PROBABILIDADE E ESTATÍSTICA

Medidas de Posição

Livro Eletrônico

SUMÁRIO

MEDIDAS DE POSIÇÃO

1. Introdução à Estatística

1.1. Conceitos Básicos

A Estatística é um ramo da matemática que estuda a coleta, a organização e a análise de dados por amostras.

Alguns conceitos fundamentais que precisamos saber antes de nos aprofundarmos em estatística são:

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

Por 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.

Por 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.

- **Unidade amostral:** é cada elemento pertencente à amostra.

Nesse caso, cada pessoa cuja altura foi medida é uma unidade amostral.

1.2. Ramos da Estatística

Existem dois ramos importantes da estatística, conhecidos como Estatística Descritiva e Estatística Inferencial.

A Estatística Descritiva tem por objetivo extrair uma amostra da população por meio das técnicas de **amostragem.**

A amostragem visa criar uma amostra que seja representativa de uma população de interesse de estudo.

Outras ferramentas importantes da Estatística Descritiva são:

- a visualização dos dados, que visa organizá-los e apresentá-los, de modo que sua compreensão seja facilitada. Para isso, costuma-se recorrer a gráficos e tabelas estatísticos.

- A análise exploratória de dados, que visa resumir suas características principais, frequentemente com métodos visuais.

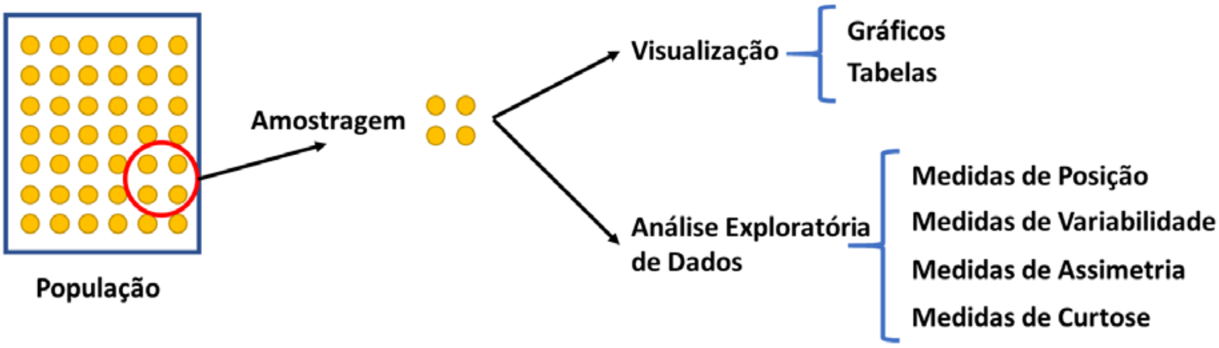


Figura 1: Fluxo da Estatística Descritiva

A análise exploratória de dados utiliza várias técnicas gráficas e métodos matemáticos. Porém, é importante deixar claro que os desafios nesse resumo dos dados são muito grandes. E é preciso ter muito cuidado com as técnicas empregadas, de modo a evitar que conclusões erradas sejam tomadas com base em uma manipulação inapropriada dos dados.

É importante também observar que a análise exploratória é feita com dados da própria amostra. Por isso, a etapa de amostragem é também essencial para o sucesso da análise exploratória. Toda a estatística trabalha, portanto, em conjunto com o objetivo de simplificar o estudo e a avaliação de dados.

A Estatística Inferencial, por sua vez, tem o objetivo contrário: ela busca utilizar uma amostra representativa de uma população para inferir sobre as características da população inteira.

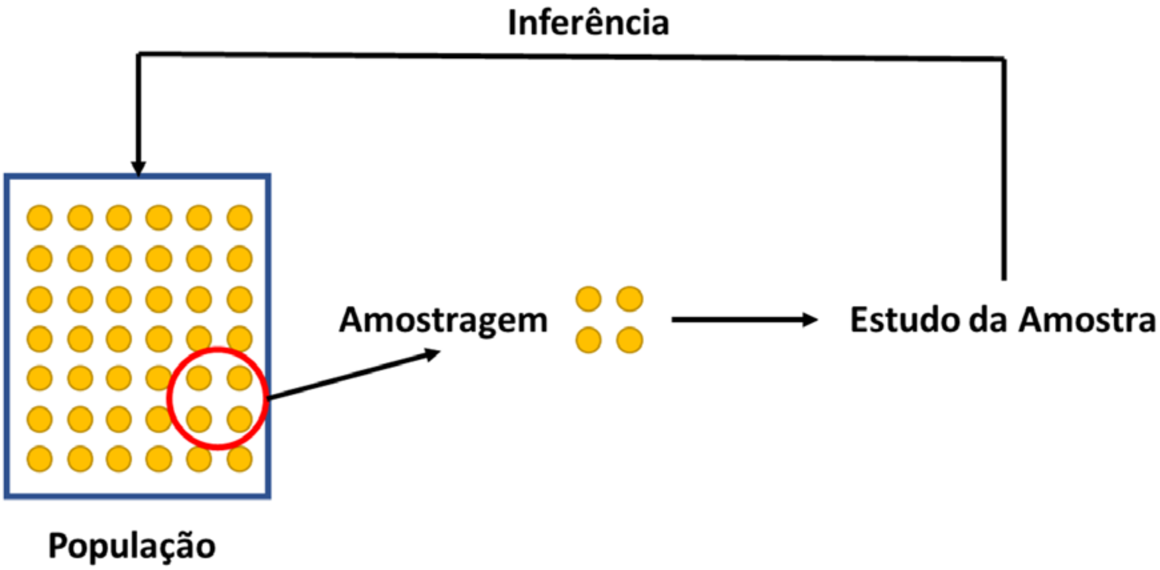


Figura 2: Fluxo da Estatística Inferencial

É natural que obter os valores reais dos parâmetros de uma população a partir de uma amostra introduzirá erros. Como forma de limitar esses erros, a Estatística Inferencial tem a importante ferramenta dos **intervalos de confiança.**

O intervalo de confiança visa estimar uma faixa segura de estimativa de um parâmetro populacional. É o que ouvimos, por exemplo, nas pesquisas eleitorais na televisão: o candidato X tem 54% das intenções de voto com margem de erro de 2% para mais ou para menos, com grau de confiança de 95%.

A margem de erro é, na verdade, a largura do intervalo de confiança, que, nesse caso, está entre 52% e 56%, com o grau de confiança de 95%.

Para o estudo dos intervalos de confiança, precisamos conhecer as principais distribuições de probabilidade, tendo em vista que ele é probabilístico.

Outro problema muito importante são **os testes de hipóteses.** Nessa situação, deseja-se saber se uma observação não contradiz uma hipótese inicial sobre a população.

Pense, por exemplo, que você é um(a) pesquisador(a) e há uma tese de que a altura média do homem brasileiro é 1,73 m. Porém, suponha que lhe foi apresentado um conjunto de 100 pessoas, cuja altura média é igual a 1,95 m. Será que essa observação contradiz a hipótese de que a altura média do brasileiro é 1,73 m? Ou essa amostra pode ter sido realmente produzida pelo mero acaso?

Esse é o grande problema dos testes de hipóteses, que são essenciais para muitas ciências, como a medicina. São muito utilizados para determinar se um determinado medicamento realmente é útil no combate a uma doença ou se pode causar efeitos colaterais.

1.3. Variáveis Estatísticas

As variáveis estatísticas são características dos elementos da amostra que nos interessa averiguar.

São classificadas em dois grupos principais:

- **variáveis qualitativas:** não podem ser mensuradas numericamente e se baseiam em qualidades.

Ex.: cor dos olhos de um indivíduo (verde, azul, castanho, preto).

- **Variáveis quantitativas:** são aquelas que podem ser mensuradas numericamente.

Ex.: comprimento de uma fita (pode ser 1,50 m ou 2,50 m).

Um caso especial que pode trazer confusão aos alunos são variáveis que, embora sejam expressas por números, estes **não** representam medidas.

Pense, por exemplo, no número de uma casa. A sua casa tem um número, por exemplo, 71. Mas esse número não representa nenhuma medida de nenhuma propriedade da sua casa. É apenas uma forma de identificação. Portanto, o número da casa é uma grandeza qualitativa.

Quando o número que caracteriza uma grandeza for somente **uma identificação**, essa grandeza é **qualitativa.**

As variáveis qualitativas podem ser classificadas em:

- **ordinais:** podem ser ordenadas.

Ex.: classe social (A, B, C, D ou E); grau de instrução (fundamental, médio ou superior).

- **Nominativas:** não podem ser hierarquizadas ou ordenadas, não admitindo, portanto, nenhuma comparação do tipo menor ou maior. Podem ser comparadas apenas por igualdade ou diferença.

Ex.: cor dos olhos (verde, azul, castanho, preto); sexo (masculino ou feminino).

Por sua vez, as variáveis quantitativas podem ser classificadas em:

- **discretas:** o conjunto de possíveis resultados é enumerável.

Ex.: número de filhos (0, 1, 2, 3 etc.); número de alunos em uma sala de aula (0, 1, 2, 3 etc.).

- **Contínuas:** podem assumir qualquer número real, inclusive, números fracionais e irracionais.

Ex.: nível de açúcar no sangue, altura.

Um conjunto enumerável pode ser entendido como um conjunto que pode ser associado aos números naturais. Dessa forma, existe o primeiro, o segundo, o terceiro etc.

Uma variável discreta só pode assumir um conjunto enumerável de valores. É o caso do número de filhos de uma pessoa. Uma pessoa pode ter 0 filho, 1 filho, 2 filhos, mas não pode ter 2½ filhos (dois filhos e meio).

Por outro lado, uma variável contínua pode assumir qualquer valor em um intervalo de números reais.

Por exemplo, a altura de uma pessoa pode ser 2 m, mas também pode ser 1,57 m ou 1,837 m. E assim por diante.

Dessa maneira, vamos sintetizar as quatro categorias de variáveis estatísticas.

Figura 3: Categorias de Variáveis Estatísticas

É importante destacar que **a classificação de uma mesma variável pode mudar**, dependendo do contexto da pesquisa.

Por exemplo, a idade de uma pessoa é, a princípio, uma variável contínua. Uma pessoa pode ter 29 anos, 4 meses, 11 dias, 18 horas, 14 minutos e 53 segundos.

Porém, no contexto de uma pesquisa, é bastante comum que o pesquisador queira anotar a idade da pessoa somente em anos. Inclusive, na maioria das vezes em que somos questionados a respeito da nossa idade, dizemos simplesmente "29 anos". Dessa forma, a idade se transformou em uma variável.

Imagine também que você está fazendo uma pesquisa sobre grau de instrução de tenentes e capitães em um quartel militar. Todos eles são formados na mesma instituição, por isso não faz sentido avaliar o grau de instrução entre superior, médio e fundamental.
Porém, pode fazer sentido avaliar o grau de instrução de um tenente pelo número de cursos diferentes que ele já participou ou pela quantidade de horas-aula em instrução após a formatura. Nesse caso, o grau de instrução do tenente se transformaria em uma variável quantitativa.

Outro conceito que precisamos conhecer é o de **observação de uma variável estatística.** A observação consiste no valor assumido daquela variável para cada elemento da amostra.

Por exemplo, quando forem medidas as alturas de pessoas e obtiverem-se os valores: 2 m, 1,57 m e 1,84 m. Cada uma dessas medidas corresponde a uma observação diferente da variável altura.

1.4. Principais Formas de Apresentação dos Dados

Por enquanto, as três formas básicas que precisamos conhecer para a apresentação dos dados estatísticos são:

- **rol:** quando os dados são listados um a um.

Por exemplo, considere que X é a idade dos funcionários em uma empresa. Se a empresa tem poucos funcionários, pode ser conveniente simplesmente listar a idade de todos eles. Por exemplo:
{28, 27, 28, 36, 34, 51, 43, 29, 32, 38}

- **Distribuição de frequências tabular:** é utilizada para variáveis discretas com muitas observações.

Pense, por exemplo, que uma empresa deseja estudar o número de filhos de seus funcionários e que, em vez de 10 funcionários, como no exemplo anterior, a empresa tivesse 80. Nesse caso, não seria muito adequado simplesmente listá-los na forma de rol, pois seria um rol muito longo. No entanto, uma visualização em forma de tabela ou gráfico seria bem mais interessante.

Número de filhos	Quantidade de casais (frequência absoluta)
0	17
1	33
2	21
3	7
4	2
5	0

O número de casais é a frequência absoluta dos valores possíveis para a variável. Ele significa que 17 casais têm 0 filho e 33 casais têm 1 filho.

Esse número poderia ser expresso também em porcentagem em relação ao total de 80 casais. 17 casais no total de 80 representam um percentual de 21,25%. Para fazer essa conta, basta dividir.

$$f_0 = \frac{17}{80} = 0{,}2125 = 21{,}25\%$$
$$f_1 = \frac{33}{80} = 0{,}4125 = 41{,}25\%$$
$$f_2 = \frac{21}{80} = 0{,}2625 = 26{,}25\%$$
$$f_3 = \frac{7}{80} = 0{,}0875 = 8{,}75\%$$
$$f_4 = \frac{2}{80} = 0{,}025 = 2{,}5\%$$

Número de filhos	Frequência absoluta	Frequência relativa	Frequência acumulada
0	17	21,25%	21,25%
1	33	41,25%	62,5%
2	21	26,25%	88,75%
3	7	8,75%	97,5%
4	2	2,5%	100%
5	0	0%	100%

Por fim, também podemos expressar em termos da **frequência acumulada**, que é a soma das frequências relativas dos valores menores ou iguais ao valor em questão. Para o último elemento da amostra, a frequência acumulada será sempre igual a 100%

Por exemplo, a frequência acumulada até 1 filho seria a soma das frequências relativas de 0 filho e de 1 filho.

$$F_1 = f_0 + f_1 = 21{,}25\% + 41{,}25\% = 62{,}5\%$$

$$F_2 = f_0 + f_1 + f_2 = 62{,}5\% + 26{,}25\% = 88{,}75\%$$

$$F_3 = f_0 + f_1 + f_2 + f_3 = 88{,}75\% + 8{,}75\% = 97{,}5\%$$

$$F_4 = f_0 + f_1 + f_2 + f_3 + f_4 = 97{,}5\% + 2{,}5\% = 100\%$$

- **Distribuição de frequências categorizada:** acontece com variáveis contínuas ou mesmo com variáveis discretas que podem apresentar muitos valores. Nesse caso, criam-se categorias que abrangem uma faixa de valores, e não um valor único. É o caso da faixa salarial. Vejamos um exemplo.

Faixa salarial	Número de empregados
1 a 3	50
3 a 5	32
5 a 7	12
7 a 9	6

1.5. Medidas de Posição e de Variabilidade

As medidas de posição e variabilidade são grandezas muito úteis no estudo de uma variável estatística.

A medida de posição é um número representativo de uma amostra.

Uma medida de posição é um número que, sozinho, representa essa amostra. Dentre elas, destacam-se **as medidas de tendência central**, que são grandezas que visam estimar um ponto central na amostra. As grandezas mais conhecidas desse tipo são: a média aritmética, a moda e a mediana.

Por exemplo, considere o PIB de três países. O PIB anual é essencialmente a soma de todos os bens e serviços finais produzidos em um determinado país, em um determinado ano.

