# PROBABILIDADE E ESTATÍSTICA

## Variância e Desvio-Padrão

Livro Eletrônico

GRAN CURSOS ONLINE

# SUMÁRIO

# VARIÂNCIA E DESVIO-PADRÃO

## 1. Medidas de Variabilidade

Já vimos que a medida de posição traz uma informação bastante útil, pois resume a amostra a uma única afirmação.

Porém, considere as três populações a seguir:

{3, 3, 3, 3, 3}

{1, 2, 3, 4, 5}

{-1, 1, 3, 5, 7}

É fácil perceber que todas essas populações têm média igual a 3. Porém, **elas são bastante diferentes entre si**.

A primeira delas é **homogênea**. Ou seja, todos os elementos são iguais e nenhum deles destoa da média.

A segunda amostra já perde um pouco essa característica. Há um desvio em relação à média de todos os elementos mostrados. No caso da terceira amostra, existem vários elementos que destoam bastante da média. Ela se inicia em -1 e termina em 7.

Por isso, em Estatística, tem-se a necessidade de criar um novo conceito que possa traduzir a questão da homogeneidade de uma amostra ou população. Esse conceito é chamado de **variância**.

### 1.1. Variância e Desvio-Padrão Populacional

O primeiro ponto que devemos observar no cálculo da variância é que **existem fórmulas diferentes quando tratamos de uma população e quando tratamos de uma amostra**.

Por isso, vamos nos recordar dos conceitos de população e amostra:

- **População:** corresponde a todo o conjunto de pessoas, itens ou eventos sobre os quais se deseja conhecer uma determinada propriedade ou fazer uma inferência;

**Exemplo:** suponha que você deseja conhecer a altura média dos brasileiros. Nesse caso, a população será todo o conjunto de brasileiros.

- **Amostra:** na maioria das situações, é inviável medir as propriedades de todo o conjunto em estudo. Por isso, é comum selecionar um subconjunto de elementos da população para realizar o estudo. Esse subconjunto é denominado amostra.

**Exemplo:** suponha que você é um funcionário do IBGE e fez uma pesquisa com várias pessoas e aferiu sua altura.

O conjunto das pessoas que foram selecionadas para medir a altura é a amostra.

É muito comum que **a própria questão te informe se o conjunto em estudo se trata de uma população ou de uma amostra**.

A variância busca medir de quanto os termos se distanciam de uma determinada medida de posição $\bar{x}$.

$$S^2 = \sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{N}$$

A variância é representada por S² mesmo. Não tente tirar a raiz, tá certo? ☺

Os termos $(x_i - \bar{x})$ são os desvios de cada elemento da amostra em relação à medida de posição $\bar{x}$.

Mas, por que elevamos ao quadrado?

Vamos fazer um exemplo e você vai ser capaz de entender.

Por exemplo, calculemos a variância padrão da população {1, 2, 3, 4, 5}. Devemos fazer a conta em três passos:

- Calculamos a média da população:

$$\mu = \frac{1+2+3+4+5}{5} = \frac{15}{5} = 3$$

- Calculamos os desvios de todas as observações em relação à média:

| X | X – μ | (X – μ)² |
|---|---|---|
| **1** | 1 – 3 = **-2** | |
| **2** | 2 – 3 = **-1** | |
| **3** | 3 – 3 = **0** | |
| **4** | 4 – 3 = **1** | |
| **5** | 5 – 3 = **2** | |

- Elevamos os desvios calculados ao quadrado:

| X | X – μ | (X – μ)² |
|---|---|---|
| 1 | 1 – 3 = **–2** | (**–2**)² = **4** |
| 2 | 2 – 3 = **–1** | (**–1**)² = **1** |
| 3 | 3 – 3 = **0** | (0)² = **0** |
| 4 | 4 – 3 = **1** | (**1**)² = **1** |
| 5 | 5 – 3 = **2** | (2)² = **4** |

Por fim, somamos todos os quadrados dos desvios e dividimos pelo número de elementos na amostra.

$$S^2 = \frac{4+1+0+1+4}{5} = \frac{10}{5} = 2$$

Agora, somos capazes de entender por que a variância leva em consideração os elementos da amostra ao quadrado. Observe que, quando elevamos ao quadrado, tanto os desvios abaixo da média como os desvios acima da média **são somados** e contribuem para elevar a variância.

O que aconteceria se, em vez disso, nós simplesmente somássemos todos os desvios em relação à média?

A resposta dessa pergunta já foi estudada no nosso curso. Nós vimos no capítulo sobre médias que a soma de todos os desvios em relação à média é sempre nula.

| X | X – μ |
|---|---|
| 1 | 1 – 3 = **–2** |
| 2 | 2 – 3 = **–1** |
| 3 | 3 – 3 = **0** |
| 4 | 4 – 3 = **1** |
| 5 | 5 – 3 = **2** |

$$S = -2 - 1 + 0 + 1 + 2 = 0$$

Dessa forma, simplesmente somar todos os desvios em relação à média seria uma informação irrelevante para a estatística. É sempre zero. É por isso que tomamos os quadrados dos desvios.

Outra informação interessante que a variância nos proporciona é sobre **a heterogeneidade** de uma população (ou amostra).

Quanto mais heterogênea for uma população, maior será a sua variância. Tomemos um exemplo {-1, 1, 3, 5, 7}. Para o cálculo da variância dessa população, temos que a média dos termos é:

$$\mu = \frac{-1+1+3+5+7}{5} = \frac{15}{5} = 3$$

Agora, tomemos os desvios em relação à média e os quadrados desses desvios.

| x | x – μ | (x – μ)² |
|---|---|---|
| **–1** | – 1 – 3 = **–4** | (**–4**)² = **16** |
| **1** | 1 – 3 = **–2** | (**–2**)² = **4** |
| **3** | 3 – 3 = **0** | (0)² = **0** |
| **5** | 5 – 3 = **2** | (2)² = **4** |
| **7** | 7 – 3 = **4** | (4)² = **16** |

Finalmente, tomamos a soma dos quadrados dos desvios e dividimos pelo total de elementos existentes na população.

$$S^2 = \frac{16+4+0+4+16}{5} = \frac{40}{5} = 8$$

Observe que a variância populacional aumentou de 2 para 8. Isso reflete que a segunda população é mais heterogênea do que a primeira.

Quanto maior a variância populacional, mais heterogênea é essa população.

Façamos a comparação entre as duas amostras estudadas nesse capítulo:

| Amostra | Variância | Característica |
|---|---|---|
| **{1, 2, 3, 4, 5}** | 2 | Mais Homogênea |
| **{–1, 1, 3, 5, 7}** | 8 | Mais Heterogênea |

Tabela 1: Comparação entre Amostras e suas Variâncias

### 1.1.1. Desvio-Padrão

O único problema da variância é que ela tem, em geral, unidade de grandeza diferente dos elementos da população ou amostra.

Por exemplo, quando calculamos a variância da população {-1, 1, 3, 5, 7}, encontramos que a média é igual a 3 e a variância é igual a 8.

O número 8 é bastante superior aos desvios encontrados nessa população. Por isso, fica bastante contraintuitivo.

É por isso que as pessoas normalmente utilizam o desvio-padrão, que corresponde à raiz quadrada da variância.

$$\sigma = \sqrt{S^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

O grande objetivo do desvio-padrão é que o desvio-padrão será mais compatível com os desvios encontrados na população ou amostra. No caso da amostra em apreço, tem-se .

Observe que 2,83 é um valor mais razoável em relação aos desvios encontrados. Analisaremos os desvios em relação à média que podem ser encontrados na população.

| Elemento | Desvio |
|---|---|
| **–1** | –1 – 3 = –4 |
| **1** | 1 –3 = –2 |
| **3** | 3 –3 = 0 |
| **5** | 5 –3 =+2 |
| **7** | 7 –3 =+4 |

Tabela 2: Desvios

Note que 2,83 é um número bem mais representativo desses desvios do que 8.

É por isso que o desvio-padrão é uma medida bem mais representativa da população do que a variância. Assim, ela será mais utilizada nos contextos práticos da Estatística.

## Direto do Concurso

**001.** (CESPE/2018/POLÍCIA FEDERAL/PERITO CRIMINAL FEDERAL/ÁREA 6) Considerando que a análise de uma amostra de minério de chumbo tenha apresentado os seguintes resultados percentuais (%): 8,10; 8,32; 8,12; 8,22; 7,99; 8,31, julgue o item a seguir, relativo a esses dados.

A variância dos dados em apreço é dada pelo valor do desvio-padrão ao quadrado.

É isso mesmo. A relação entre a variância e o desvio-padrão é que o desvio-padrão é igual à raiz quadrada da variância, e que a variância é igual ao quadrado do desvio-padrão.

**Certo.**

**002.** (CESPE/CEHAP-PB/2009/ADMINISTRADOR) O custo médio nacional para a construção de habitação com padrão de acabamento normal, segundo levantamento realizado em novembro de 2008, foi de R$ 670,00 por metro quadrado, sendo R$ 400,00/m² relativos às despesas com materiais de construção e R$ 270,00/m² com mão de obra. Nessa mesma pesquisa, os custos médios regionais apontaram para os seguintes valores por metro quadrado: R$ 700,00 (Sudeste), R$ 660,00 (Sul), R$ 670,00 (Norte), R$ 640,00 (Centro-Oeste) e R$ 630,00 (Nordeste).

*Sistema Nacional de Pesquisa de Custos e Índices da Construção Civil. SINAPI/IBGE, nov./2008 (com adaptações).*

O desvio-padrão dos custos médios regionais por metro quadrado foi:

a) inferior a R$ 30,00.
b) superior a R$ 30,01 e inferior a R$ 40,00.
c) superior a R$ 40,01 e inferior a R$ 50,00.
d) superior a R$ 50,01.

Nesse caso, temos uma população de regiões. De fato, no Brasil, só existem 5 regiões: Sudeste, Sul, Norte, Centro-Oeste e Nordeste. Portanto, realmente estamos falando de uma população, e não de uma amostra.

Dessa forma, para o cálculo da variância populacional, procedemos ao cálculo da média populacional, que é igual à soma de todas as observações dividida pelo total de elementos do conjunto.

$$\mu = \frac{700 + 660 + 670 + 640 + 630}{5} = \frac{3300}{5} = 660$$

Agora, podemos calcular os desvios em relação à média e os quadrados desses desvios.

| X | X – μ | (X – μ)² |
|---|---|---|
| **700** | 700 – 660 = 40 | (40)² = 1600 |
| **660** | 660 – 660 = 0 | (0)² = 0 |
| **670** | 670 – 660 = 10 | (10)² = 100 |
| **640** | 640 – 660 = –20 | (–20)² = 400 |
| **630** | 630 – 660 = –30 | (–30)² = 900 |

Por fim, basta tomar que a variância é igual à média dos quadrados dos desvios:

$$S^2 = \frac{1600 + 0 + 100 + 400 + 900}{5} = \frac{3000}{5} = 600$$

O desvio-padrão, por sua vez, corresponde à raiz quadrada da variância.

$$\sigma = \sqrt{600} \cong 24{,}5$$

Nem precisávamos, na verdade, ter calculado a raiz quadrada de 600. Basta olhar que, nas alternativas, a letra A dizia que o desvio-padrão era inferior a R$ 30,00. E, de fato, 30² = 900. Portanto, a raiz quadrada de 600 é realmente inferior a 30.

**Letra a.**

---

### 1.1.2. Equações para o Cálculo da Variância Populacional

Olhando a expressão da variância, o primeiro fato a reparar é que ela é uma média aritmética.

$$S^2 = \sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{N} = E[(X - \mu_X)^2]$$

A variância consiste **na média aritmética do quadrado dos desvios**. E é justamente por isso que podemos representar a variância também como **a esperança do quadrado dos desvios**, como mostrado na equação acima.

Já estudamos uma importante propriedade do operador esperança: **a esperança da soma é igual à soma das esperanças.**

Com base nisso, podemos chegar a uma expressão muito útil lembrando-nos que o valor esperado é um operador linear. Primeiramente, devemos aplicar o produto notável do quadrado da diferença.

$$S^2 = E[(X - \mu_X)^2] = E[X^2 - 2\mu_X X + \mu_X^2]$$

Agora, podemos usar que o valor esperado da soma é igual à soma dos valores esperados:

$$S^2 = E[X^2] - E[2\mu_X X] + \mu_X^2$$

Por fim, quando temos uma constante multiplicando uma variável, ela pode ser tirada para fora do valor esperado.

$$S^2 = E[X^2] - 2\mu_X E[X] + \mu_X^2$$

Agora, notemos que a esperança E[X] é igual à própria média $\mu_X$.

$$S^2 = E[X^2] - 2\mu_X^2 + \mu_X^2 = E[X^2] - \mu_X^2$$

Chegamos a uma importante relação:

| $S^2$ | = | $E[X^2]$ | – | $\mu_X^2$ |
|---|---|---|---|---|
| Variância | = | **Média dos Quadrados** | – | **Quadrado da Média** |

Essa é uma expressão bastante interessante. A variância populacional é igual à média dos quadrados menos o quadrado da média.

Essa expressão é bastante útil. Algumas questões aparecem com somatórios imensos que assustam à primeira vista. Porém, se a banca fornecer o somatório dos quadrados, você poderá mais facilmente calcular a variância.

