1. (FCC/SEFAZ BA/2019) O coeficiente de variação de Pearson correspondente a uma população P_1 com média aritmética igual a 20 e tamanho 20 é igual a 30%. Decide-se excluir de P_1, em um determinado momento, dois elementos iguais a 11 cada um, formando uma nova população P_2. A variância relativa de P_2 é igual a

a) 10/147.

b) 4/49.

c) 16/147.

d) 8/49.

e) 4/441.

Gabarito

Variância Relativa

1. LETRA A