28 – Considere a seguinte distribuição de freqüência relativa à idade dos freqüentadores de uma academia de ginástica, no horário da tarde:

Faixa Etária (Classe)	Nº Freqüentadores (n_i)
15 |—— 25	100
25 |—— 35	50
35 |—— 45	40
45 |—— 55	30
Total	**220**

Leia atentamente as afirmativas abaixo, assinale com (**F**) aquelas que considera **FALSAS** e com (**V**) aquelas que considera **VERDADEIRAS** e marque a alternativa que apresenta a seqüência considerada **CORRETA**.

( ) A mediana desses dados é 27 e a média é 30.
( ) A média desses dados é 30 e o desvio-padrão é 118,72.
( ) A mediana desses dados é 30 e a variância é 118,72.
( ) A variância desses dados é 118,72 e a média é 30.

A ( ) V, F, F, V.
B ( ) V, V, F, F.
C ( ) F, F, F, F.
D ( ) F, V, V, F.

<u>RESOLUÇÃO</u>:

<u>Mediana</u>: Podemos encontrar a mediana por fórmula ou por interpolação.

Criando a coluna de freqüência acumulada crescente (F_{AC}), temos:

Faixa Etária (Classe)	Nº Freqüentadores (n_i)	F_{AC}
15 |—— 25	100	100
25 |—— 35	50	150
35 |—— 45	40	190
45 |—— 55	30	220
Total	**220**	**-**

Vemos então que a classe da mediana (onde estará a 110ª observação) será a 2ª classe, cujo limite inferior (L) é 25 e a freqüência absoluta simples é 50.

<u>Por fórmula</u>: $\boxed{\mathbf{Md = L + \frac{\left(\frac{n}{2} - \sum f\right)}{F_{Md}} \cdot h}} \Rightarrow \text{Md} = 25 + \frac{(110-100)}{50} \cdot 10 \Rightarrow \text{Md} = 25 + \frac{100}{50} \Rightarrow$ **Md = 27**;

<u>Por interpolação</u>: Até a classe anterior temos 100 observações e para chegar na mediana precisamos de mais 10 observações. Fazendo a proporção:

$\frac{\text{amplitude da classe}}{\text{Freq. na classe}} = \frac{\text{amplitude procurada}}{\text{Freq. correspondente}} \Rightarrow \frac{10}{50} = \frac{X}{10} \Rightarrow 5X = 10 \Rightarrow X = 2.$

Somando esse valor ao limite inferior da classe da mediana, temos: Md = 2 + 25 ⇒ **Md = 27**;

<u>Para encontrar a **média**</u>: Podemos utilizar o método normal ou o simplificado (reduzido).

<u>Pelo método normal</u>: onde X é o ponto médio de cada classe (pm).

Faixa Etária (Classe)	**Nº Freqüentadores (n_i)**	**X (pm)**	**X·F**
15 ├── 25	100	20	2.000
25 ├── 35	50	30	1.500
35 ├── 45	40	40	1.600
45 ├── 55	30	50	1.500
Total	**220**	**-**	**6.600**

Média: $\overline{X} = \frac{6.600}{220} \Rightarrow \overline{X}$ = **30**;

<u>Pelo método reduzido</u>: Criamos a variável transformada Z, o que facilitará bastante quando formos calcular a variância.

Faixa Etária (Classe)	**Nº Freqüentadores (n_i)**	**Z (pmt)**	**Z·F**
15 ├── 25	100	–1	–100
25 ├── 35	50	0	0
35 ├── 45	40	1	40
45 ├── 55	30	2	60
Total	**220**	**-**	**0**

A variável transformada Z será dada pela transformação: $\boxed{Z = \frac{X - 30}{10}}$ gerando os pontos médios de classe transformados (pmt): –1, 0, 1 e 2.

A média de Z será igual a zero, pois $\overline{Z} = \frac{0}{220}$ = 0.

Para voltar à variável original (X) teremos que fazer: $\boxed{X = 10 \cdot Z + 30}$.

Então, pelas propriedades da média: temos: $\boxed{\overline{X} = 10 \cdot \overline{Z} + 30}$.

Logo, $\overline{X} = 10 \cdot 0 + 30 \Rightarrow \overline{X}$ = **30** (mesmo resultado do processo normal).