## Direto do Concurso

**003.** (FCC/TRT-9ª REGIÃO/PR/2020/ANALISTA JUDICIÁRIO/ESTATÍSTICA) Uma população com 16 valores estritamente positivos X1, X2, X3, ..., X16, correspondente a um determinado atributo, apresenta as seguintes informações:

$$\sum_{i=1}^{16} X_i = 192 \text{ e } \sum_{i=1}^{16} X_i^2 = 4.464$$

O elemento $X_{10}$, tal que $X_{10}$ = 12, é retirado da população. Os valores da variância da primeira população e da nova população formada são, respectivamente, iguais a:

a) 144 e 134,40.

b) 144 e 144.

c) 135 e 144.

d) 135 e 135.

e) 135 e 126.

Nesse tipo de questão, muitos alunos costumam se assustar quando vêem os somatórios. Porém, veja que, na verdade, eles facilitam muito a resolução de questões.
Basta que você se lembre-se de que a variância é igual à média dos quadrados menos o quadrado das médias.

$$S^2 = E[X^2] - \mu^2$$

A média da população pode ser calculada como a soma de todos os elementos dividida pelo número de elementos da população.

$$\mu = \frac{\sum X_i}{N} = \frac{192}{16} = 12$$

Da mesma forma, a esperança de X² é igual ao somatório de X² dividido pelo total do número de elementos da amostra.

$$E[X^2] = \frac{\sum X_i^2}{N} = \frac{4464}{16} = 279$$

Agora, vamos aplicar os valores calculados na expressão da variância.

$$S^2 = E[X^2] - \mu^2$$

$$S^2 = 279 - 12^2 = 279 - 144 = 135$$

Portanto, a variância da população original é igual a 135.
Na nova população, com a exclusão de um dos elementos, o elemento $X_{12} = 10$ deve ser abatido tanto da soma dos termos como da soma dos quadrados.

$$\sum X_i = 192 - 12 = 180$$

$$\sum X_i^2 = 4464 - 12^2 = 4464 - 144 = 4320$$

Por fim, podemos calcular a média populacional e a esperança de X², observando que o número de elementos da população diminui para 15, pois um elemento foi retirado.

$$\mu = \frac{\sum X_i}{N} = \frac{180}{15} = 12$$

$$E[X^2] = \frac{\sum X_i^2}{N} = \frac{4320}{15} = 288$$

Finalmente, devemos utilizar novamente a expressão da variância, que é igual à média dos quadrados menos o quadrado da média.

$$S^2 = E[X^2] - \mu^2 = 288 - 12^2 = 288 - 144 = 144$$

**Letra c.**

---

### 1.1.3. Por que Usar a Média?

Na realidade, qualquer medida de posição pode ser utilizada como referência $\bar{x}$ para o cálculo da variância.

Porém, é possível demonstrar, com o uso de Cálculo Diferencial, que a média aritmética é a medida de posição que minimiza o valor da variância.

Por isso, a média aritmética é sempre utilizada na variância.

Registre essa informação, pois ela pode ser cobrada em questões teóricas.

---

A média aritmética é o valor de $\bar{x}$ que minimiza a função de custo correspondente à variância:

$$S^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 \rightarrow S_{MIN}^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

---

Vamos entender um pouco melhor isso com um exemplo. Vejamos a seguinte distribuição:

{1, 3, 3, 3, 4, 5, 6, 7}

Podemos calcular para essa distribuição diversas medidas de posição, como a média, a moda e a mediana.

**Moda:** é a observação mais frequente, no caso, Mo = 3.

**Mediana:** como a amostra tem 8 termos, a mediana será:

$$Md = \frac{x_4 + x_5}{2} = \frac{3+4}{2} = \frac{7}{2} = 3{,}5$$

**Média:** pela definição, é a soma de todos as observações dividida pelo número de elementos no conjunto.

$$Me = \frac{1+3+3+3+4+5+6+7}{8} = \frac{32}{8} = 4$$

Façamos o cálculo da soma dos quadrados dos desvios em relação à moda.

| X | (X – Mo) | (X – Mo)² |
|---|---|---|
| **1** | (1 – 3) = –2 | (–2)² = 4 |
| **3** | (3 – 3) = 0 | 0² = 0 |
| **3** | (3 – 3) = 0 | 0² = 0 |
| **3** | (3 – 3) = 0 | 0² = 0 |
| **4** | (4 – 3) = 1 | 1² = 1 |
| **5** | (5 – 3) = 2 | 2² = 4 |
| **6** | (6 – 3) = 3 | 3² = 9 |
| **7** | (7 – 3) = 4 | 4² = 16 |

Dessa forma, a média dos quadrados dos desvios em relação à moda seria:

$$S^2 = \frac{4+0+0+0+1+4+9+16}{8} = \frac{34}{8} = 4{,}25$$

| X | (X – Md) | (X – Md)² |
|---|---|---|
| **1** | (1 – 3,5) = –2,5 | (–2,5)² = 6,25 |
| **3** | (3 – 3,5) = –0,5 | (–0,5)² = 0,25 |
| **3** | (3 – 3,5) = –0,5 | (–0,5)² = 0,25 |
| **3** | (3 – 3,5) = –0,5 | (–0,5)² = 0,25 |
| **4** | (4 – 3,5) = 0,5 | (0,5)² = 0,25 |
| **5** | (5 – 3,5) = 1,5 | (1,5)² = 2,25 |
| **6** | (6 – 3,5) = 2,5 | (2,5)² = 6,25 |

| X | (X – Md) | (X – Md)² |
|---|---|---|
| 7 | (7 – 3,5) = 3,5 | (3,5)² = 12,25 |

Façamos, agora, o cálculo da média dos quadrados dos desvios em relação à mediana.

$$S^2 = \frac{6{,}25 + 0{,}25 + 0{,}25 + 0{,}25 + 0{,}25 + 2{,}25 + 6{,}25 + 12{,}25}{8} = \frac{28}{8} = 3{,}5$$

| X | (X – Me) | (X – Me)² |
|---|---|---|
| 1 | (1 – 4) = –3 | (–3)² = 9 |
| 3 | (3 – 4) = –1 | (–1)² = 1 |
| 3 | (3 – 4) = –1 | (–1)² = 1 |
| 3 | (3 – 4) = –1 | (–1)² = 1 |
| 4 | (4 – 4) = 0 | (0)² = 0 |
| 5 | (5 – 4) = 1 | (1)² = 1 |
| 6 | (6 – 4) = 2 | (2)² = 4 |
| 7 | (7 – 4) = 3 | (3)² = 9 |

Façamos, agora, o cálculo da média dos quadrados dos desvios em relação à mediana.

$$S^2 = \frac{9 + 1 + 1 + 1 + 0 + 1 + 4 + 9}{8} = \frac{26}{8} = 3{,}25$$

Observe que o menor valor possível foi calculado justamente tomando a média como referência. Portanto, ao utilizar a média para o cálculo da variância, sempre chegamos ao menor valor possível.

## 1.2. Variância e Desvio-Padrão Amostral

Na Estatística, é bastante comum trabalharmos com amostras. E isso requer **atenção**, tendo em vista que existem duas formas de calcular a variância e o desvio-padrão amostral.

**Estimador de Máxima Verossimilhança:** é uma técnica de cálculo que considera que a amostra obtida é a amostra mais provável de ser obtida a partir da população original.

Por esse motivo, o estimador de máxima verossimilhança considera que a amostra obtida é um espelho da população. E, assim, o cálculo da variância e o do desvio-padrão amostral deve ser feito exatamente como aprendemos na seção anterior. Ou seja, devemos colocar N – o número de elementos da amostra no denominador.

$$S^2 = \sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{N} = E\left[(X - \mu_X)^2\right]$$

Fique atento, pois há questões de prova que falam em estimador de máxima verossimilhança.

É possível provar, no entanto, que o estimador de máxima verossimilhança é viesado ou tendencioso, ou seja, ele tende a retornar um valor $(\hat{\sigma})$ diferente do valor real do parâmetro populacional $(\sigma)$. O erro médio produzido pelo estimador de máxima verossimilhança é dado por:

$$Erro = \hat{\sigma} - \sigma = -\frac{\hat{\sigma}}{N}$$

A demonstração dessa propriedade foge ao escopo desse curso e só pode ser feita com ferramentas mais avançadas da Estatística. Mas, fique tranquilo.

Tudo o que você precisa saber sobre o estimador de máxima verossimilhança é: ele calcula a variância amostral como se fosse a variância populacional e que ele é um estimador tendencioso, ou seja, não retorna o valor real do parâmetro.

**Estimador não Tendencioso:** o estimador não tendencioso é aquele que tende realmente ao valor correto da variância populacional.

Nesse estimador, o cálculo do desvio-padrão é feito de maneira ligeiramente diferente. O denominador deve ser dividido por N-1, em vez de dividir por N.

$$\sigma = \sqrt{S^2} = \sqrt{\frac{1}{(N-1)} \sum_{i=1}^{N} (x_i - \bar{x})^2}$$

Porém, podemos aprender logo agora a diferença entre o cálculo do desvio-padrão para uma população inteira e para uma amostra.

O desvio-padrão amostral pode ser obtido a partir do desvio-padrão populacional a partir de um fator de correção, como mostrado a seguir.

$$S_{AM}^2 = \frac{N}{N-1} . S_{POP}^2$$

$$\sigma_{AM}^2 = \sqrt{\frac{N}{N-1}} . \sigma_{POP}^2$$

O fator de correção N/(N-1) é aplicado para a variância, enquanto a raiz quadrada desse fator é aplicada para o desvio-padrão.

Quando o número de elementos da amostra é muito grande, podemos considerar que a razão N/(N – 1) é muito próxima de 1. Nessa situação, o fator de ajuste é desprezível. Vejamos alguns cálculos para você entender:

| N | N/(N – 1) |
|---|---|
| **5** | 1,25 |
| **20** | 1,052632 |
| **50** | 1,020408 |
| **100** | 1,010101 |
| **1000** | 1,001001 |

Note que, quanto maior o tamanho da amostra, mais próximo de 1 será o fator N/(N – 1). Por isso, é possível que a questão te informe que esse fator de correção é desprezível.

Dessa forma, você deve prestar atenção às informações do enunciado:

**Use N no denominador:** quando o enunciado falar em uma população ou em estimador de máxima verossimilhança;

**Use N – 1 no denominador:** quando o enunciado em uma amostra e pedir o estimador não tendencioso.

Se, por acaso, o enunciado pedir a variância ou o desvio-padrão amostral e não falar nada sobre o tipo de estimador que deseja, suponha sempre o estimador não tendencioso, ou seja, coloque N – 1 no denominador.

## Direto do Concurso

**004.** (FUNRIO/SESAU-RO/2017) Para responder às próximas três questões, considere a seguinte amostra de idades: 18, 15, 24, 20, 22, 21, 19, 30, 20
A estimativa não tendenciosa usual da variância populacional é aproximadamente igual a:
**a)** 15,8.
**b)** 17,8.
**c)** 20,1.
**d)** 22,2.
**e)** 23,1.

Podemos calcular a variância pela definição. Primeiramente, vamos calcular a média aritmética da amostra.

$$\mu = \frac{18+15+24+20+22+21+19+30+20}{9} = \frac{189}{9} = 21$$

Agora, podemos calcular os desvios e os quadrados dos desvios.

| X | X – μ | (X – μ)² |
|---|---|---|
| **18** | 18 – 21 = –3 | (–3)² = 9 |
| **15** | 15 – 21 = –6 | (–6)² = 36 |
| **24** | 24 – 21 = 3 | (3)² = 9 |
| **20** | 20 – 21 = –1 | (–1)² = 1 |
| **22** | 22 – 21 = 1 | (1)² = 1 |
| **21** | 21 – 21 = 0 | (0)² = 0 |
| **19** | 21 – 19 = 2 | (2)² = 4 |
| **30** | 30 – 21 = 9 | (9)² = 81 |
| **20** | 20 – 21 = –1 | (–1)² = 1 |
| **Total** | | |

Como se tem uma amostra e desejamos a estimativa não viesada para o desvio-padrão populacional, é necessário adotar o fator de ajuste no denominador. Dessa forma, a variância pode ser calculada como a soma dos quadrados dos desvios dividida por N – 1.

$$S^2 = \frac{9+36+9+1+1+0+4+81+1}{9-1} = \frac{142}{8} = 17{,}75$$

**Letra b.**

---

**005.** (CESPE/TCE-PA/2016/AUDITOR DE CONTROLE EXTERNO) Caso, em uma amostra aleatória de tamanho *n* = 4, os valores amostrados sejam A = {2, 3, 0, 1}, a estimativa de máxima verossimilhança para a variância populacional será igual a 5/3.

---

Na estimativa de máxima verossimilhança, não é feita nenhuma correção de ajuste.

$$\widehat{\sigma^2} = \frac{(2-1{,}5)^2 + (3-1{,}5)^2 + (0-1{,}5)^2 + (1-1{,}5)^2}{4}$$
$$= \frac{(0{,}5)^2 + (1{,}5)^2 + (-1{,}5)^2 + (-0{,}5)^2}{4}$$
$$= \frac{0{,}25 + 2{,}25 + 2{,}25 + 0{,}25}{4} = \frac{5}{4}$$

**Errado.**

---

### 1.2.1. Dados em Tabelas

Um caso particular interessante que podemos ter é quando os dados são fornecidos na forma de tabelas.

Como a variância é simplesmente uma média aritmética, podemos resolver esse problema utilizando uma **média ponderada.** Isso significa que nós devemos calcular os desvios referentes a cada observação e tomar como pesos relativos as frequências absolutas ou relativas de cada observação tabelada.