País	PIB
Brasil	US$ 2 trilhões
Grécia	US$ 222 bilhões
Catar	US$ 201 bilhões

Note que esse número bruto é pouco informativo. No entanto, podemos melhorar significativamente a informação se considerarmos que a população do Brasil é de 209,5 milhões de pessoas, a população da Grécia é de 10,73 milhões e a do Catar é de 2,78 milhões.

Com base nesses dados, podemos calcular o PIB per capita, que nada mais é do que uma **média aritmética.** Consiste em todo o produto interno bruto da nação dividido pelo total de pessoas habitantes naquele território.

País	PIB	População	PIB per capita
Brasil	US$ 2 trilhões	209,5 milhões	US$ 9.651
Grécia	US$ 222 bilhões	10,73 milhões	US$ 20.689
Catar	US$ 201 bilhões	2,78 milhões	US$ 72.302

Notamos que, embora, o Brasil tenha um PIB maior, a riqueza total é baixa em relação ao total de sua população, o que significa que o indivíduo médio brasileiro tem uma renda menor que o indivíduo médio que mora na Grécia ou no Catar.

Mas também existem outras medidas de posição, conhecidas como **separatrizes**, que não buscam um ponto central.

Por exemplo, vimos que o PIB per capita da Grécia é de US$ 20.689, que é muito inferior ao PIB per capita do Catar.
Porém, imagine que os 10% mais pobres da Grécia tenham uma renda média anual de US$ 7.500, enquanto os 10% mais pobres do Catar tenham renda média anual de US$ 4.800.

Olha só que interessante. A média não necessariamente reflete uma porção específica dos indivíduos de que se deseja falar numa estatística. Se estamos fazendo um estudo sobre as condições de vida da população mais pobre de um determinado país, não faz sentido tomar simplesmente a média, que é uma medida de tendência central. É muito mais útil tomar uma medida de posição que realmente se desloque para mais próximo do grupo populacional que se deseja estudar.

Já as medidas de variabilidade visam estudar a **uniformidade** de uma variável estatística em uma população ou amostra.

Por exemplo, o índice de Gini avalia a desigualdade de renda em um país. Esse índice varia de 0%, que corresponde à perfeita igualdade, em que todas as pessoas possuem exatamente a mesma renda, a 100%, que corresponde à total desigualdade, em que toda a renda é concentrada nas mãos de um único indivíduo.

Por essa métrica, entre os países mais igualitários do mundo, encontram-se: República Checa (24,9%), Bélgica (27,4%), Azerbaijão (26,6%) e Argélia (27,6%). Por outro lado, entre os países mais desiguais, figuram: Brasil (53,9%), Chile (44,4%) e Estados Unidos (41,4%).

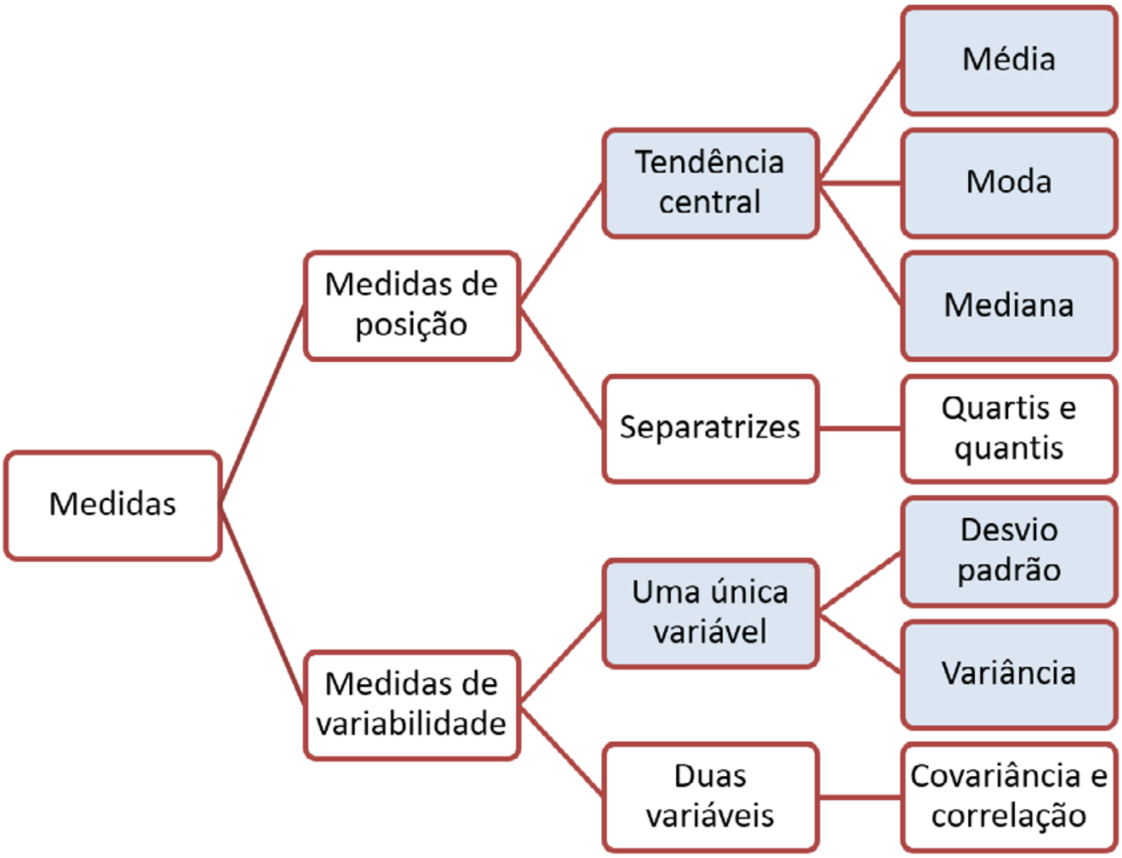
Observe, portanto, que um índice de Gini mais baixo (menor desigualdade) não significa necessariamente uma informação sobre as medidas de posição da mesma variável estatística.

Por exemplo, a Argélia e a Bélgica têm índices de desigualdade muito próximos, porém, há uma gigantesca diferença de PIB per capita entre elas (US$ 5.886 da Argélia contra US$ 41.740 da Bélgica).

Embora sejam bastante desiguais, o PIB per capita dos Estados Unidos (US$ 59.501) é maior também que o da Bélgica. Já a Suíça é mais igualitária (Gini 32,9%) e mais rica que os Estados Unidos (PIB per capita US$ 82.950).

Portanto, **não existe nenhuma relação direta** entre as medidas de posição e as medidas de variabilidade. Elas são **informações complementares.**

Vamos esquematizar as principais medidas que estudaremos ao longo do nosso curso.



Destacamos com fundo escuro as medidas que estudaremos neste capítulo, que são as mais importantes no estudo da estatística.

Direto do Concurso

001. (CESPE/SEDUC-AM/2011/ESTATÍSTICO)

RG(*)	gênero	grau de instrução	hiperatividade
35805684	M	3	N
21355706	F	5	N
43674475	M	2	S
2305386	F	3	N
⋮	⋮	⋮	⋮
98814652	M	4	N

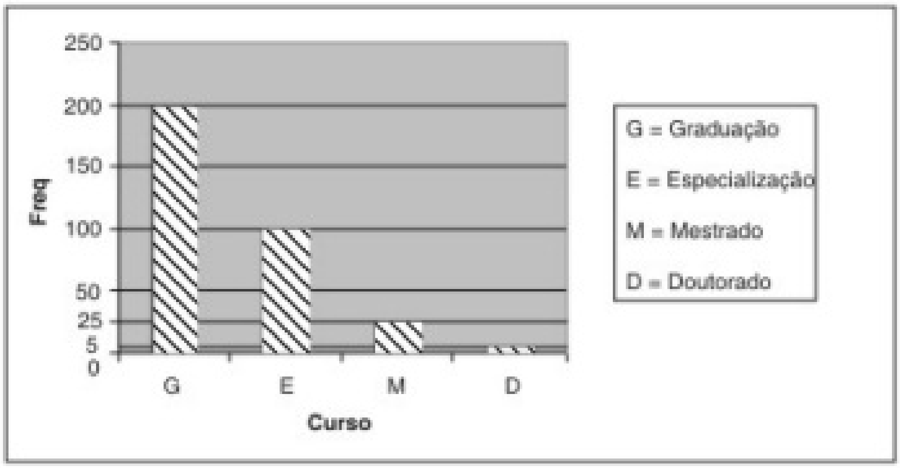
(*) números fictícios gerados por computador.

A tabela acima contém um conjunto de dados formado por quatro variáveis: RG; gênero (M = masculino; F = feminino); grau de instrução (1 = analfabeto; 2 = fundamental incompleto; 3 = fundamental completo; 4 = médio incompleto; 5 = médio completo ou superior); e hiperatividade (S = sim; N = não). Com base nessa tabela, julgue o item.
As variáveis mostradas na tabela são qualitativas.

Uma variável qualitativa é aquela que não pode ser expressa por uma medida, um valor numérico. É o caso de todas as variáveis em apreço.

Note que, mesmo no caso do grau de instrução, na verdade, ele pode ser especificado em: analfabeto, fundamental incompleto, fundamental completo, médio incompleto e médio completo ou superior. Não são, portanto, valores numéricos. Logo, essa variável é realmente qualitativa.

Certo.



002. (CESPE/SEDUC-AM/2011/ESTATÍSTICO) A qualificação dos professores é de grande importância para a qualidade da formação dos estudantes. Considerando que a figura acima apresenta a distribuição do número de professores em uma faculdade, segundo a formação acadêmica (curso), julgue o item.

A variável curso é qualitativa nominal.

Uma variável qualitativa é aquela que não pode ser expressa por uma medida, um valor numérico. É o caso da variável curso.

Porém, note que existe um grau de ordenação entre os possíveis valores dessa variável aleatória. Podemos dizer que o nível de doutorado é superior ao nível de mestrado, que é superior à especialização, que é superior à graduação.

Dessa forma, a variável não é nominal, mas sim ordinativa.

Errado.

003. (UFU-MG/2019/TÉCNICO EM ESTATÍSTICA) Considere as seguintes variáveis.

I – Tamanho de um objeto (pequeno, médio ou grande)

II – Volume de água em um rio

III – Número de clientes numa fila

IV – Número da seção de votação

V – Comprimento de um inseto

VI – Classe Social

Com relação à classificação dos dados requeridos como variáveis de pesquisa, é correto afirmar que:

a) as variáveis I, IV e VI são qualitativas.

b) as variáveis III e V são quantitativas contínuas.

c) as variáveis II e III são quantitativas discretas.

d) a variável IV é qualitativa ordinal.

Vamos analisar as afirmações.

I – Da forma como foi descrito, o tamanho não foi medido, apenas categorizado em pequeno, médio ou grande. Trata-se, portanto, de uma variável qualitativa. Ela também é ordinal, porque existe uma hierarquia entre os tamanhos.

II – O volume de um rio pode ser medido. Portanto, é uma variável quantitativa. É também uma variável contínua, porque o volume pode assumir qualquer valor, inteiro ou não. É possível que o rio tenha 14,5 m³ ou 18,33 m³.

III – O número de clientes em uma fila pode ser contado: 1, 2, 3, 4. Não é possível que exista uma fila com 2,4 pessoas. Portanto, é uma variável quantitativa discreta.

IV – Embora o número da seção de votação seja um número, ele não representa uma medida, apenas uma identificação. Quando dizemos que a seção é a número 235, isso não quer dizer nada sobre a seção específica. Portanto, trata-se de uma variável qualitativa nominativa.

V – O comprimento de um inseto pode ser medido e pode assumir qualquer valor, inteiro ou não. Portanto, é uma variável quantitativa contínua.

VI – A classe social de uma pessoa não pode ser expressa numericamente. Por exemplo, A, B, C, D ou E. Porém, existe uma ordenação entre esses valores. Trata-se, portanto, de uma variável quantitativa ordinal.

Com base nisso, vamos analisar as afirmações.

a) É isso mesmo. Afirmação correta.

b) O número de clientes em uma fila é uma variável discreta, e não contínua. Afirmação incorreta.

c) O volume de água no rio é uma variável contínua. Afirmação incorreta.

d) Essa é uma variável qualitativa nominativa. Afirmação incorreta.

Letra a.

---

2. Média Aritmética

A média aritmética é uma medida à qual já estamos bastante acostumados. É utilizada para calcular a sua média na escola e na faculdade.

Pense, por exemplo, que você faz uma faculdade, em que você precisa de média 6,5 para ser aprovado(a) em uma matéria, que tem duas provas bimestrais e um exame.
Na primeira prova, você se deu bem e tirou 8,5. Na segunda prova, você estudou um pouco menos e tirou 7,5. Crente de que já estava com boas notas, você estudou menos ainda para o exame e tirou 5,0. Qual a sua média final?

$$M = \frac{8{,}5 + 7{,}5 + 5}{3} = \frac{21}{3} = 7{,}0$$

Então, você conseguiu! Tirou média 7,0 e conseguiu ser aprovado(a).

Agora, tome a sua nota como uma variável estatística. Como você fez três provas, há três observações dessa variável.
A média aritmética foi calculada como a soma de todas as observações de uma variável dividida pela quantidade total de variáveis.

Vejamos um outro exemplo. Considere que uma empresa tenha 10.000 funcionários e deseja saber a idade média deles. Para isso, ela fez uma pesquisa com uma amostra de 10 funcionários, registrando os seguintes dados:

Tabela 1: conjunto de dados sobre a idade dos funcionários da Empresa X

Pessoa	Idade	Pessoa	Idade
1	28	**6**	51
2	27	**7**	43
3	28	**8**	29
4	36	**9**	32
5	34	**10**	38

Nesse caso, a média pode ser também calculada como a soma de todos os valores dividida pelo total de observações.

$$\bar{X} = \frac{28 + 27 + 28 + 36 + 34 + 51 + 43 + 29 + 32 + 38}{10} = \frac{346}{10} = 34{,}6$$

Portanto, a média aritmética é 34,6 anos. Uma representação a que devemos nos acostumar é $\bar{X}$ (ou X barra). Essa é uma representação muito comum para a média amostral. Geralmente, representamos:

- $\bar{X}$**:** média amostral
- μ**:** média populacional

Note a diferença entre os dois conceitos: para obtermos a média populacional, precisaríamos levar em consideração todos os elementos da população. No caso em apreço, deveríamos **ter registrado as médias de todos os funcionários da empresa.**

Como isso não aconteceu, estamos falando da média amostral.

É possível também representar a média aritmética de uma forma sintética.

$$\bar{X} = \frac{x_1 + x_2 + \cdots + x_N}{N} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Você deve ter achado estranho o sigma gigante que colocamos na igualdade ao lado direito. Porém, é importante que você se acostume com esse tipo de notação.

O sigma representa um somatório de forma mais compacta. As notações i=1 embaixo e N em cima devem ser lidas como "o somatório dos x_i em que i vai de 1 a N."

Sendo assim, temos que:

$$\sum_{i=1}^{N} x_i = x_1 + x_2 + \cdots + x_N$$

É importante você se acostumar com essa notação de somatório, pois algumas bancas gostam de colocar o sigma em provas de estatística, então você precisa estar preparado(a) para lidar com isso, ok?

Vejamos um exemplo mais simples envolvendo o operador sigma.

$$\sum_{i=1}^{4} i^2$$

Nesse caso, temos o somatório dos i^2, em que i vai de 1 a 4. Esse somatório será, portanto, igual a:

$$\sum_{i=1}^{4} i^2 = 1^2 + 2^2 + 3^2 + 4^2 = 1 + 4 + 9 + 16 = 30$$

Não tenha medo dos somatórios. As questões que os envolvem costumam ser bastante diretas.