Sabendo a mediana (= 27) e a média (= 30), já podemos julgar algumas afirmativas:

(**V**) A mediana desses dados é 27 e a média é 30.
( ) A média desses dados é 30 e o desvio-padrão é 118,72.
(**F**) A mediana desses dados é 30 e a variância é 118,72.
( ) A variância desses dados é 118,72 e a média é 30.

Assim, vemos que a resposta só pode ser *A* ou *B*, pois nas letras *C* e *D* temos a 1ª afirmativa (que é verdadeira) como F. Precisaremos então calcular a variância e para isso faremos o <u>somatório dos quadrados</u>. Podemos usar o processo normal ou o cálculo simplificado da variância (pela variável transformada Z).

<u>Pelo método normal</u>: sendo X é o ponto médio de cada classe (pm).

Faixa Etária (Classe)	**Nº Freqüentadores (n_i)**	**X (pm)**	**X·F**	**X^2·F**
15 ├── 25	100	20	2.000	40.000
25 ├── 35	50	30	1.500	45.000
35 ├── 45	40	40	1.600	64.000
45 ├── 55	30	50	1.500	75.000
Total	**220**	**-**	**6.600**	**224.000**

Variância populacional: $\sigma^2 = \frac{1}{\sum F}\left\{\sum X^2 \cdot F - \frac{\left(\sum X \cdot F\right)^2}{\sum F}\right\}$

$$\sigma^2 = \frac{1}{220}\cdot\left\{224.000 - \frac{(6.600)^2}{220}\right\} \Rightarrow \sigma^2 = \frac{1}{220}\cdot(224.000 - 198.000) \Rightarrow \sigma^2 = \frac{26.000}{220} \Rightarrow \sigma^2 = 118{,}18.$$

Teríamos então:

(**V**) A mediana desses dados é 27 e a média é 30.
(**F**) A média desses dados é 30 e o desvio-padrão é 118,72.
(**F**) A mediana desses dados é 30 e a variância é 118,72.
(**F**) A variância desses dados é 118,72 e a média é 30.

Não temos, entre as opções de resposta, a seqüência V, F, F, F. Mas temos V, F, F, V (letra A).

Então, supondo ser a variância amostral e não populacional, fazemos a correção de Bessel (n – 1) para obter S^2 (variância amostral): $S^2 = \frac{1}{\sum F - 1}\left\{\sum X^2 \cdot F - \frac{\left(\sum X \cdot F\right)^2}{\sum F}\right\}$.

$$S^2 = \frac{1}{220-1}\cdot\left\{224.000 - \frac{(6.600)^2}{220}\right\} \Rightarrow S^2 = \frac{1}{219}\cdot(224.000 - 198.000) \Rightarrow S^2 = \frac{26.000}{219} \Rightarrow \mathbf{S^2 = 118{,}72}.$$

E assim, encontrar a seqüência do gabarito da questão (letra A):

(**V**) A mediana desses dados é 27 e a média é 30.
(**F**) A média desses dados é 30 e o desvio-padrão é 118,72.
(**F**) A mediana desses dados é 30 e a variância é 118,72.
(**V**) A variância desses dados é 118,72 e a média é 30.

Mas, isso se considerarmos a distribuição como uma amostra. Entretanto, o enunciado é dúbio quanto à isso, pelo contrário, tende muito mais para que consideremos a distribuição como uma população, pois refere-se à "idade **dos** freqüentadores" o que sugere **de todos os**, induzindo o candidato a considerar a distribuição como sendo populacional. As afirmativas deveriam trazer, para que não ficasse nenhuma dúvida, "variância amostral" ou o enunciado referir-se à "idade **de alguns** freqüentadores", o que caracterizaria, sem sombra de dúvida, uma amostra. Era possível chegar à letra A porque a letra B traz a 2ª afirmativa como sendo V quando certamente será F, pois seja população ou seja amostra, o desvio padrão será um valor entre 10 e 11 e não 118,72. Assim, a questão não foi elaborada com perfeição, deixando margem para recurso.