Vejamos um exemplo para que você entenda melhor. Considere uma pesquisa feita por uma empresa sobre o número de filhos de **uma amostra** de seus funcionários.

| Número de Filhos | Quantidade de Casais (Frequência Absoluta) |
|---|---|
| **0** | 17 |
| **1** | 33 |
| **2** | 21 |
| **3** | 7 |
| **4** | 2 |
| **5** | 0 |

*Tabela 3: Número de Filhos de Casais em uma Empresa*

Devemos esclarecer que o passo a passo é exatamente o mesmo que nós vimos para quando os dados são fornecidos na forma de um rol:

Primeiramente, vamos calcular a média da distribuição, utilizando como pesos a frequência absoluta de cada categoria:

$$\mu = \frac{17.0 + 33.1 + 21.2 + 7.3 + 2.4 + 0.5}{17 + 33 + 21 + 7 + 2 + 0}$$

$$= \frac{0 + 33 + 42 + 21 + 8 + 0}{80}$$

$$\mu = \frac{104}{80} = 1{,}3$$

O segundo passo é calcular os desvios referentes a cada uma das linhas da Tabela 3:

| Número de Filhos | Quantidade de Casais (Frequência Absoluta) | Desvios |
|---|---|---|
| **0** | 17 | 0 – 1,3 = – 1,3 |
| **1** | 33 | 1 – 1,3 = – 0,3 |
| **2** | 21 | 2 – 1,3 = 0,7 |
| **3** | 7 | 3 – 1,3 = 1,7 |
| **4** | 2 | 4 – 1,3 = 2,7 |
| **5** | 0 | 5 – 1,3 = 3,7 |

Tabela 4: Cálculo dos Desvios do Número de Filhos de Casais em uma Empresa

O terceiro passo é calcular os quadrados dos desvios:

| Número de Filhos | Quantidade de Casais | Desvios | Quadrado dos Desvios |
|---|---|---|---|
| **0** | 17 | 0 – 1,3 = – 1,3 | $(1{,}3)^2 = 1{,}69$ |
| **1** | 33 | 1 – 1,3 = – 0,3 | $(0{,}3)^2 = 0{,}09$ |
| **2** | 21 | 2 – 1,3 = 0,7 | $(0{,}7)^2 = 0{,}49$ |

| Número de Filhos | Quantidade de Casais | Desvios | Quadrado dos Desvios |
|---|---|---|---|
| **3** | 7 | 3 – 1,3 = 1,7 | $(1,7)^2 = 2,89$ |
| **4** | 2 | 4 – 1,3 = 2,7 | $(2,7)^2 = 7,29$ |
| **5** | 0 | 5 – 1,3 = 3,7 | $(3,7)^2 = 13,69$ |

Tabela 5: Cálculo dos Quadrados dos Desvios

Por fim, devemos calcular a variância como a média dos quadrados dos desvios ponderada pelas frequências absolutas:

$$S^2 = \frac{17.1,69 + 33.0,09 + 21.0,49 + 7.2,89 + 2.7,29 + 0.13,69}{17 + 33 + 21 + 7 + 2 + 0 - 1}$$

$$S^2 = \frac{28,73 + 2,97 + 10,29 + 20,23 + 14,58 + 0}{80 - 1} = \frac{76,8}{7976} \cong 0,972$$

Portanto, a variância da população tratada na Tabela 3 é igual a 0,972. Observe que incluímos o −1 no denominador, porque os dados da Tabela 3 se referem a uma amostra, portanto, devemos incluir o fator de ajuste.

Por sua vez, o desvio-padrão pode ser calculado como a raiz quadrada da variância.

$$\sigma = \sqrt{0,972} \cong 0,986$$

Outro caso de particular importância é quando os dados são fornecidos na forma de **categorias.** Considere a seguinte amostra:

| Faixa Salarial | Número de Empregados |
|---|---|
| **1 \|– 3** | 50 |
| **3 \|– 5** | 32 |
| **5 \|– 7** | 12 |
| **7 \|– 9** | 6 |

Tabela 6: Faixas Salariais dos Empregados

Como a variância é essencialmente média aritmética, podemos recorrer à **técnica do ponto médio**. Seguiremos o mesmo passo a passo do cálculo da variância a que já estamos habituados:

O primeiro passo é calcular a média da distribuição, empregando a técnica do ponto médio;

| Faixa Salarial | Ponto Médio | Número de Empregados |
|---|---|---|
| **1 \|– 3** | 2 | 50 |
| **3 \|– 5** | 4 | 32 |
| **5 \|– 7** | 6 | 12 |
| **7 \|– 9** | 8 | 6 |

$$\mu = \frac{50.2 + 32.4 + 12.6 + 6.8}{50 + 32 + 12 + 6} = \frac{100 + 128 + 72 + 48}{80} = \frac{348}{80} = 4{,}35$$

O segundo passo é calcular os desvios do ponto médio de cada categoria em relação à média da distribuição;

| Ponto Médio | Desvios | Número de Empregados |
|---|---|---|
| **2** | 2 – 4,35 = –2,35 | 50 |
| **4** | 4 – 4,35 = –0,35 | 32 |
| **6** | 6 – 4,35 = 1,65 | 12 |
| **8** | 8 – 4,35 = 3,65 | 6 |

O terceiro passo é calcular os quadrados dos desvios:

| Desvios | Quadrados dos Desvios | Número de Empregados |
|---|---|---|
| **2 – 4,35 = –2,35** | $(2{,}35)^2 = 5{,}5225$ | 50 |
| **4 – 4,35 = –0,35** | $(0{,}35)^2 = 0{,}1225$ | 32 |
| **6 – 4,35 = 1,65** | $(1{,}65)^2 = 2{,}7225$ | 12 |
| **8 – 4,35 = 3,65** | $(3{,}65)^2 = 13{,}3225$ | 6 |

Por fim, tomamos a variância como a média dos quadrados dos desvios ponderadas pelas frequências absolutas de cada classe.

$$S^2 = \frac{50.5{,}5225 + 32.0{,}1225 + 12.2{,}7225 + 6.13{,}3225}{50 + 32 + 16 + 6 - 1}$$

$$S^2 = \frac{276{,}125 + 3{,}92 + 32{,}67 + 79{,}935}{80 - 1} = \frac{392{,}65}{79} \cong 4{,}97$$

Como os dados são de amostras, precisamos utilizar o ajuste N – 1 no denominador.

Por sua vez, o desvio-padrão pode ser calculado como a raiz quadrada da variância.

$$\sigma = \sqrt{4{,}97} \cong 2{,}23$$

## 1.3. Propriedades da Variância e do Desvio-Padrão

Por enquanto, aprenderemos duas propriedades importantes da variância e do desvio-padrão.

**Variância da Soma por uma Constante:** a variância não é afetada pela soma ou subtração de um valor constante.

Cuidado! Não estamos somando variáveis aleatórias.

$$Var(X + a) = Var(X)$$

A explicação para isso é que, quando somamos um parâmetro **k** a todas as observações da amostra, a média é alterada pelo mesmo fator constante **k.** E, por isso, os desvios em relação à média não são alterados.

Suponhamos que a empresa da Tabela 3 tenha doado dois filhos a mais para cada um de seus funcionários. Dessa forma, todos os funcionários que tinham 0 filhos passaram a ter 2 filhos; todos os funcionários que tinham 1 filho passaram a ter 3 filhos; e, assim, por diante.

| Frequência Absoluta | Antes da Doação | Depois da Doação |
|---|---|---|
| **17** | 0 | 2 |
| **33** | 1 | 3 |
| **21** | 2 | 4 |
| **7** | 3 | 5 |
| **2** | 4 | 6 |

*Tabela 8: Número de Filhos de Casais em uma Empresa*

Como vimos nas propriedades da média, a nova média foi também aumentada em duas unidades em relação ao valor calculado originalmente.

$$\mu = 1{,}3 + 2 = 3{,}3$$

Quando calcularmos os desvios em relação à média da variável alterada, chegaremos exatamente aos mesmos valores calculados antes da alteração.

| Frequência Absoluta | Antes da Doação | Desvios antes da Doação | Depois da Doação | Desvios depois da Doação |
|---|---|---|---|---|
| **17** | 0 | 0 – 1,3 = – 1,3 | 2 | 2 – 3,3 = – 1,3 |
| **33** | 1 | 1 – 1,3 = – 0,3 | 3 | 3 – 3,3 = – 0,3 |
| **21** | 2 | 2 – 1,3 = 0,7 | 4 | 4 – 3,3 = 0,7 |
| **7** | 3 | 3 – 1,3 = 1,7 | 5 | 5 – 3,3 = 1,7 |
| **2** | 4 | 4 – 1,3 = 2,7 | 6 | 6 – 3,3 = 2,7 |

*Tabela 9: Cálculo dos Desvios do Número de Filhos de Casais em uma Empresa*

Em questões de provas, essa propriedade pode ser cobrada da seguinte maneira:

$$Var(X + 1) = Var(X)$$

$$Var(X - 3) = Var(X)$$

**Variância do Produto por uma Constante:** a constante sai da variância elevada ao quadrado.

$$V(ar(aX) = a^2 VarX)$$

No caso de desvio-padrão, a constante fica multiplicando o desvio-padrão.

$$\sigma((aX) = a\sigma X)$$

Podemos deduzir essa propriedade olhando para a expressão que define a variância.

Sejam $x_i$ os valores das observações da variável X e μ a média populacional, o desvio-padrão da população inicial era dado por:

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

Após o aumento de todos os elementos e da média, esses passaram a ser x' e μ', teremos que o novo desvio-padrão será:

$$\sigma'^2 = \frac{\sum (x' - \mu')^2}{N} = \frac{\sum (ax - a\mu)^2}{N} = \frac{(a)^2 \sum (x - \mu)^2}{N} = (a)^2 \sigma^2$$

$$\therefore \sigma' = a.\sigma$$

Sendo assim, bastante atenção. O desvio-padrão fica multiplicado pela constante **a**, porém, no caso de variância, essa constante sai elevada ao quadrado.

Podemos utilizar conjuntamente as duas propriedades. Suponha que X seja uma variável aleatória, de modo que Var (X) = 4. Quanto valeriam?

Var (2X + 3) =?

σ (5X − 4) =?

Primeiramente, devemos observar que o desvio-padrão é igual à raiz quadrada da variância.

$$\sigma = \sqrt{4} = 2$$

O termo constante que é adicionado não influencia a variância. Portanto, podemos simplesmente ignorá-lo.

$$Var\ (2X + 3) = Var\ (2X)$$

O termo multiplicativo pode ser retirado da variância, mas note que ele sai ao quadrado.

$$Var\ (2X + 3) = Var\ (2X) = 2^2.Var(X) = 4.4 = 16$$

O termo constante subtrativo também não modifica o cálculo do desvio-padrão.

$$\sigma\ (5X - 4) = \sigma\ (5X)$$

No caso do desvio-padrão, o termo constante multiplicativo pode ser retirado, multiplicando o desvio-padrão.

$$\sigma\ (5X - 4) = \sigma\ (5X) = 5.\sigma(X) = 5.2 = 10$$

Memorize o seguinte:

| Desvio-padrão | Variância |
|---|---|
| • σ(aX) = a.σ(X) | • Var (aX) = a²Var (X) |

Figura 1: Desvio-padrão e Variância

## 1.4. Coeficiente de Variação de Pearson (CV)

O coeficiente de variação de Pearson tem o objetivo de comparar o desvio-padrão com a média.

Pode ser expresso em termos percentuais, porém, nada impede que ele seja maior que 1 ou 100%.

$$CV = \frac{\sigma}{\bar{x}}$$

De forma análoga, existe também o conceito de **variância relativa**, que é igual ao quadrado do coeficiente de variação. Ou seja:

$$VR = \frac{Variância}{Média^2} = \frac{\sigma^2}{\bar{x}^2} = \frac{Var}{\bar{x}^2}$$

Observe que a variância relativa é igual ao quadrado do coeficiente de variação.

O objetivo por trás do coeficiente de variação e da variância relativa é criar uma comparação mais justa das variações em relação à média.

Tomemos como exemplo as duas distribuições estudadas anteriormente:

| Número de Filhos | Quantidade de Casais | Faixa Salarial | Número de Empregados |
|---|---|---|---|
| **0** | 17 | | |
| **1** | 33 | **1 \|– 3** | 50 |
| **2** | 21 | **3 \|– 5** | 32 |
| **3** | 7 | **5 \|– 7** | 12 |
| **4** | 2 | **7 \|– 9** | 6 |
| **5** | 0 | | |
| **Desvio-padrão** | **0,985** | **Desvio-padrão** | **2,23** |
| **Média** | **1,35** | **Média** | **4,35** |

Tabela 7: Comparação entre Desvio-padrão e Média

Se olharmos exclusivamente o desvio-padrão, chegaríamos à conclusão de que a faixa salarial se dispersa mais do que o número de filhos por casal. Porém, quando analisamos o desvio-padrão relativo ou o coeficiente de variação de Pearson, chegamos a uma conclusão interessante:

$$\boldsymbol{CV}_1 = \frac{\boldsymbol{\sigma}_1}{\boldsymbol{\mu}_1} = \frac{0{,}985}{1{,}35} \cong 0{,}73 = 73\% \qquad \boldsymbol{CV}_2 = \frac{\boldsymbol{\sigma}_2}{\boldsymbol{\mu}_2} = \frac{2{,}23}{4{,}35} \cong 0{,}52 = 52\%$$

Observe que o desvio-padrão do número de filhos corresponde a 73% da média e que o desvio-padrão da faixa salarial corresponde a 52% da média.

Uma propriedade interessante é que **o coeficiente de variação não é alterado pela multiplicação por uma constante.** Pense, por exemplo, que tenhamos uma variável X, cuja média seja igual a μ e cujo desvio-padrão seja igual a σ.