Agora, calculemos a média aritmética das idades dos alunos da amostra selecionada.

{28, 27, 28, 36, 34, 51, 43, 29, 32, 38}

Como são 10 elementos na amostra, temos que a média aritmética é simplesmente a soma de todos os elementos dividida por 10.

$$\bar{x} = \frac{28 + 27 + 28 + 36 + 34 + 51 + 43 + 29 + 32 + 38}{10} = \frac{346}{10} = 34{,}6$$

Portanto, a média aritmética é 34,6 anos.

A média aritmética é a medida de posição mais importante e, por isso, também tem vários nomes. Pode ser conhecida também como esperança ou valor esperado.

Uma forma simples de representá-la é com o somatório de todos os elementos da amostra dividido pela quantidade de termos.

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

Direto do Concurso

004. (CESPE/TJ-PA/2020/ANALISTA JUDICIÁRIO – ESTATÍSTICA) Uma amostra aleatória dos registros de furto no município de Abaetetuba, no ano de 2017, apresenta os valores 245, 247, 238, 282 e 261. Uma estimativa não tendenciosa e eficiente para a média de furtos ocorridos em Abaetetuba no ano de 2017, considerando os dados apresentados na amostra, é

a) 238,0.
b) 254,6.
c) 260,0.
d) 282,7.
e) 308,5.

A média pode ser calculada pela soma de todos os valores dividida pelo número de elementos presentes na amostra.

$$\mu = \frac{245 + 247 + 238 + 282 + 261}{5} = \frac{1273}{5} = 254{,}6$$

Letra b.

---

005. (FCC/CÂMARA DE FORTALEZA – CE/2019/AGENTE ADMINISTRATIVO) Em um teatro com 200 lugares, houve quatro apresentações de uma peça. Na primeira apresentação foram vendidos todos os ingressos; na segunda apresentação foram vendidos 88% dos ingressos;

na terceira, 56% dos ingressos e, na quarta, 44% dos ingressos. Em média, a quantidade de ingressos vendidos por apresentação foi de

a) 72
b) 144
c) 56
d) 76
e) 140

A média do percentual de ingressos vendidos pode ser obtida pela definição, que é a soma dos percentuais dividida pelo total de observações, que é igual a 4.

$$\mu = \frac{100\% + 88\% + 56\% + 44\%}{4} = \frac{288\%}{4} = 72\%$$

Agora, podemos calcular a média dos ingressos vendidos:

$$N = 0{,}72.200 = 144$$

Letra b.

Um artifício muito útil nas questões sobre média é notar que o produto da média aritmética pela quantidade de termos presentes na amostra é igual ao somatório de todos os termos.

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

Podemos representar, simbolicamente o somatório por S.

$$S = \mu.N$$

Direto do Concurso

006. (FCC/SABESP/2018/ESTAGIÁRIO DE NÍVEL MÉDIO) Para que a média aritmética dos números: 8, 8, 1, 10, 11, 12, 7, 2, 10, 6, **x** e 5 seja 7, o valor de **x** deverá ser:

a) 2
b) 4
c) 8
d) 5
e) 3

O produto da média aritmética pela quantidade de termos é igual à soma de todos os elementos. Observe que são 12 elementos na amostra.

$$S = \mu . N$$

$$S = 7.12 = 84$$

Agora, vamos calcular o valor de **x** dentro da soma dos termos.

$$S = 8 + 8 + 1 + 10 + 11 + 12 + 7 + 2 + 10 + 6 + x + 5 = 84$$

$$80 + x = 84$$

$$\therefore x = 84 - 80 = 4$$

Letra b.

---

007. (CESPE/PREFEITURA DE SÃO CRISTÓVÃO – SE/PROFESSOR DE MATEMÁTICA) Segundo o IBGE, a massa da renda média mensal real domiciliar per capita em 2016 foi de aproximadamente R$ 264 bilhões; a população brasileira nesse ano era de aproximadamente 190 milhões de pessoas.
A partir dessas informações, julgue o item a seguir.
A renda média mensal dos brasileiros em 2016 foi superior a R$ 1.300.

A média aritmética pode ser obtida como a razão entre a soma das rendas mensais e a população brasileira total. Nesse contexto, a soma das rendas foi chamada de massa.
Para facilitar as contas, podemos escrever R$ 264 bilhões como R$ 264 000 milhões.

$$\mu = \frac{S}{N} = \frac{264000}{190} \cong 1389{,}47 > 1300$$

Certo.

2.1. Média Ponderada

Você também já conviveu com a média ponderada quando estudou na faculdade. Imagine que a sua nota bimestral seja dada por:

- uma prova com peso igual a 3;
- uma prática de laboratório com peso igual a 1.

Suponha que você tirou 6 na prova e 8 na prática de laboratório. Qual seria a sua nota final?

Para isso, devemos considerar os pesos das duas atividades. Como a prova tem peso 3, devemos multiplicar a nota da prova por 3. No denominador, devemos considerar a soma dos pesos.

$$M = \frac{6.3 + 8}{3 + 1} = \frac{18 + 8}{4} = \frac{26}{4} = 6{,}5$$

Observe que o peso maior fez que a média se aproximasse mais da sua nota da prova do que da nota da prática de laboratório. Ou seja, a prova é mais importante na determinação de sua média.

Em notação matemática de somatório, podemos escrever:

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

Na estatística, a média ponderada é muito útil quando os dados são fornecidos na forma de **distribuição de frequências** – sejam frequências absolutas, sejam frequências relativas (valor percentual). Vejamos um exemplo.

Foi feita uma pesquisa sobre o número de filhos que várias pessoas casadas possuíam e chegou-se ao seguinte resultado:

Figura 4: Gráfico sobre o Número de Filhos de Diversos Casais

O histograma da Figura 2 é bastante simples de entender. Dentre as pessoas pesquisadas, 4 pessoas não possuem filhos, 5 pessoas possuem 1 filho, 7 pessoas possuem 2 filhos, 3 pessoas possuem 3 filhos, nenhuma pessoa possui nenhum filho e 1 pessoa possui 5 filhos.

Quando se tem um gráfico desse gênero, para se calcular o número de filhos médio das pessoas entrevistadas, **deve-se utilizar as frequências** de cada uma das classes **como peso para a média**. Em outras palavras:

$$\bar{x} = \frac{\sum_{i=1}^{N} f_i x_i}{\sum_{i=1}^{N} f_i} = \frac{4.0 + 5.1 + 7.2 + 3.3 + 0.4 + 1.5}{4 + 5 + 7 + 3 + 0 + 1} = \frac{33}{20} = 1{,}65$$

É muito comum também esse tipo de informação ser fornecida na forma de uma tabela estatística. Vejamos outro exemplo com o número de filhos de um conjunto de casais de uma pesquisa.

Tabela 2: número de filhos por casal

Número de filhos	Quantidade de casais
0	17
1	33
2	21
3	7
4	2
5	0

Nesse caso, o número médio de filhos dos casais é obtido também pela média ponderada, em que a quantidade de casais serve como **peso.**

$$\bar{X} = \frac{17.0 + 33.1 + 21.2 + 7.3 + 2.4 + 0.5}{17 + 33 + 21 + 7 + 2 + 0} = \frac{0 + 33 + 42 + 21 + 8 + 0}{80}$$

$$\bar{X} = \frac{104}{80} = 1{,}3$$

Portanto, o número médio de filhos da amostra tratada na Tabela 2 é igual a 1,3 filhos.

Um ponto interessante é que o **número de filhos é uma variável discreta**. Isso significa que um casal pode ter 0, 1 ou 2 filhos, mas não é possível que o casal tenha 1,5 filho ou 2,62 filhos.

Assim, o número de filhos só pode ser um número inteiro. Mas, a média não necessariamente é um número inteiro, como vimos em nossos exemplos.

Direto do Concurso

008. (FCC/PREFEITURA DE TERESINA – PI) Uma carteira aplica 25% na ação A, 40% na ação B e o restante na ação C. Os retornos das ações A, B e C são, respectivamente, 10%, 12% e 20%. O retorno médio da carteira será

a) 14,5%.
b) 14,8%.
c) 14,6%.

d) 14,0%.
e) 14,3%.

Devemos calcular os retornos pela média aritmética, em que a composição das carteiras são os pesos.
Devemos notar que o teor da ação C na carteira pode ser obtido, impondo que a soma de todas as participações das ações é igual a 100%.

$$P_A + P_B + P_C = 1$$

$$0{,}25 + 0{,}40 + P_C = 1$$

$$\therefore P_C = 1 - 0{,}25 - 0{,}40 = 1 - 0{,}65$$

$$P_C = 0{,}35$$

Agora, vamos calcular a média usando os teores de cada ação na carteira como pesos.

$$\mu = \frac{0{,}25.10\% + 0{,}40.12\% + 0{,}35.20\%}{0{,}25 + 0{,}40 + 0{,}35} = \frac{2{,}5\% + 4{,}8\% + 7\%}{1} = 14{,}3\%$$

Letra e.

Agora, veremos o que aconteceria, se estivéssemos lidando com uma variável contínua.

2.1.1. Dados Categorizados

No caso de uma variável contínua, o gráfico de distribuição de frequências é construído em torno de **categorias.**

Uma categoria inclui um conjunto de valores possíveis. É o que acontece, por exemplo, com o salário, que é normalmente representado por meio de faixas salariais.

Suponha que foi feita uma estatística dos empregados de uma determinada empresa por faixa salarial (em função do salário mínimo).

Tabela 3: dados sobre os salários dos funcionários de uma empresa

Faixa salarial	Número de empregados
1 a 3	50
3 a 5	32
5 a 7	12
7 a 9	6

O salário é uma variável contínua. Observe que uma pessoa pode ganhar 1,5 salário mínimo ou 2,4 salários mínimos. Não há qualquer impedimento para isso.

Nesse caso, a técnica mais utilizada para a média é **a técnica do ponto médio**, que consiste em supor que todas as observações de uma determinada categoria **podem ser colocadas no ponto médio da respectiva classe.**

O ponto médio de uma classe corresponde à média aritmética dos seus extremos.

Tabela 4: dados sobre os salários dos funcionários de uma empresa

Faixa salarial	Ponto médio	Número de empregados
1 a 3	**2**	50
3 a 5	**4**	32
5 a 7	**6**	12
7 a 9	**8**	6

Agora, podemos calcular a média para o conjunto de dados categorizados fornecidos.

$$\bar{X} = \frac{50.2 + 32.4 + 12.6 + 6.8}{50 + 32 + 12 + 6} = \frac{100 + 128 + 72 + 48}{100} = \frac{348}{100} = 3{,}48$$

Essa técnica é bem interessante e muito recorrente em questões de prova. Fique de olho.

Direto do Concurso

009. (FCC/SPAG-PE/2019/ANALISTA DE PLANEJAMENTO/ADAPTADA) Durante 40 dias, foi registrado o número de pessoas atendidas por dia em um guichê de uma repartição. A tabela abaixo apresentou os dados observados sendo que não foram fornecidas as quantidades de dias em que foram atendidas uma pessoa por dia e duas pessoas por dia, indicadas na tabela por $q_1 = 11$ e $q_2 = 14$, respectivamente.

Número de pessoas atendidas	Quantidade de dias
0	9
1	q_1
2	q_2
3	5
4	1
Total	40

Tem-se que a média aritmética do número de pessoas atendidas por dia é:

a) 1,40
b) 1,50
c) 1,25
d) 1,60
e) 1,45

Devemos tomar como pesos para a média a quantidade de dias. Então, devemos multiplicar a quantidade de dias pelo número de pessoas atendidas e somar. No denominador, colocamos a soma dos pesos.

$$\mu = \frac{9.0 + 11.1 + 14.2 + 5.3 + 1.4}{40} = \frac{58}{40} = 1{,}45$$

Letra e.

010. (FCC/TCE-PR/2011/ANALISTA DE CONTROLE) A média dos salários, calculada supondo-se que todos os valores dentro de uma faixa salarial tenham seus valores iguais ao ponto médio desta faixa, em número de salários mínimos, é igual a:

Faixa salarial (em número de salários mínimos)	Frequência Absoluta
1 |— 3	200
3 |— 5	400
5 |— 7	200
7 |— 9	200

a) 4,2
b) 4,5

c) 4,6
d) 4,8
e) 5,0

Como o próprio enunciado já deu a dica, devemos considerar o ponto médio de cada faixa salarial para o cálculo da média.

Classe de salários	Ponto médio	Frequência absoluta
1 a 3	2	200
3 a 5	4	400
5 a 7	6	200
7 a 9	8	200

Calculamos a média pela técnica do ponto médio e usando as frequências absolutas de cada classe como pesos.

$$\mu = Me = \frac{200.2 + 400.4 + 200.6 + 200.8}{200 + 400 + 200 + 200} = \frac{400 + 1600 + 1200 + 1600}{1000}$$

$$\mu = \frac{4800}{1000} = 4{,}8$$

Letra d.

2.1.2. Valor Esperado

O valor esperado é um conceito muito importante na estatística. É tão importante, que mudará completamente a sua forma de ver e lidar com as probabilidades.

Consiste numa média aritmética, mas em que os pesos são valores muito especiais: são probabilidades.

Suponha que você trabalhe para uma empresa de seguros e que você precisará estimar o custo médio de uma determinada cirurgia. Com base nas suas estatísticas, você sabe que a probabilidade de que o custo da cirurgia seja dada pela Tabela 5.

Tabela 5: distribuição de probabilidades para os gastos de uma cirurgia

Probabilidade	Gastos (X)
40%	R$4.000
30%	R$6.000
20%	R$8.000
10%	R$10.000

O valor esperado pode ser obtido como a média aritmética ponderada dos gastos, usando as probabilidades como pesos. É interessante notar que a soma de todas as probabilidades é sempre igual a 1, por definição.

$$E[X] = \frac{0{,}40.4000 + 0{,}30.6000 + 0{,}20.8000 + 0{,}10.10000}{0{,}40 + 0{,}30 + 0{,}20 + 0{,}10}$$

$$E[X] = 1600 + 1800 + 1600 + 1000 = 6000$$

Dessa forma, podemos generalizar uma expressão para o cálculo dessa importante grandeza.

$$E[X] = \sum_{i=1}^{N} p_i x_i$$

O valor esperado é uma grandeza muito importante da estatística. É uma ferramenta muito útil em apostas, que lidam com a incerteza.

Suponha que eu te ofereça a seguinte aposta: eu vou rolar um dado. Se der 1 a 4, você me paga R$ 1.000, se der 5, ficamos quites; mas, se der 6, eu te pago R$ 4.900. Você aceitaria essa aposta?
Note que você tem 2/3 de chances de perder e apenas 1/6 de chances de ganhar. Você aceitaria?
Espero que você não tenha se intimidado. Embora a probabilidade de você perder seja maior, o que importa é o quanto você perde quando está errado e o quanto você ganha quando está certo. E o valor esperado leva em consideração todos esses fatores.
Se você entrar na aposta, o valor esperado é:

$$E[X] = \frac{2}{3} \cdot (-1000) + \frac{1}{6} \cdot (4900) = \frac{-2000}{3} + \frac{4900}{6}$$

$$E[X] = -\frac{4000}{6} + \frac{4900}{6} = \frac{4900 - 4000}{6} = \frac{900}{6} = 150$$

Portanto, o valor esperado da aposta é de R$ 150 com vantagem para você.