Podemos obter o mesmo resultado para a variância (amostral ou populacional) utilizando o <u>método simplificado</u>:

Faixa Etária (Classe)	Nº Freqüentadores (n_i)	Z (pmt)	Z·F	X^2·F
15 ├── 25	100	–1	–100	100
25 ├── 35	50	0	0	0
35 ├── 45	40	1	40	40
45 ├── 55	30	2	60	120
Total	**220**	**-**	**0**	**260**

A variância amostral da transformada Z será: $S_Z^2 = \frac{1}{\sum F - 1}\left\{\sum Z^2 \cdot F - \frac{\left(\sum Z \cdot F\right)^2}{\sum F}\right\}$. Logo:

$$S_Z^2 = \frac{1}{220-1}\cdot\left\{260 - \frac{(0)^2}{220}\right\} \Rightarrow S_Z^2 = \frac{260}{219} \Rightarrow S_Z^2 = 1{,}187215\,.$$

Mas queremos obter a variância da variável X. Isso será fácil, pelas propriedades da variância.

Vimos que a variável transformada Z é dada pela transformação: $\boxed{Z = \frac{X-30}{10}}$ e que para voltar à variável original (X) teremos que fazer: $\boxed{X = 10 \cdot Z + 30}$.

Então, pelas propriedades da variância, sabemos que: se a variável é multiplicada por uma constante, a sua variância fica multiplicada pelo quadrado da constante e quando é acrescida de uma constante a sua variância não se altera.

Logo: $\boxed{S_X^2 = 10^2 \cdot S_Z^2} \Rightarrow S_X^2 = 100 \cdot 1{,}1872 \Rightarrow S_X^2 = 118{,}72$ (mesmo resultado pelo processo normal).

Gabarito: LETRA **A**.

29 – O conjunto de dados abaixo refere-se à venda mensal de um determinado tipo de arroz em um supermercado, no período de 5 anos, em unidades vendidas. Leia atentamente as alternativas a seguir:

250	275	240	280	300	295	245	230	255	265	275	290
310	270	280	240	290	260	305	265	280	250	315	295
290	285	275	290	305	320	285	315	260	325	285	270
295	315	305	285	320	260	285	305	295	300	290	310
325	285	300	305	315	310	290	300	315	300	310	305

I - O 1º quartil desse conjunto de dados é 275 e a mediana é 290.
II - O 1º quartil desse conjunto de dados é 275 e o percentil de ordem 70 é 305.
III - A mediana desse conjunto de dados é 290 e o percentil de ordem 70 é 300.
IV - O 1º quartil desse conjunto de dados é 275 e o percentil de ordem 70 é 300.

Com base nessas alternativas, marque a alternativa **CORRETA**.

A ( ) Apenas as afirmativas I e IV são verdadeiras.
B ( ) As afirmativas I, II e III são verdadeiras.
C ( ) Apenas as afirmativas I e II são verdadeiras.
D ( ) Apenas a afirmativa II é verdadeira.

<u>RESOLUÇÃO</u>:

Questão fácil mas muito trabalhosa, pois para encontrar corretamente as medidas (Q_1, Md, P_{70}), teremos que primeiro organizar os 60 valores observados, criando um rol com os valores ordenados em ordem crescente. Para facilitar, vamos organizar os dados em 4 grupos de 15 elementos (quartis), pois assim ficará fácil de identificar o 1º quartil e a mediana.

230	240	240	245	250	250	255	260	260	260	265	265	270	270	275
275	275	280	280	280	285	285	285	285	285	285	290	290	290	290
290	290	295	295	295	295	300	300	300	300	300	305	305	305	305
305	305	310	310	310	310	315	315	315	315	315	320	320	325	325

1º Quartil → será a média entre 15ª e 16ª observações (ambas são iguais a 275). Logo, Q_1 = 275;

Mediana → será a média entre 30ª e 31ª observações (ambas são iguais a 290). Logo, Md = 290;

70º Percentil = 7º Decil → será a média entre a 42ª e a 43ª observações (= 305). Logo, P_{70} = 305;

E teremos a seguinte seqüência para as afirmativas: **V, V, F, F**, ou seja, apenas a I e II verdadeiras.

Gabarito: LETRA **C**.