Se, por exemplo, essa variável for multiplicada por 1,2, tanto a média como o desvio-padrão serão multiplicados pelo mesmo fator constante. Dessa forma, o novo coeficiente de variação será:

$$CV' = \frac{1{,}2.\sigma}{1{,}2.\mu} = \frac{\sigma}{\mu} = CV$$

## Direto do Concurso

**006.** (FCC/BANRISUL/2019/ESCRITURÁRIO) 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.

---

Primeiramente, vamos calcular a média amostral.

$$\bar{x} = \frac{4+5+5+8}{4} = \frac{22}{4} = 5{,}5$$

Agora, vamos utilizar a expressão do desvio-padrão populacional. Como se trata de uma população, devemos colocar no denominador o tamanho total da população N = 4.

$$\sigma = \sqrt{\frac{(4-5{,}5)^2 + (5-5{,}5)^2 + (5-5{,}5)^2 + (8-5{,}5)^2}{4}}$$

$$\sigma = \sqrt{\frac{(-1{,}5)^2 + (-0{,}5)^2 + (-0{,}5)^2 + (2{,}5)^2}{4}}$$

$$\sigma = \sqrt{\frac{(2{,}25) + (0{,}25) + (0{,}25) + 6{,}25}{4}} = \sqrt{\frac{9}{4}} = \frac{3}{2}$$

O coeficiente de variação é dado pela razão entre o desvio-padrão e a média.

$$CV = \frac{\sigma}{\bar{x}} = \frac{3/2}{5{,}5} = \frac{3/2}{11/2} = \frac{3}{2} \cdot \frac{2}{11} = \frac{3}{11}$$

**Letra a.**

---

## 1.5. Normalização

Uma variável normalizada é aquela que possui média nula e desvio-padrão igual a um.

É possível normalizar qualquer variável aleatória aplicando a seguinte transformação:

$$Z = \frac{X - \mu}{\sigma}$$

Na expressão acima, μ e σ são, respectivamente, a média e o desvio-padrão da variável aleatória X.

A variável normalizada Z indica quantos desvios padrões abaixo ou acima da média está cada observação da variável X.

A normalização é bastante utilizada por bancas organizadoras de concursos públicos e vestibulares para calcular suas notas de uma maneira mais equilibrada.

Por exemplo, quando eu fiz o vestibular de Medicina, a UPE calculava a nota dos candidatos pelo escore padrão:

$$SP = 500 + 100.Z$$

Nessa expressão, Z é a nota normalizada do candidato.

O objetivo desse procedimento é evitar que uma prova que tenha tido um elevado desvio-padrão contribua significativamente para a colocação de um candidato.

Por exemplo, se a prova de Biologia teve uma média 8 e desvio-padrão 1, temos que uma faixa dentro de um desvio-padrão das notas é de [7,9].

Por outro lado, se a prova de Português teve média menor 7, porém desvio-padrão de 2, a faixa de resultados dentro de um desvio-padrão é [5,9].

Note que um candidato que tire uma nota 5 em Biologia tem uma nota normalizada:

$$Z = \frac{5 - 8}{1} = -3$$

Isso significa que esse candidato está três desvios padrões abaixo da média. Portanto, ele teve um desempenho bastante ruim.

Por outro lado, se o candidato tirou nota 5 em Português, sua nota normalizada seria:

$$Z = \frac{5-7}{2} = -1$$

Dessa maneira, o candidato está apenas um desvio-padrão abaixo da média. Isso mostra que uma nota 5 em Português é bem menos grave (ou mais comum) que uma nota 5 em Biologia.

Sendo assim, uma nota que destoa bastante da média em Português seria punida menos severamente do que a mesma nota na prova de Biologia.

Quando a banca adota o escore padrão, é bastante importante para o candidato que ele tenha muito cuidado nas matérias que tenham desvio-padrão mais baixo. Ou seja, nas matérias em que o desempenho dos candidatos tende a ser mais homogêneo.

Nessas matérias, cada redução ou aumento da sua nota entrará para o cômputo final da sua nota com um peso maior. Portanto, cada questão será mais valiosa individualmente do que as questões de matérias em que os candidatos possuem desempenho mais heterogêneo.

## Direto do Concurso

**007.** (ESAF/AFRFB/2012) Em um concurso público a nota média da prova de inglês foi igual a 7 com desvio-padrão igual a 2. Por outro lado, a nota média da prova de lógica foi igual a 7,5 com desvio-padrão igual a 4. Naná obteve nota 8 em Inglês e nota 8 em Lógica. Nené obteve nota 7,5 em Inglês e 8,5 em Lógica. Nini obteve 7,5 em Inglês e 9 em Lógica. Com relação à melhor posição relativa ─ ou ao melhor desempenho ─, pode-se afirmar que o desempenho de:

a) Naná foi o mesmo em Inglês e Lógica.
b) Nini foi melhor em Lógica do que o de Naná em Inglês.
c) Nené foi melhor em lógica do que o de Naná em Inglês.
d) Nené foi o mesmo em Inglês e Lógica.
e) Nené foi melhor em Lógica do que em Inglês.

---

A Receita Federal poderia ter sido mais criativa e ter inventado nomes melhores para Naná, Nené e Nini.

De qualquer forma, quando a questão fala em posição relativa, ela nos pede para comparar as notas normalizadas.

| | Inglês | Lógica |
|---|---|---|
| **Naná** | $\frac{8-7}{2} = \frac{1}{2}$ | $\frac{8-7{,}5}{4} = \frac{0{,}5}{4} = \frac{1}{8}$ |

| | Inglês | Lógica |
|---|---|---|
| **Nené** | $\frac{7,5-7}{2}=\frac{0,5}{2}=\frac{1}{4}$ | $\frac{8,5-7,5}{4}=\frac{1}{4}$ |
| **Nini** | $\frac{7,5-7}{2}=\frac{0,5}{2}=\frac{1}{4}$ | $\frac{9-7,5}{4}=\frac{1,5}{4}=\frac{3}{8}$ |

Sendo assim, o desempenho relativo de Nené foi o mesmo em Inglês e Lógica.

**Letra d.**

# RESUMO

**Procedimento de Cálculo da Variância:**

- Calcule a média aritmética dos termos da amostra ou população;
- Calcule os desvios em relação à média;
- Eleve os desvios calculados ao quadrado;
- Some todos os desvios e divida por:
- **N,** se for uma população ou o estimador de máxima verossimilhança;
- **N – 1**, se for uma amostra e for pedida a estimativa não tendenciosa ou não viciada.

**Propriedades do Desvio-padrão e Variância**:

- É a raiz quadrada da variância;
- Não são afetados por uma constante de soma ou subtração:

$$\sigma(X + c) = \sigma(X)$$
$$Var(X + c) = Var(X)$$

Quando a variável é multiplicada por uma constante, o mesmo acontece com o desvio-padrão:

$$\sigma(cX) = c.\sigma(X)$$

No caso da variância, a constante sai elevada ao quadrado:

$$Var(cX) = c^2.Var(X)$$

**Coeficiente de Variação:**

- É dado pela razão entre o desvio-padrão e a média;

$$CV = \frac{\sigma}{\mu}$$

- Não é afetado por uma constante multiplicativa;

$$CV(aX) = CV(X)$$

**Normalização:** consiste em subtrair a média e dividir pelo desvio-padrão. Dessa forma, chega-se a uma variável aleatória com média nula e desvio-padrão unitário

$$Z = \frac{X - \mu}{\sigma}$$

## MAPAS MENTAIS

# Medidas de Variabilidade

## Normalização

$$Z = \frac{X - \mu}{\sigma}$$

## Procedimento de Cálculo da Variância

**N** → Se for uma população ou o estimador de máxima verossimilhança.

**N – 1** → Se for uma amostra e for pedida a estimativa não tendenciosa ou não viciada.

## Coeficiente de Variação

$$CV = \frac{\sigma}{\mu}$$

Razão entre o desvio padrão e a média.

$$CV(aX) = CV(X)$$

Não é afetado por uma constante multiplicativa.

## Propriedades do Desvio Padrão e Variância

$$\sigma(cX) = c.\sigma(X)$$

Desvio padrão

$$Var(cX) = c^2.Var(X)$$

Constante elevada ao quadrado.

$$\sigma(X + c) = \sigma(X)$$

$$Var(X + c) = Var(X)$$

# Medidas de Posição

## Média Aritmética

**Para dados em tabela**

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

**Para dados em rol**

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

**Para dados categorizados**

$$\mu = \frac{\sum_{i=1}^{N} f_i \bar{x}_i}{\sum_{i=1}^{N} f_i}$$

## Mediana

**Para dados em rol ou em Tabela N = ímpar**

Termo central $n = \frac{N+1}{2}$

Fórmula $Md = x_n$

**Para dados em rol ou em Tabela N = par**

$$n = \frac{N}{2}$$

$$Md = \frac{x_n + x_{n+1}}{2}$$

**Para dados categorizados**

$$n = \frac{N}{2} \quad Md = x_n$$

## Moda

**Fórmula de Czuber**

$$Mo = Linf + \left(\frac{\Delta f_{ant}}{\Delta f_{ant} + \Delta f_{post}}\right) \cdot h$$

**Fórmula de King**

$$Mo = Linf + \left(\frac{f_{post}}{f_{ant} + f_{post}}\right) \cdot h$$

# QUESTÕES COMENTADAS EM AULA

**001.** (CESPE/2018/POLÍCIA FEDERAL/PERITO CRIMINAL FEDERAL/ÁREA 6) Considerando que a análise de uma amostra de minério de chumbo tenha apresentado os seguintes resultados percentuais (%): 8,10; 8,32; 8,12; 8,22; 7,99; 8,31, julgue o item a seguir, relativo a esses dados.
A variância dos dados em apreço é dada pelo valor do desvio-padrão ao quadrado.

**002.** (CESPE/CEHAP-PB/2009/ADMINISTRADOR) O custo médio nacional para a construção de habitação com padrão de acabamento normal, segundo levantamento realizado em novembro de 2008, foi de R$ 670,00 por metro quadrado, sendo R$ 400,00/m² relativos às despesas com materiais de construção e R$ 270,00/m² com mão de obra. Nessa mesma pesquisa, os custos médios regionais apontaram para os seguintes valores por metro quadrado: R$ 700,00 (Sudeste), R$ 660,00 (Sul), R$ 670,00 (Norte), R$ 640,00 (Centro-Oeste) e R$ 630,00 (Nordeste).
*Sistema Nacional de Pesquisa de Custos e Índices da Construção Civil. SINAPI/IBGE, nov./2008 (com adaptações).*
O desvio-padrão dos custos médios regionais por metro quadrado foi:
a) inferior a R$ 30,00.
b) superior a R$ 30,01 e inferior a R$ 40,00.
c) superior a R$ 40,01 e inferior a R$ 50,00.
d) superior a R$ 50,01.

**003.** (FCC/TRT-9ª REGIÃO/PR/2020/ANALISTA JUDICIÁRIO/ESTATÍSTICA) Uma população com 16 valores estritamente positivos X1, X2, X3, ..., X16, correspondente a um determinado atributo, apresenta as seguintes informações:

$$\sum_{i=1}^{16} X_i = 192 \text{ e } \sum_{i=1}^{16} X_i^2 = 4.464$$

O elemento $X_{10}$, tal que $X_{10}$ = 12, é retirado da população. Os valores da variância da primeira população e da nova população formada são, respectivamente, iguais a:
a) 144 e 134,40.
b) 144 e 144.
c) 135 e 144.
d) 135 e 135.
e) 135 e 126.

**004.** (FUNRIO/SESAU-RO/2017) Para responder às próximas três questões, considere a seguinte amostra de idades: 18, 15, 24, 20, 22, 21, 19, 30, 20
A estimativa não tendenciosa usual da variância populacional é aproximadamente igual a:

a) 15,8.
b) 17,8.
c) 20,1.
d) 22,2.
e) 23,1.

**005.** (CESPE/TCE-PA/2016/AUDITOR DE CONTROLE EXTERNO) Caso, em uma amostra aleatória de tamanho *n* = 4, os valores amostrados sejam A = {2, 3, 0, 1}, a estimativa de máxima verossimilhança para a variância populacional será igual a 5/3.

**006.** (FCC/BANRISUL/2019/ESCRITURÁRIO) 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.

**007.** (ESAF/AFRFB/2012) Em um concurso público a nota média da prova de inglês foi igual a 7 com desvio-padrão igual a 2. Por outro lado, a nota média da prova de lógica foi igual a 7,5 com desvio-padrão igual a 4. Naná obteve nota 8 em Inglês e nota 8 em Lógica. Nené obteve nota 7,5 em Inglês e 8,5 em Lógica. Nini obteve 7,5 em Inglês e 9 em Lógica. Com relação à melhor posição relativa — ou ao melhor desempenho —, pode-se afirmar que o desempenho de:
a) Naná foi o mesmo em Inglês e Lógica.
b) Nini foi melhor em Lógica do que o de Naná em Inglês.
c) Nené foi melhor em lógica do que o de Naná em Inglês.
d) Nené foi o mesmo em Inglês e Lógica.
e) Nené foi melhor em Lógica do que em Inglês.

# QUESTÕES DE CONCURSO

**008.** (CESPE/SEFAZ-DF/2020) A partir de uma amostra aleatória simples de tamanho n, sabe-se que a média aritmética de uma variável X foi igual a 3. Considerando que os valores possíveis para a variável X sejam -1 e +4, julgue o item que se segue.
O desvio-padrão amostral da variável X foi igual ou superior a 2.