É interessante também a resposta que, certa vez, recebi de um aluno: "Se você se comprometer a fazer pelo menos 100 vezes a mesma aposta, eu aceito, sim."

Muito boa a resposta, não? Realmente, se você lançar o dado somente uma vez, você pode ter azar e, com isso, perder R$ 1.000. Porém, se você repetir o mesmo experimento várias vezes, as probabilidades vão se impor. E você vai ganhar mais dinheiro do que perder.

Agora, vamos a um ponto polêmico.

Imagine que você está fazendo uma prova de questões Certo ou Errado, em que uma questão errada anula uma questão certa. Suponha, agora, que existe um item em que você está na dúvida. Vale mais a pena deixar em branco ou chutar?
O que você faria?
Bom. A opinião mais comum é deixar a questão em branco. E a razão para isso é muito simples: o ser humano é avesso ao risco. Isso significa que nós temos muito mais medo de perder do que vontade de ganhar.

Para nós, é muito mais doloroso saber que podemos perder 1 ponto de uma questão, para a qual nós nos esforçamos e passamos noites a fio estudando para saber marcar corretamente na hora da prova, do que ganhar 1 ponto de uma questão que você nem estava completamente preparado(a) para fazer, não é mesmo?

Porém, eu queria apresentar uma visão diferente: a visão estatística. Suponha que você não sabe nada de uma questão. Por exemplo, imagine que eu estivesse numa prova de certo e errado em árabe. Eu nem mesmo sei localizar onde se encontra o símbolo do certo e o símbolo do errado.

المتوسط الحسابي هو دائمًا رقمينتمي إلى العينة.
أطخ ( ) حق ( )

Nessa situação, em que você não tem absolutamente nenhuma ideia de como resolver a questão, a sua chance de acertar é exatamente igual a 50%. Se você acertar a questão, você ganha 1 ponto. Se você errar, você perde 1 ponto. Dessa forma, o valor esperado da sua pontuação na questão é:

$$E[X] = 0{,}50.(+1) + 0{,}50.(-1) = 0{,}50 - 0{,}50 = 0$$

Portanto, nessa situação, seria indiferente chutar ou deixar a questão em branco. Mas, o que aconteceria, se você soubesse alguma coisa? Você estudou a matéria, então, você está acima das pessoas que não sabem nada, não é mesmo?
Se a sua probabilidade de acertar uma questão for igual a 60% e a probabilidade de errar for igual a 40%, qual seria o valor esperado do seu chute?

$$E[X] = 0{,}60.(+1) + 0{,}40.(-1) = 0{,}60 - 0{,}40 = 0{,}20$$

Observe que o valor esperado é positivo. Isso significa que, na média, você ganhará mais pontos acertando as questões no chute do que deixando-as em branco.

Direto do Concurso

011. (IFPA/2019/ESTATÍSTICO) Uma variável aleatória discreta Z tem função de probabilidade dada por:

Z	-2	-1	0	1	2
P(Z)	0,3	0,1	0,2	0,3	0,5

Pode-se afirmar que a média da variável aleatória Z é igual a:
a) 0,9
b) 0,6
c) -0,6
d) -0,9
e) 1

A média da variável aleatória Z, ou seu valor esperado, pode ser obtido com o somatório dos produtos, os valores possíveis da variável Z e suas probabilidades de ocorrência respectivas.

$$E[Z] = 0{,}3.(-2) + 0{,}1.(-1) + 0{,}2.0 + 0{,}3.(+1) + 0{,}5.(+2)$$

$$E[Z] = -0{,}6 - 0{,}1 + 0 + 0{,}3 + 1{,}0 = 1{,}3 - 0{,}7 = 0{,}6$$

Letra b.

2.2. Propriedades da Média

Vamos estudar em detalhes algumas propriedades da média.

2.2.1. As Quatro Operações

Uma conclusão muito importante sobre a média é que **esta é influenciada pelas quatro operações.**

Isso significa que:

- quando acrescentamos uma constante a todas as observações da amostra, essa mesma constante será também somada à média;
- quando multiplicamos todas as observações por uma constante, a média também será multiplicada pela mesma constante.

O que foi dito para a soma também vale para a subtração, e o que foi dito para a multiplicação também vale para a divisão.

Vamos ver um exemplo de como essas propriedades funcionam?

Imagine que a mesma pesquisa mostrada na Tabela 1 para avaliar as idades dos funcionários de uma empresa venha a ser aplicada 5 anos mais tarde. Se o quadro de funcionários da empresa se mantiver inalterado nesse período, chegaríamos aos valores registrados na Tabela 6.

Tabela 6: conjunto de dados sobre a idade dos funcionários da empresa X

Pessoa	Idade	Idade 5 anos depois
1	28	33
2	27	32
3	28	33
4	36	41
5	34	39
6	51	56
7	43	48
8	29	33

Pessoa	Idade	Idade 5 anos depois
9	32	37
10	38	43

Depois de 5 anos, a idade das pessoas terá aumentado em 5 unidades. Quem tinha 28 anos passará a ter 33 anos. Quem tinha 51 anos passará a ter 56 anos. E, assim, por diante.
Já havíamos calculado anteriormente a idade média original, encontrando o valor de 34,6 anos.

$$\bar{X} = \frac{28 + 27 + 28 + 36 + 34 + 51 + 43 + 29 + 32 + 38}{10} = \frac{346}{10} = 34{,}6$$

Agora, vamos calcular a mesma idade média daqui a 5 anos. Para isso, note que somamos 5 unidades em todas as observações.

$$\bar{X} = \frac{33 + 32 + 33 + 41 + 39 + 56 + 48 + 34 + 37 + 43}{10} = \frac{396}{10} = 39{,}6$$

Portanto, a idade média também aumentou em 5 unidades, passando de 34,6 anos para 39,6 anos.

Para verificarmos a propriedade da média em relação ao produto, podemos recorrer ao exemplo mostrado na Tabela 3, que tratou dos salários dos empregados de uma determinada empresa.

Suponha que, devido ao excelente resultado anual, o dono resolveu duplicar os salários de todos os funcionários. Nessa situação, o que acontecerá com a média salarial?

Como todos os salários foram duplicados, os 50 empregados que estavam na faixa de 1 a 3 salários mínimos passaram à faixa de 2 a 6 salários mínimos – a categoria também foi duplicada. Os 32 empregados que estavam na faixa de 3 a 5 salários mínimos passaram à faixa de 6 a 10 salários mínimos. E, assim, por diante.

Com base nessas informações, construímos a Tabela 7.

Tabela 7: os salários dos funcionários de uma empresa após a duplicação

Faixa salarial original	Faixa salarial depois do aumento	Número de empregados
1 a 3	**2 a 6**	50
3 a 15	**6 a 10**	32
5 a 7	**10 a 14**	12
7 a 9	**14 a 18**	6

Já havíamos calculado a média salarial, considerando as faixas salariais originais, utilizando a técnica do ponto médio.

$$\bar{X} = \frac{50.2 + 32.4 + 12.6 + 6.8}{50 + 32 + 12 + 6} = \frac{100 + 128 + 72 + 48}{100} = \frac{348}{100} = 3{,}48$$

Para calcular a nova média salarial, precisamos calcular os novos pontos médios das novas categorias. Observe que, como as faixas salariais foram duplicadas, os pontos médios também serão duplicados.

Tabela 8: os salários dos funcionários de uma empresa após a duplicação

Faixa salarial	Ponto médio	Número de empregados
2 a 6	**4**	50
6 a 10	**8**	32
10 a 14	**12**	12
14 a 18	**16**	6

Dessa forma, a nova média salarial pode ser obtida como a média ponderada, utilizando o número de empregados em cada categoria como pesos.

$$\bar{X} = \frac{50.4 + 32.8 + 12.12 + 6.16}{50 + 32 + 12 + 6} = \frac{200 + 256 + 144 + 96}{100} = \frac{696}{100} = 6{,}96$$

Perceba, assim, que a nova média também foi duplicada, passando de 3,48 para 6,96.

2.2.2. Combinação entre Duas Variáveis

A média aritmética é tão importante que algumas propriedades devem ser destacadas.

A média aritmética é um **operador linear**. Isso significa que ele atende à propriedade de que **a média da soma de duas variáveis aleatórias é igual à soma das médias.**

$$\mu[X+Y]=\mu[X]+\mu[Y]$$

Já vimos um caso particularmente importante quando ocorre a soma de uma variável aleatória com uma constante. Nesse caso, basta adicionar o valor dessa constante.

$$\mu[X+a]=\mu[X]+a$$

É importante citar que **não há garantias** de que a média do produto seja igual ao produto das esperanças. As únicas propriedades que realmente podemos sempre garantir do valor esperado são as duas relacionadas à sua linearidade.

Vamos trazer um exemplo. Primeiro, vamos considerar X e Y como duas variáveis aleatórias, cujos valores são listados na Tabela 9.

Tabela 9: cálculo de valor esperado

	X	Y	X − + Y
	23	16	39
	14	13	27
	18	11	29
	19	8	27
	8	25	33
Soma	**82**	**73**	**155**

As médias de X e Y podem ser calculadas como a soma das observações, que já foi destacada na Tabela 9.

$$\mu[X]=\frac{23+14+18+19+8}{5}=\frac{82}{5}=16{,}4$$

$$\mu[Y]=\frac{16+13+11+8+25}{5}=\frac{73}{5}=14{,}6$$

A média da soma X + Y pode ser também calculada pela mesma técnica.

$$\mu[X+Y] = \frac{39+27+29+27+33}{5} = \frac{155}{5} = 31$$

Observe que o valor da média da soma (31) coincidiu com a soma das médias (16,4 + 14,6 = 31).

$$\mu[X+Y] = \mu[X] + \mu[Y]$$

$$31 = 16{,}4 + 14{,}6$$

2.2.3. Existência e Unicidade

A média aritmética existe e é sempre única para qualquer tipo de amostra.

Essa é uma observação interessante a respeito da média aritmética e a diferencia de outras medidas de posição. Por exemplo, uma amostra qualquer pode não ter nenhuma moda ou ainda ter várias modas, como será visto mais adiante.

A média não necessariamente pertence ao conjunto de valores da variável. Voltemos ao exemplo do número de filhos por casal.

Tabela 10: número de filhos por casal

Número de filhos	Quantidade de casais
0	17
1	33
2	21
3	7
4	2
5	0

Nesse caso, o número médio de filhos dos casais é obtido também pela média ponderada, em que a quantidade de casais serve como **peso.**

$$\bar{X} = \frac{17.0+33.1+21.2+7.3+2.4+0.5}{17+33+21+7+2+0} = \frac{0+33+42+21+8+0}{80}$$

$$\bar{X} = \frac{104}{80} = 1{,}3$$

A média de filhos dos casais da amostra é igual a 1,3.

Observe que um casal pode ter 0 filho, 1 filho, 2 filhos, 3 filhos etc. Mas não é possível que um casal tenha 1,3 filho. Não existe nenhum casal que tenha a quantidade de filhos igual à média obtida.

2.2.4. Sensibilidade

A média envolve todos os elementos da amostra. Por isso, é bastante sensível à presença de um **outlier.**

Obs.: O outlier é um elemento que destoa dos demais.

Por exemplo, voltemos ao caso de estudo sobre o número de filhos por casal.

Tabela 11: número de filhos por casal

Número de filhos	Quantidade de casais
0	4
1	5
2	7
3	3
4	0
5	1

A média aritmética é calculada como a média ponderada do número filhos, utilizando a quantidade de casais como peso.

$$\bar{x} = \frac{4.0 + 5.1 + 7.2 + 3.3 + 0.4 + 1.5}{4 + 5 + 7 + 3 + 0 + 1} = \frac{33}{20} = 1{,}65$$

Suponha que a esse grupo fosse adicionado um casal que tivesse 20 filhos. Esse casal é um **outlier**, pois ele tem um número de filhos muito superior ao dos outros casais. Enquanto os demais têm de 0 a 4 filhos, esse casal é o único que tem 20 filhos.

Ao adicionar o outlier, o que aconteceria com a média?

Para descobrirmos, precisamos atualizar os dados da Tabela 11.

Tabela 12: número de filhos por casal com um outlier

Número de filhos	Quantidade de casais
0	4
1	5
2	7
3	3
4	0
5	1
20	1

Calculando a nova média ponderada:

$$\bar{x} = \frac{4.0 + 5.1 + 7.2 + 3.3 + 0.4 + 1.5 + 1.20}{4 + 5 + 7 + 3 + 0 + 1} = \frac{53}{21} = 2{,}53$$

Observe que, com **a adição de um único elemento outlier** à amostra, **a média passou de 1,65 para 2,65** – um aumento bastante expressivo.

Perceba que, quando existem outliers, a média pode ser bastante **enganosa.**

Imagine que você conta para a sua mãe que você vai trabalhar em uma empresa, cujos funcionários têm uma média de filhos igual a 2,65.

Ela ficará muito animada, imaginando a probabilidade de ter 3 netos, não é? Pelo menos, a minha ficaria.

Porém, não é exatamente isso. O que aconteceu é que apenas 5 pessoas na empresa possuem três filhos ou mais: três delas possuem exatamente três filhos, uma delas possui cinco filhos e **uma delas é um caso extremo** e possui 20 filhos. A imensa maioria das pessoas naquela empresa possui até dois filhos.

Imagine que você esteja procurando uma escola de ensino médio para seu filho e você procura informações sobre a renda média dos ex-alunos daquela instituição.
Durante suas pesquisas, você encontra que o salário médio dos alunos recém-formados que vieram da Lakeside High School é um número impressionante, superior a US$ 2 milhões por ano.

Após constatar essa quantia inigualável, você fica entusiasmado(a) e procura um meio de colocar o seu filho para estudar nessa escola a todo custo. Você já começa a imaginar o seu filho dirigindo uma Lamborghini dentro de alguns anos.
Até que você descobre que, por coincidência, Bill Gates e Paul Allen, ambos multimilionários, foram ex-alunos dessa instituição.

Dessa maneira, realmente não é estranho que o salário médio dos alunos daquela escola seja superior a US$ 2 milhões por ano. Trata-se de uma pequena escola que teve dois alunos que se tornaram multimilionários.

Ainda que todos os demais ex-alunos daquela instituição estivessem desempregados e sem gerar nenhuma renda, somente esses dois elementos seriam suficientes para levar o salário médio às alturas, como aconteceu nessa ilustração.

Portanto, a informação de que o salário médio dessa escola é altíssimo não quer dizer nada sobre a qualidade dessa escola. Porém, revela algo muito interessante: **a média é sensível à presença de valores extremos.** Podemos dizer, ainda, que **a média não é uma estimativa robusta.**

2.2.5. Média das Médias

Suponha que as duas tabelas sobre o número de filhos de casais se refiram a funcionários de uma mesma empresa que pertencem a dois setores diferentes.

Tabela 13: número de filhos por casal em dois setores de uma empresa

Número de filhos	Quantidade de casais	Número de filhos	Quantidade de casais
0	4	**0**	17
1	5	**1**	33
2	7	**2**	21
3	3	**3**	7
4	0	**4**	2
5	1	**5**	0
Setor A **Média: 1,65** **Tamanho: 20 pessoas**		**Setor B** Média: 1,30 Tamanho: 80 pessoas	

Já calculamos separadamente as médias para o número de filhos por casal em cada uma das duas situações. Podemos também obter **a média global** do número de filhos por casal nessa empresa, considerando os dois setores.