30 – Uma prova de matemática foi aplicada a alunos de diversas turmas de uma mesma série de uma determinada escola. Os resultados estão apresentados na tabela abaixo:

Turma	Notas									
A	10	8	9	9	8	10	10	8		
B	9	7	9	7	9	9	7	8	7	
C	10	4	7	5	9	7				
D	6	8	5	7	8	7	8			
E	8	4	3	2	7	9	4	8	8	7

Leia atentamente as afirmativas abaixo, assinale com (**F**) aquelas que considera **FALSAS** e com (**V**) aquelas que considera **VERDADEIRAS** e marque a alternativa que apresenta a seqüência considerada **CORRETA**.

( ) A turma **E** apresenta a menor média e a segunda maior variância.
( ) A turma **A** apresenta a maior média e a menor variância.
( ) Apenas uma das turmas avaliadas apresenta variância superior à média.
( ) As turmas **C** e **D** apresentam a mesma média, mas a turma **C** apresenta menor variabilidade.

A ( ) F, V, V, F.
B ( ) V, F, F, V.
C ( ) V, F, V, V.
D ( ) F, V, V, V.

RESOLUÇÃO:

Vamos primeiro encontrar as médias das amostras para cada uma das turmas.

Turma A → $\overline{A} = \frac{10+8+9+9+8+10+10+8}{8} \Rightarrow \overline{A} = \frac{72}{8} \Rightarrow \overline{A} = 9$;

Turma B → $\overline{B} = \frac{9+7+9+7+9+9+7+8+7}{9} \Rightarrow \overline{B} = \frac{72}{9} \Rightarrow \overline{B} = 8$;

Turma C → $\overline{C} = \frac{10+4+7+5+9+7}{6} \Rightarrow \overline{C} = \frac{42}{6} \Rightarrow \overline{C} = 7$;

Turma D → $\overline{D} = \frac{6+8+5+7+8+7+8}{7} \Rightarrow \overline{D} = \frac{49}{8} \Rightarrow \overline{D} = 7$;

Turma E → $\overline{E} = \frac{8+4+3+2+7+9+4+8+8+7}{10} \Rightarrow \overline{E} = \frac{60}{10} \Rightarrow \overline{E} = 6$;

Encontrando as variâncias amostrais para cada uma das turmas.

Turma A:

A	10	8	9	9	8	10	10	8	**SOMATÓRIO**
$A-\overline{A}$	1	–1	0	0	–1	1	1	–1	**0**
$\left(A-\overline{A}\right)^2$	1	1	0	0	1	1	1	1	**6**

$$S_A^2 = \frac{\sum\left(A-\overline{A}\right)^2}{n-1} \Rightarrow S_A^2 = \frac{6}{8-1} \Rightarrow S_A^2 = \frac{6}{7} \Rightarrow S_A^2 \cong 0{,}86;$$

Turma B:

B	9	7	9	7	9	9	7	8	7	**SOMATÓRIO**
$B-\overline{B}$	1	–1	1	–1	1	1	–1	0	–1	**0**
$(B-\overline{B})^2$	1	1	1	1	1	1	1	0	1	**8**

$$S_B^2=\frac{\sum(B-\overline{B})^2}{n-1}\Rightarrow S_B^2=\frac{8}{9-1}\Rightarrow S_B^2=\frac{8}{8}\Rightarrow S_B^2=1{,}00;$$

Turma C:

C	10	4	7	5	9	7	**SOMATÓRIO**
$C-\overline{C}$	3	–3	0	–2	2	0	**0**
$(C-\overline{C})^2$	9	9	0	4	4	0	**26**

$$S_C^2=\frac{\sum(C-\overline{C})^2}{n-1}\Rightarrow S_C^2=\frac{26}{6-1}\Rightarrow S_C^2=\frac{26}{5}\Rightarrow S_C^2=5{,}20;$$

Turma D:

D	6	8	5	7	8	7	8	**SOMATÓRIO**
$D-\overline{D}$	–1	1	–2	0	1	0	1	**0**
$(D-\overline{D})^2$	1	1	4	0	1	0	1	**8**

$$S_D^2=\frac{\sum(D-\overline{D})^2}{n-1}\Rightarrow S_D^2=\frac{8}{7-1}\Rightarrow S_D^2=\frac{8}{6}\Rightarrow S_D^2\cong 1{,}33;$$