Questão bastante interessante. Uma forma de obter a média aritmética igual a 3 é {−1, +4, +4, +4, +4}. Dessa forma, temos a média aritmética da variável:

$$\mu = \frac{-1+4+4+4+4}{5} = \frac{16-1}{5} = \frac{15}{5} = 3$$

Podemos aproveitar para calcular

$$Var = \frac{(-1-3)^2 + 4.(4-3)^2}{5} = \frac{(-4)^2 + 4.(1)^2}{5} = \frac{16+4}{5} = \frac{20}{5} = 4$$

**Certo.**

**009.** (CESPE/SEFAZ-ES/2013/AUDITOR-FISCAL DA RECEITA ESTADUAL) Considerando que a tabela acima mostra o tempo, em minutos, gasto para a realização de auditorias em seis balanços contábeis, assinale a opção correta a respeito de medidas de variabilidade.

| auditoria | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| tempo | 60 | 90 | 30 | 40 | 50 | 90 |

a) O desvio-padrão amostral foi inferior a 30 minutos.
b) A amplitude total, que representa a diferença entre as observações nas extremidades do conjunto de dados, foi igual a 30 minutos.
c) O desvio médio absoluto em torno da média amostral foi superior a 25 minutos.
d) A variância amostral desse conjunto de dados foi inferior a 600 minutos².
e) O coeficiente de variação, que representa a amplitude entre o quartil superior e o inferior, foi igual a 80 minutos.

Para calcular o desvio-padrão, primeiramente devemos tomar a média das observações.

$$\sigma^2 = \frac{0+900+900+400+100+900}{6} = \frac{3200}{6-1} = \frac{3200}{5} = 640$$

$$\sigma = \sqrt{640} \cong 25{,}3 < 30$$

O próximo passo é montar uma tabela para calcular os desvios em relação à média e os seus respectivos quadrados.

| X | (X – μ) | (X – μ)² |
|---|---|---|
| **60** | (60 – 60) = 0 | $0^2 = 0$ |
| **90** | (90 – 60) = 30 | $30^2 = 900$ |
| **30** | (30 – 60) = –30 | $(-30)^2 = 900$ |
| **40** | (40 – 60) = –20 | $(-20)^2 = 400$ |
| **50** | (50 – 60) = –10 | $(-10)^2 = 100$ |
| **90** | (90 – 60) = 30 | $(30)^2 = 900$ |

$$\sigma^2 = \frac{0 + 900 + 900 + 400 + 100 + 900}{6} = \frac{3200}{6-1} = \frac{3200}{5} = 640$$

$$\sigma = \sqrt{640} \cong 25{,}3 < 30$$

**Letra a.**

---

**010.** (FCC/SEFIN-RO/2010/AUDITOR-FISCAL DE TRIBUTOS ESTADUAIS) A média aritmética de todos os salários dos funcionários em uma repartição pública é igual a R$ 1.600,00. Os salários dos funcionários do sexo masculino apresentam um desvio-padrão de R$ 90,00 com um coeficiente de variação igual a 5%. Os salários dos funcionários do sexo feminino apresentam um desvio-padrão de R$ 60,00 com um coeficiente de variação igual a 4%. Escolhendo aleatoriamente um funcionário desta repartição, a probabilidade dele ser do sexo feminino é igual a:

a) 1/2
b) 1/3
c) 3/4
d) 3/5
e) 2/3

---

Uma questão bem sofisticada. O coeficiente de variação é dado pela razão entre a média e o desvio-padrão. Como foram fornecidos os coeficientes de variação e o desvio-padrão dos grupos masculino e feminino, podemos calcular a média de ambos os grupos:

$$CV = \frac{\sigma}{\mu} \therefore \mu = \frac{\sigma}{CV}$$

Agora, calculemos a média dos grupos feminino e masculino:

$$\mu_{feminino} = \mu_f = \frac{\sigma}{CV} = \frac{60}{0{,}04} = 1500$$

$$\mu_{masculino} = \mu_m = \frac{\sigma}{CV} = \frac{90}{0{,}05} = 1800$$

Sejam p a probabilidade de a pessoa ser do sexo feminino e (1 – p) a probabilidade de a pessoa ser do sexo masculino.

$$\mu = p.1500 + (1 - p).1800 = 1600$$

$$1500p + 1800 - 1800p = 1600$$

$$1800 - 300p = 1600$$

$$300p = 1800 - 1600 = 200$$

$$\therefore p = \frac{200}{300} = \frac{2}{3}$$

**Letra e.**

---

**011.** (CESPE/TELEBRAS/2015/ANALISTA SUPERIOR ADMINISTRATIVO) Considerando que os possíveis valores de um indicador *X*, elaborado para monitorar a qualidade de um serviço de cabeamento residencial para a comunicação de dados, sejam elementos do conjunto {0, 1, 2, 3, 4, 5} e que uma amostra aleatória de 5 residências tenha apontado os seguintes indicadores: 4, 4, 5, 4 e 3, julgue o próximo item.

A variância amostral dos indicadores observados foi igual a 0,5.

---

Vamos seguir o passo a passo. Primeiramente, precisamos calcular a média.

$$\mu = \frac{4 + 4 + 5 + 4 + 3}{5} = \frac{20}{5} = 4$$

Agora, vamos montar uma tabela para calcular os desvios em relação à média.

| X | (X – μ) | (X – μ)² |
|---|---|---|
| **4** | (4 – 4) = 0 | 0² = 0 |
| **4** | (4 – 4) = 0 | 0² = 0 |
| **5** | (4 – 4) = 1 | 1² = 1 |
| **4** | (4 – 4) = 0 | 0² = 0 |
| **3** | (3 – 4) = 0 | (–1)² = 1 |

Podemos, agora, obter a variância amostral.

$$\sigma^2 = \frac{0+0+1+0+1}{5-1} = \frac{2}{4} = \frac{1}{2}$$

**Certo.**

---

**012.** (CESPE/SEFAZ-ES/2008/CONSULTOR DO EXECUTIVO) Um Tribunal de Contas realizou, em determinado ano, um levantamento estatístico para estimar a proporção de processos regulares e irregulares. Como a população de processos irregulares existentes é muito grande, foi tomada uma amostra aleatória simples de 400 processos irregulares. Os processos amostrados foram classificados em dois tipos: A e B. Sabe-se que a quantidade de irregularidades por processo do tipo B segue, aproximadamente, uma distribuição normal, e os resultados por tipo de processo estão na tabela abaixo.

| tipos de processo | total de processos irregulares | quantidade média de irregularidades encotradas por processo | variância amostral da quantidade de irregularidades encontradas por processo |
|---|---|---|---|
| A | 320 | 7 | 14,0 |
| B | 80 | 15 | 12,8 |

O coeficiente de variação da quantidade de irregularidades encontradas por processo do tipo B é superior a 0,75.

---

O coeficiente de variação pode ser obtido como a razão entre o desvio-padrão e a média.

$$CV = \frac{\sigma}{\mu}$$

No enunciado, foi fornecida a variância. Mas, o desvio-padrão pode ser calculado como a raiz quadrada da variância.

$$\sigma^2 = 12{,}8 \therefore \sigma = \sqrt{12{,}8}$$

Como o desvio-padrão é uma raiz quadrada bem difícil de calcular, podemos recorrer à variância relativa que é igual ao quadrado do coeficiente de variação.

$$VR = \frac{\sigma^2}{\mu^2} = \frac{12{,}8}{(15)^2} = \frac{12{,}8}{225} = \frac{51{,}2}{900} \cong 0{,}05689$$

Como a variância relativa é igual ao quadrado do desvio-padrão, o enunciado dizer que o coeficiente de variação é superior a 0,75 é equivalente a ele dizer que a variância relativa é superior ao quadrado de 0,75.

$$CV > 0{,}75 \therefore VR > (0{,}75)^2 = 0{,}5625$$

Dessa forma, como calculamos previamente, a variância relativa foi de 0,057 aproximadamente, que é muito inferior ao valor proposto no enunciado. Portanto, a afirmação do enunciado está incorreta.

**Errado.**

---

**013.** (FCC/SEFAZ-SP/2006/GESTÃO TRIBUTÁRIA) Considerando as respectivas definições e propriedades relacionadas às medidas de posição e de variabilidade, é correto afirmar:

a) Concedendo um reajuste de 10% em todos os salários dos empregados de uma empresa, tem-se também que a respectiva variância fica multiplicada por 1,10.

b) Definindo coeficiente de variação (CV) como sendo o quociente da divisão do desvio-padrão pela respectiva média aritmética (diferente de zero) de uma sequência de valores, tem-se então que CV também poderá ser obtido dividindo a correspondente variância pelo quadrado da média aritmética.

c) Subtraindo um valor fixo de cada salário dos funcionários de uma empresa, tem-se que o respectivo desvio-padrão dos novos valores é igual ao valor do desvio-padrão dos valores anteriores.

d) Dividindo todos os valores de uma sequência de números estritamente positivos por 4, tem-se que o respectivo desvio-padrão fica dividido por 2.

e) Em qualquer distribuição de valores em estudo, a diferença entre a mediana e a moda é sempre diferente de zero.

---

a) Errada. Reajustar os salários em 10% é o mesmo que multiplicar todos os salários por 1,10. Ao multiplicar uma variável aleatória por um termo **a,** sua variância fica multiplicada por **a².** Portanto, nesse caso, a variância ficaria multiplicada por 1,21, portanto, seria acrescentada em 21%.

b) Errada. A razão entre a variância e o quadrado da média é também conhecida como variância relativa, que corresponde ao quadrado do desvio-padrão.

c) Certa. Uma das propriedades mais importantes do desvio-padrão é que, quando acrescentamos um valor fixo a todas as observações de uma variável aleatória, o desvio-padrão não se altera.

d) Errada. Quando dividimos todos as observações da variável por 4, o desvio-padrão também fica dividido por 4. A variância ficaria dividida por $4^2 = 16$.

e) Errada. Nem sempre. Vejamos um exemplo oposto:

$$\{1, 3, 3, 4\}$$

Nesse caso, a moda é igual a 3, pois é a observação mais frequente (aparece duas vezes).

**Letra c.**

---

**014.** (CESPE/SEFAZ-MT/2004/AGENTE DE TRIBUTOS ESTADUAIS) Um órgão do governo recebeu pela Internet denúncias de sonegação de impostos estaduais contra 600 pequenas empresas. Denúncias contra outras 200 pequenas empresas foram encaminhadas pessoalmente para esse órgão. Para a apuração das denúncias, foram realizadas auditorias nas 800 empresas denunciadas. Como resultado dessas auditorias, foi elaborada a tabela abaixo, que apresenta um quadro das empresas denunciadas e os correspondentes débitos fiscais ao governo. Das empresas denunciadas, observou-se que apenas 430 tinham débitos fiscais.

| | valor do débito fiscal (VDF), em R$ mil, apurado após auditoria na empresa denunciada | | | | |
|---|---|---|---|---|---|
| forma de recebimento da denúncia | 0 < VDF < 1 | 1 • •VDF < 2 | 2 • •VDF < 3 | 3 • •VDF • •4 | total |
| pela Internet | 60 | 100 | 50 | 30 | 240 |
| pessoalmente | 20 | 120 | 40 | 10 | 190 |
| total | 80 | 220 | 90 | 40 | 430* |

Nota: *Para as demais empresas, VDF = 0.

Com base na situação hipotética acima e de acordo com as informações apresentadas, julgue o item que se segue.

A amplitude total dos débitos fiscais é uma medida de variabilidade.

---

Sim. Com certeza. A amplitude total é uma medida de variabilidade, pois mostra o quanto as observações da amostra variam em torno da média.

Trata-se de uma medida mais simples, pois leva em consideração apenas as observações extremas, mas é uma medida de variabilidade.

**Certo.**

---

**015.** (CESPE/ANTAQ/2014/TÉCNICO EM REGULAÇÃO DE SERVIÇOS DE TRANSPORTE AQUAVIÁRIO) A tabela acima apresenta os resultados de uma pesquisa de satisfação realizada em uma amostra de usuários dos serviços de transporte fluvial prestados por uma empresa. Com base nessas informações e na tabela, julgue o próximo item.

| nota atribuída pelo passageiro | frequência |
|---|---|
| 0 | 15 |
| 1 | 30 |
| 2 | 45 |
| 3 | 50 |
| 4 | 35 |
| 5 | 5 |

O desvio-padrão da série de notas obtidas pela empresa é inferior àquele que seria obtido caso todos os usuários tivessem avaliado a empresa com as notas 2 ou 3.

Se todas as notas fossem iguais a 2 ou 3, a variabilidade delas em torno da média seria menor. Por consequência, menor também seria o desvio-padrão.

**Errado.**

**016.** (CESPE/DEPEN/2015/AGENTE FEDERAL DE EXECUÇÃO PENAL) Considerando os dados da tabela mostrada, que apresenta a distribuição populacional da quantidade diária de incidentes (N) em determinada penitenciária, julgue o item que se segue.

| quantidade diária de incidentes ($N$) | frequência relativa |
|---|---|
| 0 | 0,1 |
| 1 | 0,2 |
| 2 | 0,5 |
| 3 | 0,0 |
| 4 | 0,2 |
| total | 1 |

A amplitude total da distribuição é igual a 5, pois há cinco valores possíveis para a variável N.

A amplitude é igual à diferença entre a maior e a menor observação. Não tem nada a ver com a contagem do total de valores possíveis.

Nesse caso, a amplitude da amostra é A = 4 – 0 = 4.

**Errado.**

**017.** (CESPE/2018/POLÍCIA FEDERAL/PERITO CRIMINAL FEDERAL/ÁREA 1)

| | 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 kg.