A forma mais fácil de fazer isso é pela **média ponderada** das médias de cada um dos setores. Devemos utilizar o tamanho de cada amostra como peso.

$$Média = \frac{20.1{,}65 + 80.1{,}30}{20 + 80} = \frac{33 + 104}{100} = \frac{137}{100} = 1{,}37$$

Uma forma de nos certificarmos disso é montar uma nova tabela calculando o total de casais que possuem 0, 1, 2, 3, 4 ou 5 filhos. Observe que temos 4 casais no Setor A com 0 filho e 17 casais no Setor B com 17 filhos. No total, são 21 casais com 0 filho. E, assim, por diante.

Tabela 14: número de filhos por casal em dois setores de uma empresa

Número de filhos	Quantidade de casais Setor A	Quantidade de casais no Setor B	Quantidade de casais no total
0	4	17	21
1	5	33	38
2	7	21	28
3	3	7	10
4	0	2	2
5	1	0	1

A média do número de filhos por casal pode ser calculada como uma média ponderada, seguindo a mesma técnica que já vimos anteriormente.

$$Média = \frac{21.0 + 38.1 + 28.2 + 10.3 + 2.4 + 1.5}{21 + 38 + 28 + 10 + 2 + 1} = \frac{137}{100} = 1{,}37$$

DIRETO DO CONCURSO

012. (CESPE/SEDUC-AL/2018/PROFESSOR DE MATEMÁTICA) **Situação hipotética**: A média aritmética dos pesos dos 60 alunos de uma sala de aulas é igual a 51,8 kg. Nessa sala, a média aritmética do peso dos meninos é de 62 kg e das meninas, 45 kg. **Assertiva**: Nesse caso, essa sala de aulas tem 24 meninos e 36 meninas.

A média global dos pesos dos alunos da turma é igual à média aritmética dos pesos dos meninos e das meninas ponderada pela quantidade de meninos e meninas na sala.

Supondo que a sala tenha realmente 24 meninos e 36 meninas, como afirmado pela assertiva, a média global seria:

$$\mu = \frac{24.62 + 36.45}{24 + 36} = \frac{1488 + 1620}{60} = \frac{3108}{60} = 51{,}8$$

Portanto, a proporção de 24 meninos e 36 meninas realmente é coerente com a média igual a 51,8 kg.

Certo.

---

013. (CESPE/ANP/2013/ESPECIALISTA EM REGULAÇÃO) Considere duas amostras aleatórias de tamanhos diferentes, em que a soma dos valores observados sejam S_1 e S_2 e as respectivas médias $\overline{x_1}$ e $\overline{x_2}$. Nesse caso, a média aritmética global, ou seja, a média dos valores observados nas duas amostras é igual à média das médias $\overline{x_1}$ e $\overline{x_2}$.

A média global deve ser calculada como a média aritmética ponderada pelo número de elementos de cada conjunto das médias $\overline{x_1}$ e $\overline{x_2}$.

$$\bar{x} = \frac{n_1\overline{x_1} + n_2\overline{x_2}}{n_1 + n_2}$$

Sendo assim, não basta tomar a média das médias $\overline{x_1}$ e $\overline{x_2}$, é preciso tomar a média aritmética ponderada.

É bem verdade que, com boa vontade, deveríamos entender que a média aritmética ponderada é também uma forma de média. Portanto, a questão não está completamente errada.

Porém, para provas de concursos, considere que, quando a questão falar em média, está falando da média aritmética simples, ou seja, sem pesos.

Errado.

2.2.6. Desvios em Relação à Média

A soma de todos os desvios em relação à média aritmética é sempre igual a zero.

Trata-se de uma propriedade bastante particular que diferencia a média aritmética de todas as outras medidas de posição. Essa propriedade é muito importante, pois está na base da definição de desvio-padrão.

O desvio corresponde à diferença entre cada valor observado e a média aritmética. Voltemos novamente ao exemplo do número de filhos por casal, cuja média já calculamos, igual a 1,65.

Tabela 15: desvio do número de filhos por casal em relação à média aritmética

Número de filhos	Quantidade de casais	Desvio
0	4	0 – 1,65 =–1,65
1	5	1 – 1,65 = –0,65
2	7	2 – 1,65 = 0,35
3	3	3 – 1,65 = 1,35
4	0	4 – 1,65 = 2,35
5	1	5 – 1,65 = 3,35

Vamos, agora, tomar a soma de todos os desvios em relação à média. Note que são 4 casais com desvio igual a –1,65; 5 casais com desvio iguais a –0,65; 7 casais com desvio igual a 0,35; 3 casais com desvio igual a 1,35; e 1 casal com desvio igual a 3,35.

$$S = 4.(-1{,}65) + 5.(-0{,}65) + 7.(0{,}35) + 3.(1{,}35) + 1.(3{,}35)$$

$$S = \ -6{,}6 - 3{,}25 + 2{,}45 + 4{,}05 + 3{,}35 = -9{,}85 + 9{,}85 = 0$$

Essa é uma propriedade geral que sempre pode ser verificada para qualquer média aritmética.

2.3. Outros Tipos de Médias

Vamos, agora, estudar a média geométrica e a média harmônica. Esses tipos de média aparecem menos em questões de provas, porém vale a pena estudá-las.

2.3.1. Média Geométrica

A média geométrica consiste na raiz N-ésima dos produtos dos N termos.

$$\bar{x} = \sqrt[N]{x_1 x_2 \ldots x_N}$$

Para o caso da amostra de 10 elementos que mostramos anteriormente:

$$\bar{x} = \sqrt[10]{28.27.28.36.34.51.43.29.32.38} = 33{,}9$$

Portanto, a média geométrica é de 33,9 anos.

Em termos práticos, é muito difícil que uma questão de prova cobre uma média geométrica com mais de dois termos, porque não faz sentido você ter de extrair uma raiz de índice 10 durante uma prova de concurso.

No entanto, é importante saber que a média geométrica existe e, principalmente, saber a desigualdade das médias, que ainda veremos neste capítulo.

2.3.2. Média Harmônica

A média harmônica consiste no inverso da média aritmética dos inversos.

$$\frac{1}{\bar{x}} = \frac{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_N}}{N}$$

Outra forma de representar a média harmônica é passar o denominador N para o outro lado.

$$\frac{N}{\bar{x}} = \frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_N}$$

A média aritmética deve ser utilizada para grandezas racionais, quando se utilizar uma ponderação pelo numerador.

Como exemplo, tomemos a velocidade de um veículo, que é definida como a razão entre o espaço percorrido e o tempo gasto.

$$v = \frac{\Delta S}{\Delta t}$$

Se a ponderação for pelo espaço percorrido, utiliza-se a média harmônica

Se a ponderação for pelo tempo gasto, utiliza-se a média aritmética.

Para exemplificar, considere duas situações:

- um veículo percorre uma estrada a 120 km/h nos primeiros 200 km do percurso e a 60 km/h nos 200 km finais do percurso;
- um veículo percorre uma estrada a 120 km/h na primeira hora do percurso e a 60 km/h na segunda hora do percurso.

Qual a velocidade média nos dois casos?

Observe que a velocidade é igual à razão entre o espaço percorrido e o tempo. Como o espaço percorrido está no numerador, na primeira situação, deve-se utilizar a média harmônica.

$$\frac{2}{v_1} = \frac{1}{120} + \frac{1}{60}$$

$$\frac{2}{v_1} = \frac{1+2}{120} = \frac{3}{120} \therefore v_1 = \frac{2.120}{3} = 80\ km/h$$

Na segunda situação, observe que o tempo está no denominador da expressão da velocidade. Nesse caso, deve-se utilizar a velocidade média.

$$v_2 = \frac{120 + 60}{2} = \frac{180}{2} = 90\ km/h$$

Um exemplo de aplicação da média harmônica é a densidade de ligas metálicas. A densidade é a razão entre a massa e o volume de um objeto.

De maneira geral, a pureza do ouro em uma liga metálica é dada em proporção de massa, que está no denominador. Por isso, a densidade da liga deve ser calculada pela média harmônica.

Se a ponderação for feita pela massa, utiliza-se a média harmônica

$$d = \frac{m}{v}$$

Se a ponderação for pelo tempo gasto, utiliza-se a média aritmética.

Com o conceito de densidade, podemos determinar se uma aliança (ou qualquer outra joia) é realmente formada por ouro 18 quilates. Para isso, devemos saber que a densidade do ouro puro, também conhecido como 24 quilates, é igual a 19,3 g/cm³. Porém, o ouro puro é raramente utilizado para a fabricação de joias, pois é um metal muito mole.

O normal é utilizar ligas, em que se misturam outros metais ao ouro, como a prata e o cobre. A mistura de ouro e prata produz o ouro amarelo, enquanto a mistura de ouro e cobre produz o ouro rosa.

O teor mais utilizado de ouro nas joias é justamente o ouro 18 quilates, que corresponde a uma medida de pureza do metal em massa. Se o ouro 24 quilates é o ouro 100% em massa, o ouro 18 quilates será:

$$p_{ouro} = \frac{18}{24} = \frac{3}{4} = 0{,}75$$

Supondo que uma peça seja formada exclusivamente por ouro e prata, o ouro 18 quilates é formado por 75% em massa de ouro e 25% em massa de prata. A densidade da peça de ouro 18 quilates será calculada pela média harmônica. Ou seja, o inverso da média aritmética ponderada dos inversos das densidades do ouro e da prata.

$$\frac{1}{d} = \frac{0{,}75}{19{,}3} + \frac{0{,}25}{10{,}5}$$

$$\frac{1}{d} = 0{,}03886 + 0{,}0238 = 0{,}06267$$

$$\therefore d = \frac{1}{0{,}06267} \cong 15{,}96\ g/cm^3$$

Se você tem uma aliança ou outra joia de ouro 18 quilates, você pode conferir se a densidade da peça é realmente coerente com esse valor encontrado.

Se for encontrada uma densidade menor, é um sinal de que a peça pode ter um teor maior de prata do que os 25% prometidos.

A densidade da peça pode ser determinada facilmente em casa por um experimento simples. Tudo o que você precisará é uma balança, um barbante e um copo de água.

- pese a barra de ouro seca com o auxílio da balança;
- encha um copo de água com volume suficiente para submergir a barra de ouro;
- pese a amostra de água e tare a balança;
- amarre o ouro com um barbante e o pendure dentro da água. A barra precisa estar completamente submersa, mas não pode tocar as paredes do recipiente;
- a densidade da peça é aproximadamente igual à razão entre o peso da barra seco e o peso da barra submersa.

Figura 5: determinação da densidade de uma aliança de ouro 18 quilates

No caso da minha aliança, a densidade calculada experimentalmente é:

$$d = \frac{5,80}{0,36} = 16,1\ g/mL$$

Observe que o valor encontrado foi até mesmo um pouco superior ao valor previsto teoricamente pelas nossas contas, que era 15,96 g/mL. Isso significa que pode ser até mesmo que o ourives tenha utilizado um teor ligeiramente superior de ouro ao que foi prometido: 75%. Logo, a minha aliança é provavelmente verdadeira.

2.3.3. Desigualdade das Médias

Em uma amostra qualquer, a média aritmética é sempre maior ou igual à média geométrica que, por sua vez, é maior ou igual à média harmônica.

$$MA \geq MG \geq MH$$

A igualdade somente se verifica quando a amostra for composta por todos os elementos iguais.

Assim, se a amostra tiver pelo menos um elemento distinto dos demais, pode-se afirmar que a média aritmética será maior que a geométrica, que será maior que a harmônica.

Por conta disso, na escola, utiliza-se a média aritmética, pois ela permite que os alunos tenham melhores resultados. Isso facilita bastante que eles passem de ano.

Um caso particular dessa relação acontece quando a amostra é composta por apenas dois elementos. Para a estatística, esse é um caso irrelevante, pois as nossas amostras serão bem longas, porém isso pode ser abordado em questões de probabilidade.

Quando a amostra é composta por apenas dois elementos, tem-se que a média geométrica é a média da aritmética com a harmônica. Vejamos.

A média geométrica é a raiz quadrada do produto dos dois elementos. Portanto:

$$MG = \sqrt{x_1 x_2} \therefore MG^2 = x_1 x_2$$

Por outro lado, a média harmônica é o inverso da média aritmética dos inversos.

$$\frac{1}{MH} = \frac{\frac{1}{x_1} + \frac{1}{x_2}}{2} = \frac{x_1 + x_2}{2x_1 x_2}$$

Note que a média aritmética (MA) acabou de aparecer no lado direito.

$$\frac{1}{MH} = \frac{MA}{MG^2} \therefore MG^2 = MA.MH \therefore MG = \sqrt{MA.MH}$$

Direto do Concurso

014. (FCC/SEFAZ-GO/2018/AUDITOR-FISCAL) Os matemática definem diferentes tipos de médias entre dois números positivos e, para cada aplicação, escolhem qual o tipo mais adequado a ser utilizado. A média harmônica H entre os números positivos a e b, por exemplo, é definida como o inverso da média aritmética dos inversos desses números, ou seja,

$$H = \frac{\frac{1}{a} + \frac{1}{b}}{2}$$

A média aritmética dos números 5 e 20 supera a média harmônica desses mesmos números em

a) 4 unidades.
b) 4,25 unidades.
c) 4,5 unidades.
d) 4,75 unidades.
e) 5 unidades.

Para calcular a média harmônica, primeiro vamos calcular a soma dos inversos.

$$\frac{1}{20}+\frac{1}{5}=\frac{1+4}{20}=\frac{5}{20}=\frac{1}{4}$$

Portanto, a média aritmética dos inversos é:

$$\frac{1/4}{2}=\frac{1}{8}$$

Por fim, a média harmônica é o inverso da média aritmética dos inversos.

$$H=\frac{1}{1/8}=8$$

Agora, vamos comparar com a média aritmética comum, que é simplesmente a soma dos dois números dividida por 2.

$$A=\frac{5+20}{2}=\frac{25}{2}=12{,}5$$

Logo, a diferença entre elas é:

$$A-H=12{,}5-8=4{,}5$$

Portanto, a média aritmética supera a média harmônica em 4,5 unidades.

Letra c.

---

015. (CEPERJ/SEFAZ-RJ/2013/ANALISTA DE CONTROLE INTERNO) Um filme foi exibido em um cinema em 8 diferentes sessões, ao longo de todo o dia. O número de presentes em cada sessão é mostrado na tabela abaixo:

Sessão	Número de presentes	Sessão	Número de presentes
1	88	5	94
2	102	6	82
3	90	7	80
4	76	8	68

O número médio de pessoas por sessão corresponde a:

a) 68
b) 72
c) 76
d) 81
e) 85

O número médio pode ser obtido com a soma de todas as observações da variável aleatória dividida pelo número de sessões.

$$\mu = \frac{88 + 102 + 90 + 76 + 94 + 82 + 80 + 68}{8} = \frac{680}{8} = 85$$

Letra e.

016. (VUNESP/PREFEITURA DE SÃO JOSÉ DOS CAMPOS/2012/AUDITOR TRIBUTÁRIO MUNICIPAL – GESTÃO TRIBUTÁRIA) A média aritmética de alturas de 10 alunos de um time de futebol é 175 cm. Dois novos alunos entram para o time, e a nova média de alturas passa a ser 178 cm. Se a diferença entre as alturas desses dois novos jogadores é 6 cm, o maior dos dois mede, em cm,

a) 188.
b) 190.
c) 192.
d) 194.
e) 196.