Turma E:

E	8	4	3	2	7	9	4	8	8	7	**SOMATÓRIO**
$E-\overline{E}$	2	–2	–3	–4	1	3	–2	2	2	1	**0**
$(E-\overline{E})^2$	4	4	9	16	1	9	4	4	4	1	**56**

$$S_E^2=\frac{\sum(E-\overline{E})^2}{n-1}\Rightarrow S_E^2=\frac{56}{10-1}\Rightarrow S_E^2=\frac{56}{9}\Rightarrow S_E^2\cong 6{,}22;$$

Para facilitar e não confundir, podemos agrupar as medidas obtidas num único quadro:

TURMA →	**A**	**B**	**C**	**D**	**E**
MÉDIA	9	8	7	7	6
VARIÂNCIA	0,86	1,00	5,20	1,33	6,22

Podemos agora, julgar as afirmativas:

(**F**) A turma **E** apresenta a menor média e a <u>segunda maior</u> variância.
(**V**) A turma **A** apresenta a maior média e a menor variância.
(**V**) Apenas uma das turmas avaliadas apresenta variância superior à média.
(**F**) As turmas **C** e **D** apresentam a mesma média, mas a turma **C** apresenta <u>menor</u> variabilidade.

Gabarito: LETRA **A**.

31 – Considere o seguinte diagrama de ramo-e-folhas:

2	5	8	8		
3	0	1	2	6	
4	5	5	5	7	
5	0	5	8	8	9
6	3	4	8		
7	3				

A média, a mediana e a moda desse conjunto de dados são, respectivamente:

A ( ) 47, 45, 45.
B ( ) 47, 47, 45.
C ( ) 47, 46, 45.
D ( ) 45, 45, 45.

<u>RESOLUÇÃO</u>:

Essa era a questão mais fácil e menos trabalhosa da prova. Quem estudou pela aula virtual “Formas de apresentação dos dados - Diagrama de Ramos e Folhas” (Toque 6 da Seção Toque de Mestre), não teve a menor dificuldade em resolvê-la.

Basta lembrar que, no diagrama, os números à esquerda são os ramos (2, 3, 4, 5, 6 e 7) e os números à direita são as folhas. Juntando em cada ramo as folhas correspondentes, podemos transformar o diagrama no seguinte rol:

25, 28, 28, 30, 31, 32, 36, 45, 45, 45, | 47, 50, 55, 58, 58, 59, 63, 64, 68, 73.

X_{10} | X_{11}

$$Md = \frac{X_{10} + X_{11}}{2} \Rightarrow Md = \frac{45 + 47}{2} \Rightarrow \mathbf{Md = 46}.$$

Vemos então que a distribuição tem 20 observações e, sendo esse número par, a mediana será a média aritmética dos dois termos centrais, no caso, o 10° e o 11°.

Não precisamos nem verificar qual é a Moda (que foi repetida em todas as opções de resposta) ou calcular a média, pois em apenas uma das opções de resposta a mediana é igual a 46 (letra C), mas podemos ver nesse rol que a observação com maior freqüência é igual a 45 (Moda) e que a soma das observações será igual a 940. Esse total dividido pelo número de observações nos fornecerá a média, igual a 47.

Caso o diagrama tivesse um número maior de observações, ao invés de escrever o rol (é prático porque são apenas 20 observações) seria melhor, aproveitando o próprio diagrama, criar uma coluna de freqüência e outra de freqüência acumulada crescente para localizar a mediana (vide o exemplo na aula virtual Toque 6.

Gabarito: LETRA **C**.

<u>COMENTÁRIOS SOBRE A PROVA</u>:

Uma questão muito fácil (31) e três questões muito trabalhosas (na questão 29 poderia já ter sido fornecido o rol das 60 observações), o que não mede conhecimento. Para uma banca desconhecida do público concurseiro, como é o caso do IADE, a prova de Estatística até surpreendeu, está num bom nível e sem falhas graves, mas muito distante do nível de excelência e criatividade nas questões de bancas como FCC, NCE e ESAF.

Parabéns aos aprovados e futuros colegas.

PEDRO BELLO