Primeiramente, vamos calcular a média amostral.

$$\mu = \frac{10 + 22 + 18 + 22 + 28}{5} = \frac{100}{5} = 20$$

Agora, vamos aplicar a expressão da variância **amostral**. Como temos uma amostra e vamos calcular a variância amostral, devemos lembrar que, no denominador, aparece o termo N – 1, em que N é o número de elementos da amostra.

$$Var = \frac{(10-20)^2 + (22-20)^2 + (18-20)^2 + (22-20)^2 + (28-20)^2}{5-1} =$$

$$Var = \frac{(-10)^2 + (2)^2 + (-2)^2 + (2)^2 + (8)^2}{4}$$

$$Var = \frac{100 + 4 + 4 + 4 + 64}{4} = \frac{176}{4} = 44$$

O desvio-padrão, por sua vez, é igual à raiz quadrada da variância.

$$\sigma = \sqrt{Var} = \sqrt{44} < 7$$

De fato, a raiz quadrada de 44 é menor que 7. Podemos verificar isso facilmente, pois $7^2 = 49$. Portanto, a raiz de quadrada de 44 deve ser um número menor que 7.
**Certo.**

**018.** (ESAF/ANAC/2016/ANALISTA ADMINISTRATIVO) Os valores a seguir representam uma amostra
3 3 1 5 4 6 2 4 8

Então, a variância dessa amostra é igual a

a) 4,0
b) 2,5.
c) 4,5.
d) 5,5
e) 3,0

Vamos usar o procedimento para o cálculo de variância. Primeiramente, precisamos calcular a média amostral.

$$\mu = \frac{3+3+1+5+4+6+2+4+8}{9} = \frac{36}{9} = 4$$

Agora, vamos calcular todos os desvios em relação à média.

| X | X – μ | (X – μ)² |
|---|---|---|
| **3** | (3 – 4) = –1 | (–1)² = 1 |
| **3** | (3 – 4) = –1 | (–1)² = 1 |
| **1** | (1 – 4) = –3 | (–3)² = 9 |
| **5** | (5 – 4) = 1 | (1)² = 1 |
| **4** | (4 – 4) = 0 | (0)² = 0 |
| **6** | (6 – 4) = 2 | (2)² = 4 |
| **2** | (2 – 4) = –2 | (–2)² = 4 |
| **4** | (4 – 4) = 0 | (0)² = 0 |
| **8** | (8 – 4) = 4 | (4)² = 16 |

Agora, devemos somar todos os desvios em relação à média. Como temos uma amostra, devemos observar que devemos dividir pelo número de elementos da amostra menos 1 – pois devemos fazer a correção para o desvio-padrão amostral.

$$V\left(arX = \frac{1+1+9+1+0+4+4+0+16}{9-1} = \frac{36}{8} = 4{,}5\right)$$

**Letra c.**

**019.** (ESAF/MTUR/2014) Uma variável aleatória x qualquer possui média igual a 4 metros. Sabendo-se que a média do quadrado dos valores de x é igual a 36 m², então a variância relativa de x é igual a:

a) 1,25 m²
b) 1,25 m
c) 1,25
d) 5 m²
e) 5

A variância de uma população pode ser calculada pela média do quadrado menos o quadrado da média. Observe que a questão também colocou as unidades nas respostas, portanto, vamos tomar cuidado com as unidades.

$$Var = E[X^2] - \mu^2 = 36 - 4^2 = 36 - 16 = 20\ m^2$$

Agora, vamos calcular a variância relativa, que é dada pela razão entre a variância e a média ao quadrado.

$$VR = \frac{Var}{\mu^2} = \frac{20\ m^2}{(4\ m^2)} = \frac{20\ m^2}{16\ m^2} = 1{,}25$$

**Letra c.**

**020.** (CESPE/TCE-PA/2016/AUDITOR DE CONTROLE EXTERNO) 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.

| 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 variância de *X* é inferior a 2,5.

Vamos seguir o procedimento de cálculo da variância. O primeiro passo é calcular a média.

$$\mu = 0{,}3.0 + 0{,}1.1 + 0{,}2.2 + 0{,}1.3 + 0{,}3.4$$

$$\mu = 0 + 0{,}1 + 0{,}4 + 0{,}3 + 1{,}2 = 2{,}0$$

Para calcular a variância, devemos tomar a média dos desvios de cada observação em relação à média.

| X | (X – μ) | (X – μ)² | Frequência Relativa |
|---|---|---|---|
| **0** | (0 – 2) = –2 | (–2)² = 4 | 0,3 |
| **1** | (1 – 2) = –1 | (–1)² = 1 | 0,1 |
| **2** | (2 – 2) = 0 | (0)² = 0 | 0,2 |
| **3** | (3 – 2) = 1 | (1)² = 1 | 0,1 |
| **4** | (4 – 2) = 2 | (2)² = 4 | 0,3 |

Por fim, vamos calcular a média do quadrado dos desvios. Para isso, basta utilizar as frequências relativas como pesos para os quadrados dos desvios.

$$Var = 0{,}3.4 + 0{,}1.1 + 0{,}2.0 + 0{,}1.1 + 0{,}3.4$$

$$Var = 1{,}2 + 0{,}1 + 0 + 0{,}1 + 1{,}2 = 2{,}6$$

**Errado.**

---

**021.** (FCC/PREFEITURA DO RECIFE-PE/2019/ANALISTA DE GESTÃO ADMINISTRATIVA) 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.

---

a) Errada. Quando multiplicamos todos os elementos de P por 16, o desvio-padrão será multiplicado também por 16.
b) Certa. Quando dividimos todos os elementos de P por 2, a variância da população será igual à variância de P dividida por $2^2$. Dividir por $2^2 = 4$ é o mesmo de multiplicar por 0,25. Isso acontece, porque:

$$\frac{1}{4} = 0{,}25$$

c) Errada. Ao adicionar uma constante aditiva K > 0, a média aritmética também é acrescida de K, porém, o desvio-padrão e a variância se mantêm inalterados.
d) Errada. Item interessante. Vejamos em que situações a variância é igual ao desvio-padrão.

$$\sigma^2 = \sigma$$

$$\sigma^2 - \sigma = 0$$

$$\sigma(\sigma - 1) = 0$$

$$\therefore \sigma = 0 \; ou \; \sigma = 1$$

Se todos os elementos de P são iguais, o desvio-padrão será igual a zero. Nesse caso, realmente, o desvio-padrão da amostra será igual à variância.
Porém, o erro da afirmação é dizer que essa é a única situação. É possível também que o desvio-padrão seja igual à variância, quando ambos forem iguais a 1.
e) Errada. Ao se subtrair uma constante K, o desvio-padrão se mantém inalterado. Afirmação incorreta.
**Letra b.**

---

**022.** (CESPE/IPHAN/2018/ANALISTA I/ÁREA 2) 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.

O coeficiente de variação é dado pela razão entre o desvio-padrão e a média amostral.

$$CV = \frac{\sigma}{\mu} = \frac{1,2}{3} = 0,4$$

De fato, o coeficiente de variação é superior a 0,3 e inferior a 0,5.
**Certo.**

**023.** (FCC/ARTESP/2017/ESPECIALISTA EM REGULAÇÃO DE TRANSPORTE) 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.

O primeiro ponto que nos chama a atenção nessa questão é que ela fala em **uma amostra.** Portanto, devemos usar o fator de correção (N−1)/N.
Como temos um rol relativamente grande, a forma mais simples de resolver é usar a expressão de que a variância é igual à média dos quadrados menos o quadrado da média, multiplicada pelo fator de correção para o desvio-padrão amostral.

$$E[X^2] = \frac{36^2 + 28^2 + 12^2 + 5^2 + 3^2 + 2^2 + 2^2 + 4^2 + 9^2 + 11^2 + 22^2 + 38^2}{12}$$

$$E[X^2] = \frac{1296 + 784 + 144 + 25 + 9 + 4 + 4 + 16 + 81 + 121 + 484 + 1444}{12}$$

$$E[X^2] = \frac{4412}{12} = \frac{1103}{3}$$

Agora, vamos calcular a média aritmética.

$$\mu = \frac{36+28+12+5+3+2+2+4+9+11+22+38}{12} = \frac{172}{12} = \frac{43}{3}$$

Vamos calcular a expressão da variância populacional.

$$\sigma_{POP}^2 = E[X^2] - \mu^2$$

$$\sigma_{POP}^2 = \frac{1103}{3} - \left(\frac{43}{3}\right)^2 = \frac{1103}{3} - \frac{1849}{9} = \frac{3309 - 1849}{9} = \frac{1460}{9}$$

Agora, vamos aplicar o fator de correção para a variância.

$$\sigma_{AM}^2 = \frac{N}{N-1} \cdot \sigma_{POP}^2$$

$$\sigma_{AM}^2 = \frac{12}{11} \cdot \frac{1460}{9} = \frac{4}{11} \cdot \frac{1460}{3} \cong 177$$

Por fim, vamos extrair a raiz quadrada.

$$\sigma = \sqrt{177}$$

O modo mais simples de extrair a raiz quadrada é testando as alternativas.

$$(13{,}30)^2 = 176{,}89$$

Portanto, a letra A é realmente muito próxima.

$$\sigma = \sqrt{177} \cong 13{,}30$$

**Letra a.**

---

**024.** (FGV/IBGE/2017/ANALISTA CENSITÁRIO/MÉTODOS QUANTITATIVOS/ADAPTADA) 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 medida adimensional de dispersão de A é superior à de B.

---

a) Errada. O desvio-padrão dos preços em B é:

$$\sigma_{maçã} = \sqrt{400} = 20\ maçãs$$

b) Errada. Precisamos saber com quantas peras se compra uma maçã.

$$tx = \frac{100}{120} \cong 0{,}833$$

c) Certa. Vamos calcular os coeficientes de variação.

$$CV_A = \frac{\sigma_{pera}}{\mu_{pera}} = \frac{25}{100} = 0{,}25$$

$$CV_B = \frac{\sigma_{maçã}}{\mu_{maçã}} = \frac{20}{120} = 0{,}167$$

Observe que o coeficiente de variação de A é realmente superior ao coeficiente de variação de B.
**Letra c.**

---

**025.** (CESPE/IPHAN/2018/ANALISTA I/ÁREA 2) 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.
Para esse conjunto de valores, a variância é igual a 3.

---

A variância é igual ao quadrado do desvio-padrão.

$$Var = \sigma^2 = (1{,}2)^3 = 1{,}44$$

Portanto, afirmação está errada.
**Errado.**

---

**026.** (CESPE/2018/POLÍCIA FEDERAL/PERITO CRIMINAL FEDERAL/ÁREA 6) Considerando que a análise de uma amostra de minério de chumbo tenha apresentado os seguintes resultados percentuais (%): 8,10; 8,32; 8,12; 8,22; 7,99; 8,31, julgue o item a seguir, relativo a esses dados.

O coeficiente de variação da análise é dado pela razão entre o desvio-padrão e a média, multiplicada por 100%

É isso mesmo. O coeficiente de variação é igual à razão entre o desvio-padrão e a média. Para ser expresso em termos percentuais, podemos multiplicá-lo por 100%.

**Certo.**

**027.** (CESPE/2018/POLÍCIA FEDERAL/PERITO CRIMINAL FEDERAL/ÁREA 6) Considerando que a análise de uma amostra de minério de chumbo tenha apresentado os seguintes resultados percentuais (%): 8,10; 8,32; 8,12; 8,22; 7,99; 8,31, julgue o item a seguir, relativo a esses dados.

O desvio-padrão da análise em apreço é dado pela raiz quadrada do valor médio dividido pelo número de amostras, no caso, 6.

Essa não é a definição de desvio-padrão. A expressão dada pelo enunciado foi:

$$Exp = \sqrt{\frac{\mu}{6}}$$

Porém, o desvio-padrão corresponde à raiz quadrada da variância. Como se tem uma amostra, devemos lembrar de subtrair uma unidade no denominador.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{6}(x_i - \mu)^2}{6 - 1}}$$

Portanto, a expressão citada no enunciado não corresponde ao desvio-padrão.

**Errado.**

**028.** (FGV/IBGE/2017/ANALISTA CENSITÁRIO/MÉTODOS QUANTITATIVOS) Uma amostra de cinco indivíduos é extraída aleatoriamente de uma dada população, obtendo-se os seguintes valores:

$X_1 = 3$, $X_2 = 5$, $X_3 = 4$, $X_4 = 7$ e $X_5 = 11$

Então a variância amostral e a estimativa não tendenciosa da variância populacional seriam iguais a, respectivamente:

**a)** 8 e 10;
**b)** 9 e 8;
**c)** 8 e 12;

**d)** 8 e 6,67;
**e)** 6,67 e 10.

A banca adotou uma nomenclatura bastante arcaica aqui. A variância amostral deve ser sempre calculada com N – 1 no denominador, pois essa é a estimativa não tendenciosa para o desvio-padrão populacional.
No entanto, o objetivo dessa questão, era que o aluno calculasse a variância amostral com N no denominador e a estimativa não tendenciosa para o desvio-padrão populacional.
Para fazer as contas, vamos primeiramente obter a média.

$$\bar{x} = \frac{3+5+4+7+11}{5} = \frac{30}{5} = 6$$

Agora, façamos as contas do desvio-padrão.

$$\sigma^2 = \frac{(3-6)^2 + (5-6)^2 + (4-6)^2 + (7-6)^2 + (11-6)^2}{5}$$

$$\sigma^2 = \frac{(-3)^2 + (-1)^2 + (-2)^2 + (1)^2 + 5^2}{5} = \frac{9+1+4+1+25}{5} = \frac{40}{5} = 8$$

Agora, vamos aplicar o fator de ajuste para obter a estimativa não tendenciosa.

$$\sigma^2 = \frac{N}{N-1} \cdot 8 = \frac{5}{4} \cdot 8 = 10$$

**Letra a.**

---

**029.** (FCC/MPE-AP/2012/ANALISTA MINISTERIAL/ADMINISTRAÇÃO) Ao considerar uma curva de distribuição normal, com uma média como medida central, temos a variância e o desvio-padrão referentes a esta média. Em relação a estes parâmetros,
**a)** a variância é uma medida cujo significado é a metade do desvio-padrão.
**b)** a variância é calculada com base no dobro do desvio-padrão.
**c)** o desvio-padrão é a raiz quadrada da variância.
**d)** a média dividida pelo desvio-padrão forma a variância.
**e)** a variância elevada ao quadrado indica qual é o desvio-padrão.