A média pode ser obtida como a razão entre a soma de todas as alturas dividida pelo total de alunos. Portanto, na situação inicial, a média da altura dos alunos pode ser obtida como a soma de todas as 10 alturas, que representaremos por S, e o número de alunos presentes no time originalmente, que é igual a 10.

$$\mu = \frac{S}{10} = 175$$

Assim, podemos calcular a soma da altura S.

$$\therefore S = 10.175 = 1750$$

A seguir, com o ingresso de dois alunos de alturas iguais a **x** e **y**, a nova média de alturas passou a ser:

$$\mu_2 = \frac{S + x + y}{12} = 178$$

$$\therefore S + x + y = 12.178$$

Utilizando o valor calculado previamente para a soma das alturas dos alunos, temos:

$$1750 + x + y = 2136$$

$$\therefore x + y = 2136 - 1750$$

$$\therefore x + y = 386$$

Sabemos também que a diferença de alturas entre os dois alunos é igual a:

$$x - y = 6$$

Chegamos, portanto, a um sistema com duas equações e duas incógnitas. Podemos somá-las.

$$(x + y) + (x - y) = 386 + 6$$

$$2x = 392$$

$$\therefore x = \frac{392}{2} = 196$$

Letra e.

---

3. Moda

A moda é a segunda medida de posição que nós vamos estudar. Como o próprio nome diz, ela se refere **ao valor observado mais frequente** da amostra.

Quando os dados de uma amostra são apresentados na forma de um rol, basta contar cada um deles:

{1, 1, **2, 2, 2,** 3, 3, 4, 5, 6}

Como destacado no rol em estudo, o número **2** é visto por três vezes. Portanto, a moda da amostra é 2.

Tome cuidado para não cair na seguinte pegadinha: a moda de uma amostra é o valor mais frequentemente observado, e não a quantidade de vezes que esse valor aparece no rol.

Nesse caso, o número **2** aparece três vezes. Portanto, a moda é igual a 2, e não igual a 3.

Direto do Concurso

017. (FCC/ALESE/2018/ANALISTA LEGISLATIVO – ECONOMIA/ADAPTADA) Em um grupo de pessoas encontramos as seguintes idades: 20, 30, 50, 39, 20, 25, 41, 47, 36, 45, 41, 52, 18, 41. A moda é:

a) 42.
b) 45.
c) 41.
d) 20.
e) 39.

Notemos que o número 20 se repete duas vezes.
20, 30, 50, 39, **20**, 25, 41, 47, 36, 45, 41, 52, 18, 41.
Por sua vez, o número 41 se repete três vezes.
20, 30, 50, 39, 20, 25, **41**, 47, 36, 45, **41**, 52, 18, **41.**
Nenhum outro número se repete três vezes ou mais. Portanto, a moda é realmente o número 41.
Letra c.

3.1. Distribuições Multimodais

Ao contrário da média aritmética, a moda nem sempre existe. E, quando existe, pode não ser única.

Podemos classificar uma distribuição pelo número de modas que ela apresenta.

- **Distribuição amodal:** são as distribuições que não possuem nenhuma moda.

X – = {1, 2, 4, 8, 16, 32, 64}

Nessa amostra, não conseguimos observar nenhum valor que se repita em quantidade de vezes maior que os demais. Por isso, a distribuição não apresenta moda.

- **Distribuição unimodal:** possui uma única moda;

É o caso da distribuição estudada no começo desta seção.

Y = {1, 1, **2, 2, 2,** 3, 3, 4, 5, 6}

Essa distribuição apresenta um valor único de moda Mo(Y) = 2.

Distribuição bimodal: possui duas modas;

Z = {1, 1, **2, 2, 2,** 3, 3, **4, 4, 4,** 5, 6}

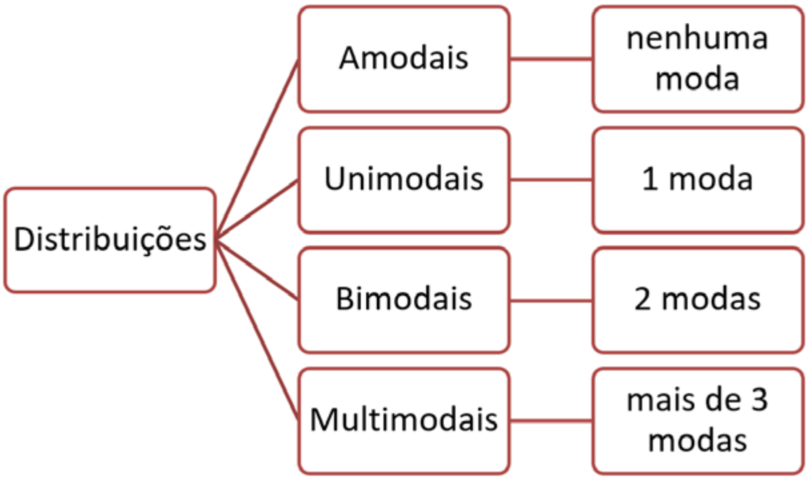
Nessa distribuição, observamos que existem dois valores de moda: **2** e **4**. Esses dois valores se repetem exatamente três vezes, que é mais do que qualquer outro elemento dentro da amostra.

- **Distribuição multimodal:** possui três ou mais modas.

W = {1, 1, **2, 2, 2,** 3, 3, **4, 4, 4,** 5, **6, 6, 6**}

Nesse caso, podemos enxergar três valores de moda: **2, 4 e 6.** Os três valores se repetem exatamente três vezes.

Essa distribuição também pode ser chamada de **trimodal**, em referência às suas três modas.



3.2. Moda em Distribuições de Frequências

Quando a moda é fornecida em distribuições tabeladas ou em rol, é a medida de posição mais simples de calcular.

Por exemplo, retornando à Figura 2, temos que:

Figura 6: gráfico sobre o número de filhos de diversos casais

Como 7 pessoas possuem 2 filhos e não existe nenhuma classe com um número maior de filhos, então a moda é exatamente igual a 2.

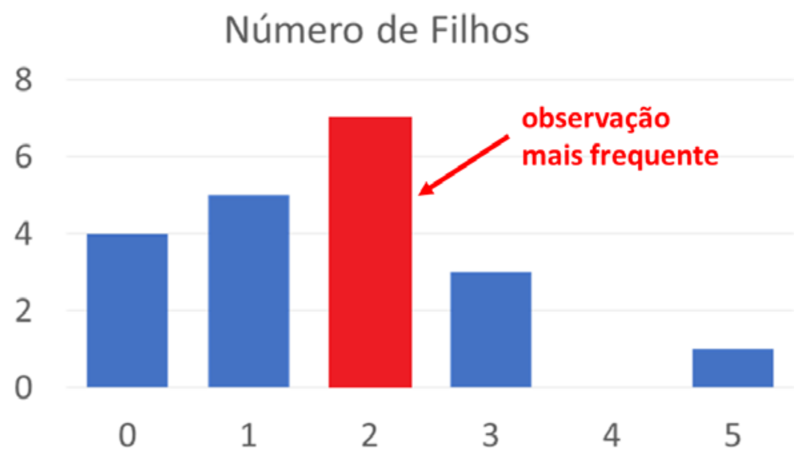


Figura 7: visualização da moda

Representamos assim:

$$Mo = 2$$

Podemos concluir, portanto, que a moda é sempre um elemento da amostra. Nesse caso, é o elemento n = 2 filhos.

A moda também pode ser obtida facilmente quando os valores são fornecidos na forma de tabela.

Tabela 16: número de filhos por casal

Número de filhos	Quantidade de casais
0	17
1	33
2	21
3	7
4	2
5	0

Nesse caso, devemos tomar simplesmente a observação mais frequente.

Número de Filhos	Quantidade de Casais
0	17
1	33
2	21
3	7
4	2
5	0

observação
mais frequente

Figura 8: visualização da moda em uma tabela

Nesse caso, a moda é igual a 1, porque 33 casais possuem exatamente 1 filho e não há nenhuma outra categoria que tenha um número maior de filhos.

3.3. Propriedades da Moda

A moda compartilha as propriedades que vimos para a média aritmética em relação às quatro operações.

- Quando acrescentamos uma constante a todas as observações da amostra, essa mesma constante será também somada à moda.
- Quando multiplicamos todas as observações por uma constante, a moda também será multiplicada pela mesma constante.

Para exemplificar, suponha que uma empresa tenha entregado um novo filho para cada um dos casais participantes da pesquisa mostrada na Tabela 2. Com isso, o número de filhos após a doação aumentaria em uma unidade para cada pessoa.

Tabela 17: número de filhos por casal

Número de filhos inicial	Número de filhos após a doação	Quantidade de casais
0	**1**	17
1	**2**	**33**
2	**3**	21
3	**4**	7
4	**5**	2
5	**6**	0

Em laranja, marcamos a observação mais frequente, que corresponde exatamente à moda.

Observe que, na situação original, a moda amostral era igual a 1. Depois que o número de filhos de cada um dos participantes da pesquisa foi aumentado em uma unidade, a moda também foi aumentada em uma unidade, passando a 2.

Da mesma forma, vejamos o que aconteceria, se, por acaso, em vez de doar um filho para cada participante da pesquisa, a empresa resolvesse duplicar o número de filhos de cada família. Ou seja, se a família não tem nenhum filho, ela continuaria com 0 filho. Se a família tem 1 filho, passaria a ter 2. Se tivesse 2 filhos, passaria a ter 4. E, assim, por diante.

Tabela 18: número de filhos por casal

Número de filhos inicial	Número de filhos após a doação	Quantidade de casais
0	**0**	17
1	**2**	**33**
2	**4**	21
3	**6**	7
4	**8**	2
5	**10**	0

A propriedade mais interessante da moda é **a robustez.** Ao contrário da média, a moda não é sensível à presença de valores extremos. Dessa vez, vamos introduzir novamente uma pessoa na amostra que possui 20 filhos.

Tabela 19: número de filhos por pessoa

Número de filhos	Número de pessoas
0	4
1	5
2	**7**
3	3

Número de filhos	Número de pessoas
4	0
5	1
20	1

Mesmo com a introdução do outlier de 20 filhos, a moda não se alterou, permanecendo igual a 2 filhos. Diz-se, portanto, que **a moda é robusta**, ou seja, **é resistente à inclusão de valores extremos.**

3.4. Moda em Dados Categorizados

O caso mais complicado de cálculo de moda em uma distribuição é quando temos uma variável contínua, cujos valores são fornecidos em categorias.

Tabela 20: dados sobre os salários dos funcionários de uma empresa

Faixa salarial	Número de empregados
1 a 3	32
3 a 5	50
5 a 7	12
7 a 9	6

Nesse caso, não existe uma **definição de moda.** São feitas apenas estimativas.

O primeiro conceito que precisamos conhecer é a **classe modal**, que é a categoria que apresenta a maior frequência absoluta. Nesse caso, o maior número de empregados.

Na distribuição em apreço, a faixa salarial de 3 a 5 salários mínimos apresenta a maior frequência absoluta – são 50 empregados nessa faixa salarial.

Em todas as estimativas que aprenderemos neste capítulo, consideramos que a moda pertence à classe modal.

3.4.1. Moda Bruta

Na moda bruta, toma-se simplesmente o ponto médio da classe modal. Na distribuição mostrada na Tabela 20, teríamos:

$$Mo = \frac{3+5}{2} = \frac{8}{2} = 4$$

3.4.2. Moda de Czuber

A Fórmula de Czuber é uma das expressões mais conhecidas e importantes para a estimativa da Moda em uma distribuição de frequências categorizada. É a única fórmula que leva em consideração a frequência da classe modal.

O primeiro passo para determinar a moda de Czuber dessa distribuição é determinar a **classe modal**, que é a classe com a maior frequência de empregados. Nesse caso, é a classe de 3 a 5 salários mínimos, que conta com 50 empregados.

Faixa Salarial	Número de Empregados
1 a 3	32
3 a 5	50
5 a 7	12
7 a 9	6

classe modal (3 a 5)
classe anterior (1 a 3)
classe posterior (5 a 7)

Figura 9: classe modal

Devemos destacar:

- **L$_{inf}$:** o limite inferior da classe modal, ou seja, o menor valor possível da classe;
- **h:** amplitude da classe modal, ou seja, a diferença entre o limite superior e o inferior da classe;
- **Δf$_{ant}$:** a diferença entre as frequências da classe modal e da classe anterior;
- **Δf$_{post}$:** a diferença entre as frequências da classe modal e da classe posterior.

No exemplo em apreço, registramos:

$$L_{inf} = 3$$

$$h = 5 - 3 = 2$$

$$\Delta f_{ant} = 50 - 32 = 18$$

$$\Delta f_{ant} = 50 - 12 = 38$$

A expressão da Moda de Czuber é:

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

Entenda que faz bastante sentido que Δf_{ant} esteja no numerador. Quanto maior for o valor de Δf_{ant}, isso significa que a diferença entre o número de elementos na classe modal e na classe anterior é muito grande. Portanto, nessa situação, a moda deve se localizar mais distante da classe anterior.

E isso é bastante coerente. Se Δf_{ant} for elevado, realmente isso aumentará o valor da moda, contribuindo para que a moda se afaste da classe anterior e se aproxime da classe posterior.

Agora, vamos às contas:

$$Mo = 3 + \left(\frac{18}{18+38}\right) \cdot 2$$

$$Mo = 3 + \left(\frac{18}{56}\right) \cdot 2 = 3 + 0{,}3214.2 = 3 + 0{,}643 = 3{,}643$$

3.4.3. Moda de King

Quando falamos de moda em uma distribuição em dados categorizados, geralmente utilizamos o conceito de Czuber. Se não houver previsão expressa sobre isso na questão de prova, utilize a Moda de Czuber.

Porém, é possível também que uma eventual questão mais inovadora venha a falar sobre o conceito de Moda de King.

Analogamente ao que fizemos na Moda de Czuber, devemos primeiramente identificar a **classe modal.**

	Faixa Salarial	Número de Empregados	
	1 a 3	32	classe anterior
classe modal	3 a 5	50	
	5 a 7	12	classe posterior
	7 a 9	6	

Figura 10: classe modal

Agora, devemos destacar:

- **L_{inf}:** o limite inferior da classe modal, ou seja, o menor valor possível da classe;
- **h:** amplitude da classe modal, ou seja, a diferença entre o limite superior e o inferior da classe;
- **f_{ant}:** a frequência da classe anterior à classe modal;
- **f_{post}:** a frequência da classe posterior à classe modal;

No exemplo em apreço, registramos:

$$L_{inf} = 3$$

$$h = 5 - 3 = 2$$

$$f_{ant} = 32$$

$$f_{ant} = 12$$

A expressão da Moda de Czuber é:

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

Entenda que faz bastante sentido que f_{post} esteja no numerador dessa expressão. Quanto maior for a frequência da classe posterior, isso significa que ela terá uma maior quantidade de elementos. Portanto, é natural esperar que a moda esteja mais próxima dela.

De fato, quanto maior f_{post}, maior será o valor registrado pela expressão da Moda de King, contribuindo para aproximar a moda da classe posterior.

Agora, vamos às contas:

$$Mo = 3 + \left(\frac{12}{32 + 12}\right) \cdot 2$$

$$Mo = 3 + \left(\frac{12}{44}\right) \cdot 2 = 3 + 0{,}272.2 = 3 + 0{,}545 = 3{,}545$$

Quando calculamos a Moda de Czuber para essa mesma distribuição, encontramos o valor 3,643, que é muito próximo do valor obtido para a Moda de King.