A relação entre o desvio-padrão e a variância é conhecida e deve estar no seu sangue de concurseiro. O desvio-padrão é a raiz quadrada da variância, ou ainda que a variância é o desvio-padrão ao quadrado.
Exatamente o que está dito na letra C.
**Letra c.**

---

**030.** (FGV/SEFAZ-RJ/2013/ANALISTA DE CONTROLE INTERNO) 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

Primeiramente, calcularemos a média dessa população.

$$\mu = \frac{2+4+2+4+2+4+2+4}{8} = \frac{4.(2+4)}{8} = \frac{4.6}{8} = 3$$

Agora, calcularemos os desvios. Como se trata de uma população, devemos dividir a variância pelo número de elementos da amostra (N = 8).

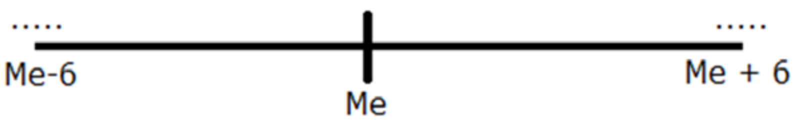
$$S^2 = \frac{4.(2-3)^2 + 4.(4-3)^2}{8} = \frac{4.(-1)^2 + 4.1^2}{8} = \frac{8}{8} = 1$$

$$\sigma = \sqrt{1} = 1$$

**Letra b.**

**031.** (CESPE/PF/2012) Se a amplitude observada em um conjunto de dados formado por 10 elementos for igual a 12, então a variância desse conjunto de dados será inferior a 120.

Questão interessante, pois nos obriga a pensar no caso em que existe maior variância.
A variância será máxima quando os termos estiverem o mais distante possível da média. Nesse caso, teríamos 5 à esquerda e 5 à direita.




Nesse caso, todos os elementos distam 6 da média. Portanto, temos que a variância máxima será igual a:

$$S^2_{MAX} = \frac{5.(-6)^2 + 5.6^2}{10} = \frac{180+180}{10} = 36 < 120$$

**Certo.**

**032.** (CESPE/MEC/2009/AGENTE ADMINISTRATIVO) Em uma escola, há 2 mil estudantes distribuídos em 100 turmas: 50 são do turno matutino e as outras 50, do turno vespertino. 60% dos alunos estão matriculados no período matutino e 40% estão matriculados no período vespertino.
A média das idades dos estudantes matriculados no turno vespertino é 10% superior à média das idades dos estudantes do turno matutino. A variância das idades daqueles que estudam no turno matutino $(\sigma_M^2)$.
A média das idades dos estudantes matriculados no turno vespertino é 10% superior à média das idades dos estudantes do turno matutino. A variância das idades daqueles que estudam no turno matutino $(\sigma_M^2)$ é igual à variância das idades dos estudantes do turno vespertino $(\sigma_V^2)$. Com base nessas informações, julgue o seguinte item.
Com respeito à dispersão, a variância da distribuição das idades dos 2 mil estudantes é igual a $(\sigma_M^2)$ .

Questão bastante interessante. A forma mais simples resolver essa questão é entender que os dois turnos vespertino e matutino são dois estratos da população de alunos daquela escola.
Dissemos que são estratos, porque o fato de um aluno estudar pela manhã ou pela tarde influencia na sua média de idade.
Como são estratos, eles necessariamente são mais homogêneos que a população inteira. Logo, termos que $\sigma^2 > \sigma_M^2$.
**Errado.**

**033.** (FCC/TRF-2ª REGIÃO/2012/ANALISTA JUDICIÁRIO/ESTATÍSTICA) 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

Primeiramente, calcularemos a variância populacional.

$$Var = \sigma^2 = 0{,}2^2 = 0{,}04$$

Sabemos que a variância é igual à média dos quadrados menos o quadrado da média. A medida dos quadrados é dada pela razão entre a soma dos quadrados e o tamanho da população.

$$E[X^2] = \frac{65,6}{20} = 3,28$$

Agora, calcularemos a média aritmética.

$$Var = E[X^2] - (E[X])^2$$

$$\therefore 0,04 = 3,28 - (\mu)^2 \therefore \mu^2 = 3,28 - 0,04 = 3,24$$

$$\therefore \mu = \sqrt{3,24} = \sqrt{\frac{324}{100}} = \frac{18}{10} = 1,8$$

**Letra c.**

---

**034.** (FGV/AL-BA/2014/AUDITOR) A média das idades de um grupo de 4 amigos é 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

---

Questão interessante falando sobre os efeitos de um deslocamento no tempo sobre a média e a variância.
Como todos os amigos aumentarão sua idade em 5 anos, a média também aumentará 5 anos, portanto, passará a 41 anos.
Por outro lado, as diferenças de idades entre esses amigos, que vão compor o desvio-padrão, não variam. Sendo assim, o desvio-padrão continuará igual a 2. Como a questão pediu a variância, essa é o quadrado do desvio-padrão, portanto, Var = $2^2$ = 4.
**Letra a.**

---

**035.** (FCC/TRF-2ª REGIÃO/2012/ANALISTA JUDICIÁRIO/ESTATÍSTICA) Um censo realizado em duas empresas Alfa e Beta revelou que os coeficientes de variação correspondentes dos salários de seus empregados foram 10% e 5%, respectivamente. Sabe-se que a soma das médias aritméticas dos salários das duas empresas é igual a R$ 3.400,00 e o desvio-padrão da empresa Beta é igual a 9⁄16 do desvio-padrão da empresa Alfa. A soma dos respectivos valores das variâncias, em (R$) 2, das duas empresas, é igual a:

a) 33.700.
b) 35.000.
c) 43.100.
d) 51.200.
e) 62.500.

Os coeficientes de variação correspondem à razão entre o desvio-padrão e a média.

$$CV_\alpha = \frac{\sigma_\alpha}{\bar{\alpha}} = 0{,}10 \therefore \sigma_\alpha = 0{,}10.\bar{\alpha}$$

Podemos calcular o desvio-padrão para a empresa beta.

$$CV_\beta = \frac{\sigma_\beta}{\bar{\beta}} = 0{,}05 \therefore \sigma_\beta = 0{,}05.\bar{\beta}$$

Usando o fato de que o desvio-padrão de beta é 9/16 do desvio-padrão de alfa, temos:

$$\sigma_\beta = 0{,}05\,\bar{\beta} = \frac{9}{16}.0{,}10.\bar{\alpha} \therefore \bar{\beta} = \frac{9}{16}.\frac{0{,}10}{0{,}05}\,\bar{\alpha} = \frac{9}{8}\bar{\alpha}$$

E agora a informação da média aritmética:

$$\bar{\alpha} + \bar{\beta} = 3400$$

Substituindo a primeira na segunda, temos:

$$\frac{9}{8}\bar{\alpha} + \bar{\alpha} = 3400 \therefore \frac{17}{8}\bar{\alpha} = 3400 \therefore \bar{\alpha} = \frac{3400.8}{17} = 200.8 = 1600$$

Agora, podemos obter os desvios padrão de cada empresa.

$$\sigma_\alpha = 0{,}10.\bar{\alpha} = 0{,}10.1600 = 160$$

$$\bar{\alpha} + \bar{\beta} = 6800 \therefore \bar{\beta} = 3400 - 1600 = 1800$$

$$\sigma_\beta = 0{,}05.\bar{\beta} = 0{,}05.1800 = 90$$

Como a questão pediu a soma das variâncias, precisamos nos lembrar que a variância é igual ao quadrado do desvio-padrão.

$$\sigma_\alpha^2 + \sigma_\beta^2 = 160^2 + 90^2 = 25600 + 8100 = 33700$$

**Letra a.**

**036.** (FGV/TCM-SP/2015/AGENTE DA FISCALIZAÇÃO/ECONOMIA) Para estimar o desvio--padrão (variabilidade) do preço de um determinado item de consumo, essencial à administração pública, é extraída uma pequena amostra aleatória simples, através de consultas a quatro fornecedores distintos. Supondo que os valores encontrados foram R$4,00, R$5,00, R$8,00 e R$11,00, a estimativa para a variabilidade, obtida através do estimador não tendencioso da variância, será:

a) $\sqrt{7}$
b) $\sqrt{7,5}$
c) $\sqrt{10}$
d) 7,5
e) 10

Primeiramente, devemos encontrar a média amostral.

$$\mu = \frac{4+5+8+11}{4} = \frac{28}{4} = 7$$

Para calcular a variância, devemos nos lembrar que, como se trata de uma amostra, devemos dividir por N – 1:

$$S^2 = \frac{(4-7)^2+(5-7)^2+(8-7)^2+(11-7)^2}{4-1} = \frac{(-3)^2+(-2)^2+1^2+4^2}{3}$$

$$S^2 = \frac{9+4+1+16}{3} = \frac{30}{3} = 10$$

Como a questão pediu o desvio-padrão, esse é igual à raiz quadrada da variância.

$$\sigma = \sqrt{10}$$

**Letra c.**

**037.** (FGV/SENADO FEDERAL/2008/ESTATÍSTICO) 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%.

Nesse caso, ao reajustar os salários em 20%, a média será reajustada em 20% e também todos os desvios também serão reajustados em 20%. Provaremos que o desvio-padrão também aumentará 20%.

Sejam $x_i$ os valores das observações da variável X e μ a média populacional, o desvio-padrão da população inicial era dado por:

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

Após o aumento de todos os elementos e da média, esses passaram a ser x' e μ', teremos que o novo desvio-padrão será:

$$\sigma'^2 = \frac{\sum (x'-\mu')^2}{N} = \frac{\sum (1{,}2x - 1{,}2\mu)^2}{N} = \frac{(1{,}2)^2 \sum (x-\mu)^2}{N} = (1{,}2)^2 \sigma^2$$

$$\therefore \sigma' = 1{,}2.\sigma$$

Dessa maneira, o coeficiente de variação de Pearson:

$$CV' = \frac{\sigma'}{\mu'} = \frac{1{,}2.\sigma}{1{,}2.\mu} = \frac{\sigma}{\mu} = CV$$

**Letra c.**

---

**038.** (FGV/SUSAM/2014/ECONOMISTA) Considere as seguintes informações obtidas a partir de uma amostra aleatória:

$$\sum_{i=1}^{100} x_i^2 = 100 \qquad \sum_{i=1}^{100} x_i = 10$$

Considerando que $N \approx N - 1$, é correto afirmar que o desvio-padrão da variável X é:

a) 0
b) $\sqrt{0{,}99}$
c) 1
d) $\sqrt{0{,}01}$
e) 99/100

Uma questão, no mínimo, controversa.

Ao dizer que N é aproximadamente igual a N-1, a questão nos pede para ignorar o fator de correção que transforma o desvio-padrão populacional em amostral.

Sendo assim, devemos calcular:

$$Var = E[X^2] - \mu^2 = \frac{100}{100} - \left(\frac{10}{100}\right)^2 = 1 - 0{,}01 = 0{,}99$$

$$\therefore \sigma = \sqrt{0{,}99}$$

É, no mínimo, estranho, uma questão pedir o valor aproximado e colocar entre as opções de resposta o valor exato que poderia ser obtido aplicando o fator de correção.

$$\sigma_{EXATO} = \sqrt{\frac{100}{99}}.\sqrt{0{,}99} = \sqrt{1} = 1$$

Porém, devemos marcar a letra b), pois o comando do enunciado é bastante claro ao nos mandar ignorar o fator de correção.

**Letra b.**

---

**039.** (CESGRANRIO/CHESF/2012) Uma pesquisa gerou um conjunto de valores tais que:
a média de todos os valores é 50;
a soma dos quadrados dos valores é 150.000;
o tamanho da população é 50.
Se de cada um dos valores for subtraída a média, e, em seguida, o resultado de cada subtração for dividido por 10, obtém-se um novo conjunto de valores.
A variância desses valores transformados é:
a) 4,5
b) 5
c) 30
d) 45
e) 50

---

A variância da população original é:

$$Var(X) = E[X^2] - \mu^2 = \frac{150000}{50} - (50)^2 = 3000 - 2500 = 500$$

A transformação citada no enunciado consiste em obter uma variável Z, tal que:

$$Z = \frac{X - \mu}{10} \therefore Var(Z) = \frac{1}{10^2}Var(X) = \frac{1}{100}.500 = 5$$

**Letra b.**

---

**040.** (CESGRANRIO/PETROBRAS/2011/GEOFÍSICO JÚNIOR) Foram realizadas medidas das temperaturas máximas noturnas de 5.000 municípios de algumas regiões, gerando-se uma amostra de variância V. No entanto, descobriu-se que todos os termômetros utilizados subtraíram 4 graus em todas as medidas. Após as devidas correções, a variância obtida será:

a) 4V
b) 2V
c) V
d) V/2
e) V/4

Seja X a variável aleatória que representa a temperatura medida. A variável Z = X + 4 representa as medidas corretas. Após as devidas correções, deseja-se a variância de Z. Temos que:

$$Var(Z) = Var(X + 4) = Var(X) = V$$

**Letra c.**

**041.** (FCC/SEFAZ-BA/2019/AUDITOR-FISCAL/MUITO DIFÍCIL) 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.

Uma questão bem difícil.
O primeiro ponto é notar que $P_1$ é uma população. Portanto, estamos falando na questão do desvio-padrão populacional.
Como foi fornecido o coeficiente de variação e a média populacional, podemos calcular o desvio-padrão da população $P_1$.

$$CV_1 = \frac{\sigma_1}{\mu_1} \therefore \sigma_1 = CV_1 \cdot \mu_1 = 0{,}30.20 = 6$$

Agora, podemos obter a variância da população:

$$Var = S^2 = 6^2 = 36$$

Vamos aplicar a relação de que a variância é igual à média dos quadrados menos o quadrado da média.

$$Var = E[X^2] - \mu^2 = 36$$

Como nos foi fornecida a média da população, podemos calcular o termo E[X²].

$$E[X^2] - 20^2 = 36$$

$$E[X^2] - 20^2 = 36$$

De posse da média e também da média dos quadrados, podemos calcular as somas.

$$\mu = \frac{\sum x_i}{N} = \frac{S_X}{20} \therefore S_X = 20.\mu = 20.20 = 400$$

$$E[X^2] = \frac{\sum x_i^2}{N} = \frac{S_{X^2}}{20} \therefore S_{X^2} = 20.E[X^2] = 20.436 = 8720$$

Criamos as denominações $S_X$ e $S_{X^2}$ para facilitar e não precisar escrevermos toda hora o somatório.

Agora, vamos passar para a situação em que foram retirados dois elementos da amostra iguais a 11. Essa retirada afetará tanto a média como a média dos quadrados.

Vamos calcular os novos valores das somas dos termos da sequência e da soma dos quadrados.

$$S'_X = (Soma\ anterior) - 2.11$$

$$S'_X = (S_X) - 2.11$$

$$S'_X = 400 - 22 = 378$$

Portanto, a nova média populacional será:

$$\mu = \frac{\sum x_i}{N-2} = \frac{378}{18} = 21$$

Façamos o mesmo para a soma dos quadrados.

$$S'_{X^2} = (Soma\ anterior) - 2.(11)^2$$

$$S'_{X^2} = 8720 - 2.121$$

$$S'_{X^2} = 8720 - 242 = 8478$$

Agora, vamos calcular a nova média dos quadrados.

$$E[X^2] = \frac{S'_{X^2}}{N-2} = \frac{8478}{18} = 471$$

Por fim, vamos calcular a variância da nova população $P_2$.

$$Var = E[X^2] - \mu^2$$

$$Var = 471 - 21^2$$

$$Var = 471 - 441 = 30$$

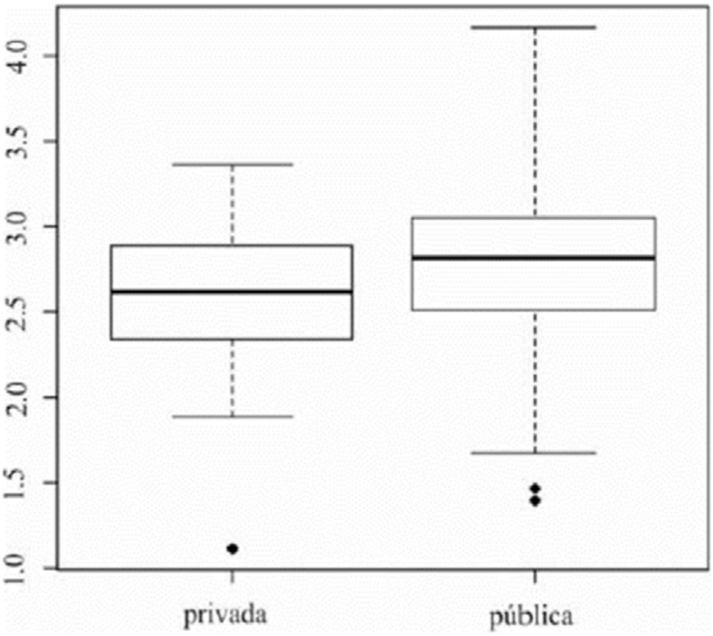
A variância relativa é dada pela razão entre a variância e o quadrado da média.

$$VR = \frac{Var}{\mu^2} = \frac{30}{21^2} = \frac{30}{441} = \frac{10}{147}$$

**Letra a.**

---

**042.** (CESPE/FUB/2015/ESTATÍSTICO) O conceito médio da graduação (G) é um indicador calculado pelo INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira) para a avaliação da qualidade dos cursos de graduação das instituições de ensino superior. A figura apresentada mostra, esquematicamente, as distribuições desse indicador nas instituições privadas e públicas, referentes ao ano de 2013, e a tabela apresenta algumas estatísticas descritivas referentes a essas distribuições.



| | privada | pública |
|---|---|---|
| média amostral | 2,6 | 2,8 |
| desvio padrão amostral | 0,36 | 0,48 |
| primeiro decil (D1) | 2,2 | 2,3 |
| primeiro quartil (Q1) | 2,3 | 2,5 |
| mediana (Q2) | 2,6 | 2,8 |
| terceiro quartil (Q3) | 2,9 | 3,1 |
| nono decil (D9) | 3,1 | 3,4 |
| mínimo | 1,1 | 1,4 |
| máximo | 3,3 | 4,2 |

Com base nessas informações, julgue o próximo item.

A comparação entre os coeficientes de variação das distribuições do indicador G nas instituições privadas e públicas permite concluir que a primeira distribuição é mais homocedástica que a segunda.

Vamos comparar os coeficientes de variação das duas variáveis.

$$CV_1 = \frac{desvio - padrão}{média} = \frac{0{,}36}{2{,}6} \cong 0{,}138 \qquad CV_2 = \frac{desvio - padrão}{média} = \frac{0{,}48}{2{,}8} \cong 0{,}171$$

A primeira distribuição apresenta um coeficiente de variação menor que o da segunda. Portanto, ela é mais homogênea ou mais homocedástica.

**Certo.**

**043.** (CESPE/TCE-ES/2012/AUDITOR) Em pesquisa realizada para se estimar o salário médio dos empregados de uma empresa, selecionou-se, aleatoriamente, uma amostra de nove empregados entre todos os empregados da empresa. Os dados de tempo de serviço, em anos, e salário, em quantidade de salários mínimos, dos indivíduos dessa amostra estão dispostos na tabela abaixo.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| tempo de serviço (anos) | 3 | 2 | 6 | 7 | 4 | 8 | 2 | 2 | 2 |
| salário (quantidade de salários mínimos) | 6 | 6 | 10 | 8 | 5 | 9 | 6 | 5 | 6 |

A partir dos dados da tabela, julgue o item seguinte.

A estimativa não viciada da variância dos salários dos indivíduos da amostra com mais de 5 anos de serviço é igual a 2/3.

Como foi pedida a variância somente dos servidores com mais de 5 anos de serviço, devemos extrair esse conjunto da amostra fornecida pelo enunciado.

| Tempo de Serviço | Salário |
|---|---|
| **6** | 10 |
| **7** | 8 |
| **8** | 9 |

Vamos seguir o procedimento para o cálculo da variância dos salários dos indivíduos:
Calculamos a média amostral:

$$\mu = \frac{10+8+9}{3} = \frac{27}{3} = 9$$

Calculemos os desvios dos salários dos funcionários em relação à média. E, a seguir, calculemos esses desvios ao quadrado.

| Salário | Desvios | Desvios ao Quadrado |
|---|---|---|
| **10** | 10 – 9 = 1 | $1^2 = 1$ |
| **8** | 8 – 9 = –1 | $(-1)^2 = 1$ |
| **9** | 9 – 9 = 0 | $(0)^2 = 0$ |

Tomemos a variância como a média dos desvios ao quadrado, observando, porém, que, como se trata de uma amostra, devemos colocar N – 1 no denominador.

$$Var = \frac{1+1+0}{3-1} = \frac{2}{2} = 1$$

**Errado.**

---

**044.** (CESPE/ABIN/2018/OFICIAL DE INTELIGÊNCIA/ADAPTADA) A quantidade diária de *e-mails* indesejados recebidos por um atendente é uma variável aleatória *X*. Para estimar sua variância, retirou-se dessa distribuição uma amostra aleatória simples de tamanho quatro, cujos valores observados foram 10, 4, 2 e 4.

Com relação a essa situação hipotética, julgue o seguinte item.

A estimativa de máxima verossimilhança para a variância de *X*, que corresponde à variância amostral, é maior ou igual a 9.

---

Note que foi pedida a estimativa de máxima verossimilhança. Portanto, nessa situação, não cabe o fator de ajuste que é utilizado para o desvio-padrão amostral.

Vamos seguir o procedimento para o cálculo da variância:

Calculamos a média amostral:

$$\mu = \frac{10 + 4 + 2 + 4}{4} = \frac{20}{4} = 5$$

Calculemos os desvios em relação à média. E, a seguir, calculemos esses desvios ao quadrado.

| Salário | Desvios | Desvios ao Quadrado |
|---|---|---|
| **10** | 10 – 5 = 5 | $5^2 = 25$ |
| **4** | 4 – 5 = –1 | $(-1)^2 = 1$ |
| **2** | 2 – 5 = –3 | $(-3)^2 = 9$ |
| **4** | 4 – 5 = –1 | $(-1)^2 = 1$ |

Tomemos a variância como a média dos desvios ao quadrado.

$$Var = \frac{25 + 1 + 9 + 1}{4} = \frac{36}{4} = 9$$

Dessa forma, a variância é exatamente igual a 9.

**Certo.**

---

**045.** (FCC/TRT-11ª REGIÃO/AM E RR) O conjunto $\{X_1, X_2, X_3, \ldots, X_{10}\}$ refere-se a uma população de tamanho 10 de elementos estritamente positivos, em que

$$\sum_{i=1}^{10} X_i = 84, \quad \sum_{i=1}^{10} X_i^2 = 720 \quad \text{e} \quad \sum_{i=1}^{10} \log(X_i) = 9{,}185.$$

Observação: log(N) é o logaritmo de N na base 10. Considere as seguintes afirmações com relação a esta população:

I – O coeficiente de variação é igual a 1/7.

II – Multiplicando todos os elementos da população por 2, o coeficiente de variação da nova população formada não se altera.

III – Dividindo todos os elementos da população por 2, a variância da nova população formada é igual a 25% da variância anterior.

Está correto o que se afirma APENAS em:

a) I e III.

b) II e III.

c) I e II.

d) III.

e) I, II e III.

Vamos analisar as afirmações.

I – Para calcular o coeficiente de variação, devemos calcular a média e o desvio-padrão da variável.

A média pode ser obtida como a soma das variáveis dividida pelo total de observações.

$$\mu = \frac{\sum X_i}{10} = \frac{84}{10} = 8{,}4$$

Já a variância pode ser calculada como a diferença entre a esperança dos quadrados e o quadrado das esperanças.

$$E[X^2] = \frac{\sum X_i^2}{10} = \frac{720}{10} = 72$$

$$Var(X) = E[X^2] - \mu^2 = 72 - (8{,}4)^2 = 72 - 70{,}56 = 1{,}44$$

Como se trata de uma população, não há necessidade de se colocar um fator de ajuste. O desvio-padrão é a raiz quadrada da variância.

$$\sigma(X) = \sqrt{1{,}44} = \sqrt{\frac{144}{100}} = \frac{12}{10} = 1{,}2$$

Por fim, o coeficiente de variação é dado pela razão entre o desvio-padrão e a média.

$$CV = \frac{\sigma(X)}{\mu} = \frac{1{,}2}{8{,}4} = \frac{1}{7}$$

Afirmação correta.

II – O coeficiente de variação não se altera quando multiplicamos a variável por uma constante. Afirmação incorreta.

A razão para isso é que tanto o desvio-padrão como a variância são multiplicados pelo mesmo fator, portanto, a razão entre eles não se altera.

III – Quando dividimos a população por 2, a variância ficará dividida por $2^2$, ou seja, a nova variância será:

$$Var(X') = \frac{1}{2^2} \cdot Var(X) = \frac{1}{4} \cdot Var(X) = 0{,}25. Var(X) = 25\%\ de\ Var\ (X)$$

Afirmação correta.

**Letra a.**

# GABARITO

1. C
2. a
3. c
4. b
5. E
6. a
7. d
8. C
9. a
10. e
11. C
12. E
13. c
14. C
15. E
16. E
17. C
18. c
19. c
20. E
21. b
22. C
23. a
24. c
25. E
26. C
27. E
28. a
29. c
30. b
31. C
32. E
33. c
34. a
35. a
36. c
37. c
38. b
39. b
40. c
41. a
42. C
43. E
44. C
45. a

**Thiago Cardoso**

Engenheiro eletrônico formado pelo ITA com distinção em Matemática, analista-chefe da Múltiplos Investimentos, especialista em mercado de ações. Professor desde os 19 anos e, atualmente, leciona todos os ramos da Matemática para concursos públicos.


LEI Nº 8.666/1993 - LICITAÇÃO
Avaliação
Comentário
Seu feedback é valioso. Você gostaria de deixar um comentário e assim nos ajudar a melhorar nossos produtos e serviços?
Obs: A avaliação da aula em pdf é exclusivamente pedagógica. Clique aqui para relatar problemas técnicos, pois serão desconsiderados deste canal.
Sim, salvar comentário.
Não, obrigado.