Vamos esquematizar a diferença entre as duas técnicas de cálculo da moda.

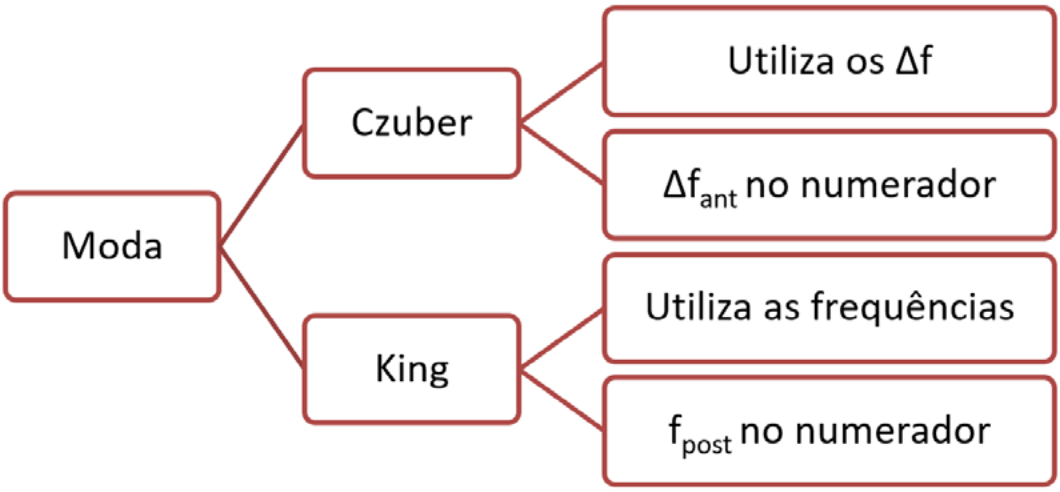


Figura 11: comparação entre a Moda de Czuber e a Moda de King

DIRETO DO CONCURSO

018. (FCC/BACEN/2006) O valor da moda, obtida com a utilização da Fórmula de Czuber, é igual a (desprezar os centavos na resposta):

Salários dos empregados da empresa XYZ em dezembro de 2005

Salários (R$)	Freqüências Simples Absolutas
1 000,00 ⊢ 2 000,00	2
2 000,00 ⊢ 3 000,00	8
3 000,00 ⊢ 4 000,00	16
4 000,00 ⊢ 5 000,00	10
5 000,00 ⊢ 6 000,00	4

a) R$3.201,00
b) R$3.307,00
c) R$3.404,00
d) R$3.483,00
e) R$3.571,00

Devemos lembrar que a Moda de Czuber é dada pelas variações de frequência. No numerador, devemos colocar a variação de frequência em relação à frequência anterior.

$$Mo = L_{inf} + \frac{\Delta f_{ant}}{\Delta f_{ant} + \Delta f_{post}} \cdot h$$

Primeiramente, devemos identificar a classe modal, que é aquela cuja frequência absoluta é a maior. Identificaremos também as frequências absolutas das classes anterior e posterior e o limite inferior da classe modal.

Salários dos empregados da empresa XYZ em dezembro de 2005

	Salários (R$)	Frequências Simples Absolutas	
	1 000,00 ⊢ 2 000,00	2	
	2 000,00 ⊢ 3 000,00	8	classe anterior
limite inferior	3 000,00 ⊢ 4 000,00	16	
	4 000,00 ⊢ 5 000,00	10	classe posterior
	5 000,00 ⊢ 6 000,00	4	

Agora, vamos utilizar a Fórmula de Czuber:

$$Mo = L_{inf} + \frac{\Delta f_{ant}}{\Delta f_{ant} + \Delta f_{post}} \cdot h$$

O termo **h** é a amplitude da classe modal, que é igual à diferença entre R$ 4.000 e R$ 3.000, ou seja, h = R$ 1.000.

$$Mo = 3000 + \frac{16 - 8}{(16 - 8) + (16 - 10)} \cdot 1000$$

$$Mo = 3000 + \frac{8}{8 + 6} \cdot 1000$$

$$Mo = 3000 + \frac{8}{14} \cdot 1000 \cong 3000 + 571 = 3571$$

Letra e.

Para fins de treino, vamos fazer as contas da estimativa da moda para essa distribuição usando as fórmulas da Moda Bruta e da Moda de King.

A Moda Bruta toma simplesmente o ponto médio da classe modal.

$$Mo = \frac{3000 + 4000}{2} = \frac{7000}{2} = 3500$$

E, agora, vamos calcular a Moda de King.

$$Mo = L_{inf} + \frac{f_{post}}{f_{ant} + f_{post}} \cdot h$$

$$Mo = 3000 + \frac{10}{10 + 8} \cdot 1000$$

$$Mo = 3000 + \frac{10}{18} \cdot 1000 \cong 3000 + 555{,}56 \cong 3555{,}56$$

Note que, nessa distribuição em particular, não houve grandes diferenças entre as três estimativas.

Porém, mesmo assim, não se iluda. Esteja preparado(a) para a possibilidade de que os três valores estejam presentes dentre as opções de resposta.

4. Mediana

Para o cálculo da mediana, deve-se dispor em ordem crescente todos os elementos da amostra. Escolhe-se como a mediana o valor que divide a amostra em **duas partes de tamanho**

igual, de modo que o número de elementos inferiores à mediana é igual ao número de elementos superiores à mediana.

Quando a amostra é formada por uma quantidade ímpar de termos, a mediana corresponderá exatamente a um elemento da amostra. Vejamos um exemplo:

{23, 27, 29, 31, 40, 46, 48}

Essa amostra tem sete elementos, que é um número ímpar. A mediana será igual ao termo central, que é:

{23, 27, 29, **31**, 40, 46, 48}

3 elementos inferiores à mediana

3 elementos superiores à mediana

O termo **31** tem uma interessante propriedade nessa sequência: ele divide a amostra em duas partes. São três elementos inferiores à mediana e três elementos superiores à mediana.

Quando os dados são listados na forma de um rol com N termos, sendo N um número ímpar, **a mediana será sempre um elemento da amostra**, que pode ser obtido com a posição:

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

Por outro lado, quando a amostra possui uma quantidade par de termos, a mediana corresponderá à média aritmética dos dois termos centrais.

{23, 27, 29, **31, 35**, 40, 46, 48}

Os dois termos centrais estão em negrito.

$$Md = \frac{x_3 + x_4}{2} = \frac{31+35}{2} = \frac{66}{2} = 33$$

Retornando à amostra, temos que metade dos elementos da amostra são inferiores à mediana, destacados em negrito, e a outra metade é superior.

33

A mediana é a média aritmética dos dois termos centrais

{23, 27, 29, **31, 35**, 40, 46, 48}

4 elementos são inferiores à mediana

4 elementos são superiores à mediana

É muito comum em questões de provas os dados serem fornecidos na forma de um rol, porém desordenado.

{28, 27, 28, 36, 34, 51, 43, 29, 32, 38}

Nesse caso, você precisará ordená-lo para calcular a mediana.

28	27	28	36	34	51	43	29	32	38
x_2	x_1	x_3	x_7	x_6	x_{10}	x_9	x_4	x_5	x_8

Como o rol tem um número par de elementos, já sabemos que a mediana será obtida pela média aritmética de dois pontos.

$$n = \frac{1+10}{2} = \frac{11}{2} = 5{,}5$$

A mediana pode ser obtida como a média entre os termos x_5 e x_6.

$$Md = \frac{x_5 + x_6}{2} = \frac{32+34}{2} = \frac{66}{2} = 33$$

Direto do Concurso

019. (ESAF/ANAC/2016/ANALISTA ADMINISTRATIVO) Os valores a seguir representam a quantidade de aviões que decolaram por hora durante as 10 primeiras horas de certo dia.

33 34 27 30 28 26 34 23 14 31

Logo, levando em consideração somente essas 10 horas, pode-se afirmar corretamente que:

a) o número médio de aviões que decolaram por hora é igual a 27.

b) o número mediano de aviões que decolaram por hora é igual a 29.

c) em 50% das horas o número de aviões que decolaram por hora ficou abaixo da média.

d) o número mediano de aviões que decolaram por hora é igual a 27.

e) em 30% das horas o número de aviões que decolaram por hora foi superior a 30.

Primeiramente, calcularemos o número médio de aviões que decolaram por hora.

$$\mu = \frac{33+34+27+30+28+26+34+23+14+31}{10} = \frac{280}{10} = 28$$

Portanto, em 5 horas, o número de aviões ficou acima da média e, em 4 horas, ficou abaixo da média. Logo, a letra c está errada.

Por outro lado, a mediana pode ser calculada colocando a sequência em ordem crescente.

14 23 26 27 **28 30** 31 33 34 34

Tomemos os dois termos centrais e sua média:

$$Md = \frac{28+30}{2} = \frac{58}{2} = 29$$

Por fim, apenas para descartar a letra e, temos que, em 4 horas (40%), o número de aviões que decolaram foi superior a 30.

Letra b.

4.1. Mediana em Distribuições de Frequências

É bastante importante saber calcular a mediana quando os dados são fornecidos na forma de um diagrama de frequências.

Para isso, voltemos aos dados do número de filhos de diversas pessoas que já estudamos na Figura 2.

Tabela 21: número de filhos por pessoa

Número de filhos	Número de pessoas
0	4
1	5
2	7
3	3
4	0
5	1

Como a amostra tem 20 elementos, isso significa que começa no primeiro elemento (x_1) e termina no vigésimo termo (x_{20}). Devemos tomar como mediana os termos centrais. Tomamos a soma 1 + 20 sobre dois.

$$\frac{1+20}{2} = 10{,}5$$

Como não existe o elemento $x_{10,5}$, podemos pegar os elementos 10 e 11 para a mediana.

$$Md = \frac{x_{10} + x_{11}}{2}$$

Agora, vamos identificar o décimo e o décimo primeiro elementos da amostra.

Tabela 22: número de filhos de pessoas numa amostra

Número de filhos	Número de pessoas	Elementos
0	4	x_1 a x_4
1	5	x_5 a x_9
2	7	x_{10} a x_{16}
3	3	x_{17} a x_{19}
4	0	
5	1	x_{20} a x_{20}

Para saber os elementos de cada classe, começamos da primeira, que tem 4 elementos, portanto, ela terá de x_1 a x_4. Na segunda classe, temos cinco elementos. Basta somar 4 + 5 = 9 e teremos que essa classe termina em x_9.

A classe seguinte tem 7 elementos. Somando 9 + 7 = 16, concluímos que a classe terminará em x_{16}. E, assim, por diante. Somente a última classe terá um único elemento, que será o x_{20}.

Concluímos, portanto, que o décimo e o décimo primeiro elementos são iguais a 2, pois possuem dois filhos.

$$Md = \frac{x_{10} + x_{11}}{2} = \frac{2+2}{2} = \frac{4}{2} = 2$$

Direto do Concurso

020. (FCC/CNMP/2015/ANALISTA DO CNMP – ESTATÍSTICO) A tabela de frequências absolutas abaixo corresponde à distribuição dos valores dos salários dos funcionários de nível médio lotados em um órgão público no mês de dezembro de 2014.

Classe de Salários (R$)	Frequências Absolutas
1.500 ⊢ 2.500	f_1
2.500 ⊢ 3.500	f_2
3.500 ⊢ 4.500	f_3
4.500 ⊢ 5.500	f_4
5.500 ⊢ 6.500	f_5
6.500 ⊢ 7.500	f_6

Observação: $f_i = -i^2 + 10i + 1,\ 1 \le i \le 6$.

O valor da mediana destes salários, obtido pelo método da interpolação linear, é, em R$, igual a:

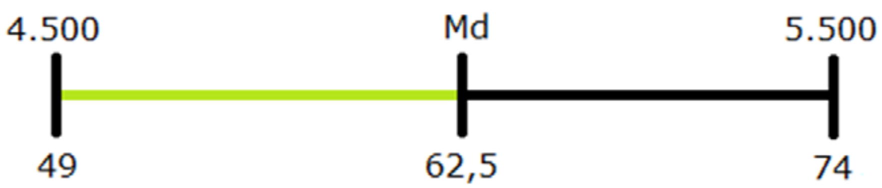
a) 5.320,00
b) 5.040,00

c) 5.260,00
d) 4.900,00
e) 5.400,00

O diferencial dessa questão é que devemos calcular as frequências absolutas pela função fornecida.

Classe de salários	Frequência absoluta	Frequência acumulada
R$1.500 a R$2.500	$-1^2+10+1 = 10$	10
R$2.500 a R$3.500	$-2^2+20+1 = 17$	27
R$3.500 a R$4.500	$-3^2+30+1=22$	**49**
R$4.500 a R$5.500	$-4^2+40+1=25$	**74**
R$5.500 a R$6.500	$-5^2+50+1 = 26$	100
R$6.500 a R$7.500	$-6^2+60+1=25$	125

Como foi fornecida uma amostra com 125 elementos, devemos tomar o $x_{62,5}$ como mediana. Foram destacadas na tabela duas classes: a classe a que pertence a mediana e a classe anterior, haja vista que a classe anterior é importante para determinar essa medida de posição. Agora, façamos a interpolação linear.



Com sabemos, o pedaço menor é proporcional ao segmento inteiro.

$$\frac{Md-4500}{62{,}5-49}=\frac{5500-4500}{74-49}\therefore\frac{Md-4500}{13{,}5}=\frac{1000}{25}=40$$

$$\therefore\ Md-4500=13{,}5.40=540\therefore\ Md=4500+540=5040$$

Letra b.

4.2. Robustez da Mediana

Assim como a moda, a mediana é uma **estimativa robusta.** Ou seja, não é sensível à presença de outliers.

Por exemplo, se adicionarmos uma pessoa que possui 20 filhos na amostra, o que acontecerá?

Tabela 23: adição de um outlier para o cálculo da mediana

Número de filhos	Número de pessoas	Elementos
0	4	x_1 a x_4
1	5	x_5 a x_9
2	7	x_{10} a x_{16}
3	3	x_{17} a x_{19}
4	0	
5	1	x_{20}
20	1	x_{21}

Ao colocar um novo elemento, o tamanho da amostra passou a 21 elementos. Dessa maneira, a mediana será calculada por:

$$\frac{1+21}{2} = \frac{22}{2} = 11$$

$$Md = x_{11} = 2$$

Perceba que a mediana não se alterou, mesmo com a adição de um outlier. Vamos reunir as três amostras.

Tabela 24: comparação dos efeitos da adição de um outlier à média, moda e mediana

Medida	Amostra normal	Amostra com outlier
Média	1,65	2,65
Moda	2	2
Mediana	2	2

Anote a conclusão a que acabamos de chegar.

A média é uma estimativa sensível à presença de outliers.
Porém, a mediana e a moda são estimativas robustas.

Devido à robustez, a mediana é muito utilizada em substituição à média em situações em que a presença de outliers é muito comum. Por exemplo, na avaliação de patrimônio.

Na minha consultoria de investimentos, não faria sentido falar em patrimônio médio dos clientes. Só por curiosidade, eu tenho 3 clientes, cujo patrimônio é superior à soma de todos os demais. Logo, a média seria sensivelmente alterada por esses três indivíduos.

Faz muito mais sentido falar em patrimônio mediano dos clientes, a fim de evitar o problema dos outliers.

4.3. Mediana em Dados Categorizados

Como já estudamos, as variáveis contínuas normalmente são apresentadas na forma de dados categorizados. Ou seja, elas são fornecidas em faixas de valores, não apenas valores únicos.

Por esse motivo, não é possível calcular a mediana seguindo a técnica aprendida anteriormente. Em vez disso, utiliza-se a técnica da interpolação linear.

Vejamos um exemplo com as notas de alunos em uma prova discursiva.

Tabela 25: exemplo de dados categorizados

Nota	Número de alunos (frequência absoluta)
0 a 1	20
1 a 2,5	16
2,5 a 4	5
4 a 5	1

Porém, para o cálculo da mediana, adotaremos um procedimento ligeiramente diferente. O total de alunos nessa classe é 42. Tomaremos como mediana o elemento que ocupa exatamente a posição definida como 42/2. **Não** se soma 1 nesse caso.

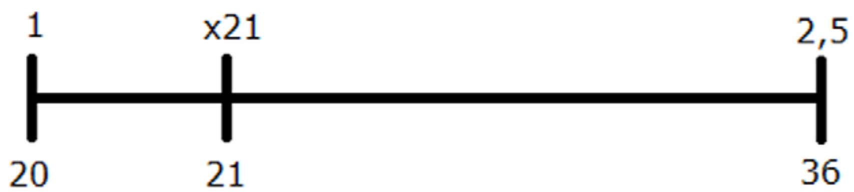
$$Md = x_{42/2} = x_{21}$$

Agora, devemos localizar a **classe mediana,** que é aquela em que está localizado o elemento 21.

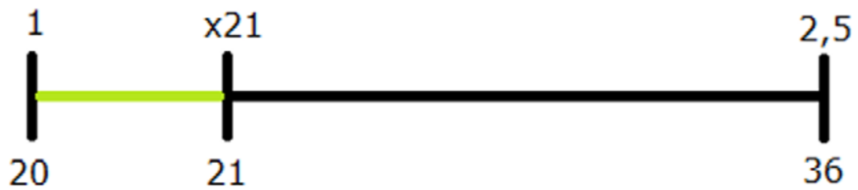
Nota	Número de alunos	Frequência acumulada	Elementos
0 a 1	20	**20**	1 a **20**
1 a 2,5	16	**36**	21 a **36**
2,5 a 4	5	41	37 a 41
4 a 5	1	42	42

Tomaremos os pontos extremos da segunda classe (onde está localizada a mediana) e da classe anterior.

Definiremos que o elemento x_{20} será o ponto extremo da primeira classe e que o elemento x_{36} será o extremo da segunda classe. E faremos uma interpolação linear para encontrar o elemento x_{21}.



Destacaremos a região entre o x_{21} e o início da interpolação linear.



A parte à esquerda (que vai do 20 ao 21) deve ser proporcional a todo o trecho preto. Sendo assim, temos:

$$\frac{x_{21} - 1}{21 - 20} = \frac{2{,}5 - 1}{36 - 20}$$

Perceba que tomamos as medidas 20 a 21 do lado esquerdo e toda a medida de 20 a 36 no lado direito da proporção.

No numerador da fração, estão as medidas da parte superior, ou seja, o valor da variável X, enquanto no denominador da fração estão as medidas da parte inferior, ou seja, a posição ocupada pela medida.

$$\frac{x_{21}-1}{1}=\frac{1,5}{16}\therefore x_{21}-1=0,09375$$

$$\therefore x_{21}=1+0,09375=1,09375$$

A interpolação linear pode parecer um pouco estranha, porém o procedimento será o mesmo para todas as questões. É só uma questão de treino.

Direto do Concurso

021. (FCC/SEFAZ-SC/2018/AUDITOR-FISCAL DA RECEITA ESTADUAL – AUDITORIA E FISCALIZAÇÃO (PROVA 1)) A tabela a seguir apresenta a distribuição de frequências dos salários, em número de salários mínimos (SM), dos funcionários de um órgão público:

Faixa salarial (SM)	Porcentagem
2 ⊢—— 4	a
4 ⊢—— 6	a + 20
6 ⊢—— 8	b
8 ⊢—— 12	b – 10

Sabe-se que: b – a = 5%,
$\bar{x}$ é a média salarial, obtida por meio dessa tabela, calculada como se todos os valores de cada faixa salarial coincidissem com o ponto médio da referida faixa, md é a mediana salarial, calculada por meio dessa tabela pelo método da interpolação linear.
Nessas condições, $\bar{x}$ + md, em anos, é igual a:
a) 9,85
b) 11,35
c) 11,05
d) 10,95
e) 11,65

Sabemos que b – a = 5%. Dessa forma, temos que:

$$b-a=5\%\therefore b=a+5\%$$

Para calcular o valor de **a**, vamos utilizar o fato de que a soma das porcentagens é igual a 100%.

$$a+(a+20\%)+b+(b-10\%)=100\%$$

$$a+(a+20\%)+(a+5\%)+(a+5\%-10\%)=100\%$$

$$a+(a+20\%)+(a+5\%)+(a-5\%)=100\%$$

$$4a + 20\% = 100\%$$

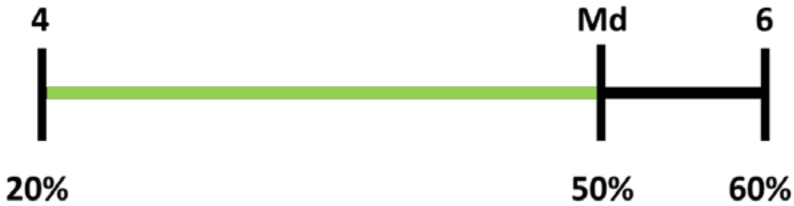
$$\therefore 4a = 100\% - 20\% = 80\%$$

$$\therefore a = \frac{80\%}{4} = 20\%$$

De posse do valor de **a**, podemos montar a tabela de distribuição de porcentagem de cada uma das faixas salariais.

Faixa salarial (SM)	Porcentagem	Porcentagem acumulada
2 a 4	20%	0% a 20%
4 a 6	**40%**	**20% a 60%**
6 a 8	25%	60% a 85%
8 a 12	15%	85% a 100%

Para calcular a mediana, primeiramente devemos identificar a classe mediana, que está em destaque, pois é a classe em que a porcentagem acumulada atinge 50%.
Podemos calcular a mediana por interpolação linear exatamente no momento em que a porcentagem acumulada atinge 50%.



$$\frac{Md - 4}{50\% - 20\%} = \frac{6 - 4}{60\% - 20\%}$$

$$\frac{Md - 4}{30\%} = \frac{2}{40\%} \therefore Md - 4 = \frac{30\%}{40\%} \cdot 2 = \frac{3}{4} \cdot 2 = 1{,}5$$

$$\therefore Md = 4 + 1{,}5 = 5{,}5$$

Já a média pode ser calculada pela técnica do ponto médio.

Faixa salarial (SM)	Ponto médio	Porcentagem
2 a 4	**3**	20%
4 a 6	**5**	40%

Faixa salarial (SM)	Ponto médio	Porcentagem
6 a 8	**7**	25%
8 a 12	**10**	15%

Usando o ponto médio, temos:

$$\bar{X} = 0{,}20.3 + 0{,}40.5 + 0{,}25.7 + 0{,}15.10$$

$$\bar{X} = 0{,}6 + 2 + 1{,}75 + 1{,}5 = 5{,}85$$

Portanto, a soma pedida no enunciado da média com a mediana é:

$$\bar{X} + Md = 5{,}85 + 5{,}5 = 11{,}35$$

Letra b.

4.4. Relação de Pearson

Em uma distribuição em dados categorizados, a moda pode ser estimada pela Relação de Pearson, que relaciona a moda (Mo), mediana (Md) e média (Me).

Segundo o estatístico Pearson, a diferença entre a média e moda é igual ao triplo da diferença entre a média e mediana.

$$(Me - Mo) = 3.(Me - Md)$$

$$Me - Mo = 3Me - 3Md$$

$$\therefore Me - 3Me + 3Md = Mo$$

$$\therefore Mo = 3Md - 2Me$$

Considerando o caso anterior, já vimos que a média pode ser calculada com o ponto médio de cada uma das classes. Vejamos o exemplo.

Nota	Ponto médio	Número de alunos
0 a 1	0,5	20
1 a 2,5	1,75	16
2,5 a 4	3,25	5
4 a 5	4,5	1

Portanto, podemos calcular a média como:

$$\mu = Me = \frac{20.0{,}5 + 16.1{,}75 + 5.3{,}25 + 1.4{,}5}{20 + 16 + 5 + 1} = \frac{58{,}75}{42} = 1{,}39881$$

Substituindo os valores calculados para a média e a mediana na Relação de Pearson, temos:

$$Mo = 3.1{,}09375 - 2.1{,}39881 \cong 0{,}48$$

Uma maneira relativamente intuitiva de calcular a média seria considerar o ponto médio da classe mais frequente. Nesse exemplo, a classe mais frequente é 0 a 1, por isso a moda estimada dessa maneira seria 0,5.

Com isso, vemos que a relação de Pearson é bastante coerente com essa forma intuitiva de calcular a moda.

Direto do Concurso

022. (FCC/INFRAERO/2011/ESTATÍSTICO) A tabela de frequências relativas abaixo corresponde à distribuição da renda mensal das pessoas que adquiriram pacotes de excursão de uma empresa de turismo em 2010. O valor da média aritmética da renda (Me) foi obtido considerando que todos os valores incluídos num certo intervalo de classe são coincidentes com o ponto médio deste intervalo. O valor da mediana (Md) foi obtido pelo método da interpolação linear.

Renda (R$)	Frequência Relativa
2.500 ⊢ 3.500	K
3.500 ⊢ 4.500	2K + 0,125
4.500 ⊢ 5.500	3K + 0,150
5.500 ⊢ 6.500	4K + 0,075
6.500 ⊢ 7.500	5K − 0,100
Total	1,000

O valor da moda (Mo), obtido pela relação de Pearson: Mo = 3Md - 2Me, é igual a:

a) R$4.250,00
b) R$4.750,00
c) R$5.000,00
d) R$5.250,00

O primeiro passo para resolver a questão é lembrar que a soma das frequências relativas deve ser sempre igual a 100%. Dessa maneira, temos que:

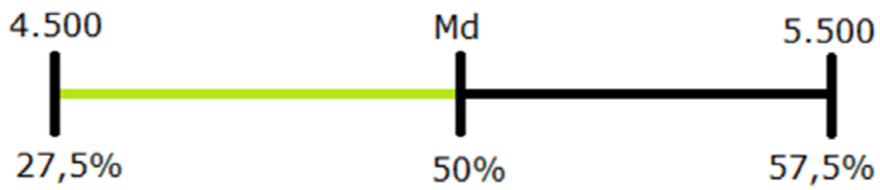
$$K + 2K + 0{,}125 + 3K + 0{,}150 + 4K + 0{,}075 + 5K - 0{,}100 = 1$$

$$15K + 0{,}25 = 1 \therefore 15K = 1 - 0{,}25 = 0{,}75 \therefore K = \frac{0{,}75}{15} = 0{,}05$$

Com base no valor de K calculado, podemos montar a tabela e determinar todas as frequências relativas envolvidas.

Renda	Ponto médio	Frequência relativa	Frequência acumulada
R$2.500 a R$3.500	R$3.00	5%	5%
R$3.500 a R$4.500	R$4.000	22,5%	27,5%
R$4.500 a R$5.500	R$5.000	30%	57,5%
R$5.500 a R$6.500	R$6.000	27,5%	85%
R$6.500 a R$7.500	R$7.000	15%	100%

Mais uma vez, destacamos os pontos relevantes para o cálculo da interpolação linear.



$$\frac{Md - 4500}{50\% - 27{,}5\%} = \frac{5500 - 4500}{57{,}5\% - 27{,}5\%} \therefore \frac{Md - 4500}{22{,}5\%} = \frac{1000}{30\%}$$

$$Md - 4500 = 1000.\frac{22{,}5\%}{30\%} = \frac{1000.3}{4} = 750 \therefore Md - 4500 = 750$$

$$\therefore Md = 4500 + 750 = 5250$$

Agora, calcularemos a média utilizando os pontos médios.

$$Me = 0{,}05.3000 + 0{,}225.4000 + 0{,}3.5000 + 0{,}275.6000 + 0{,}15.7000$$

$$= 150 + 900 + 1500 + 1650 + 1050 = 5250$$

Por fim, utilizando a relação de Pearson, temos que a moda é dada por:

$$Mo = 3Md - 2Me = 3.5250 - 2.5250 = 5250$$

Letra d.

RESUMO

Vamos agora resumir os principais conceitos apresentados nesse material.

Dados Categorizados x Dados em Tabela

Um dado em categoria quando os valores da variável aleatória são fornecidos na forma de uma faixa de valores, não na forma de um valor específico. Por exemplo, o salário é de 1 a 3 salários mínimos.

No caso de dados em categorias, deve-se utilizar:

- **média aritmética:** técnica do ponto médio;
- **mediana:** interpolação linear;
- **moda:** fórmula da Moda Bruta, de Czuber, de King ou de Pearson.

Média Aritmética

Modalidade	Fórmula	Conceito
Para dados em rol	$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$	Soma tudo e divide pela quantidade de termos.
Para dados em tabela	$\mu = \frac{\sum_{i=1}^{N} f_i x_i}{\sum_{i=1}^{N} f_i}$	Utilizar as frequências como pesos.
Para dados categorizados	$\mu = \frac{\sum_{i=1}^{N} f_i \bar{x}_i}{\sum_{i=1}^{N} f_i}$	Utilizar o ponto médio de cada categoria.

Mediana

Modalidade	Termo central	Fórmula	Conceito
Para dados em rol ou em Tabela N = ímpar	$n = \frac{N+1}{2}$	$Md = x_n$	Tomamos o termo central.
Para dados em rol ou em Tabela N = par	$n = \frac{N}{2}$	$Md = \frac{x_n + x_{n+1}}{2}$	Tomamos a média aritmética dos dois termos centrais.
Para dados categorizados	$n = \frac{N}{2}$	$Md = x_n$	Tomamos o termo $x_{N/2}$ calculado por interpolação linear

Moda

Modalidade	Conceito
Para dados em rol ou em tabela	Tomamos o termo mais frequente.
Para dados categorizados	Utilizamos Moda Bruta, Czuber, King ou Pearson.

- A Moda Bruta toma o ponto médio da classe modal.
- A Relação de Pearson estabelece que a distância entre a média e mediana é o triplo da distância entre a média e a moda.

$$Me - Mo = 3.(Me - Md)$$

$$Me - Mo = 3Me - 3Md$$

$$\therefore Mo = 3Md - 2Me$$

- A Moda de Czuber e a de King tomam como referência o limite inferior e a amplitude da classe modal.

$$Mo = \overset{\text{limite inferior}}{\overset{\downarrow}{Linf}} + \left(\frac{\quad}{\quad}\right) \cdot \overset{\text{amplitude}}{\overset{\downarrow}{h}}$$

- A Fórmula de Czuber leva em consideração as diferenças de frequências entre a classe modal e as duas vizinhas.

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

Quanto maior a frequência da classe anterior, menor é o valor da moda. Por isso, Δf_{ant} está no denominador. Quanto maior f_{ant}, menor será o valor de $\Delta f_{ant} = f - f_{ant}$, que é a diferença entre a frequência da classe modal e da classe anterior.

- A Fórmula de King leva em consideração somente as frequências da classe anterior e da classe posterior à classe modal.

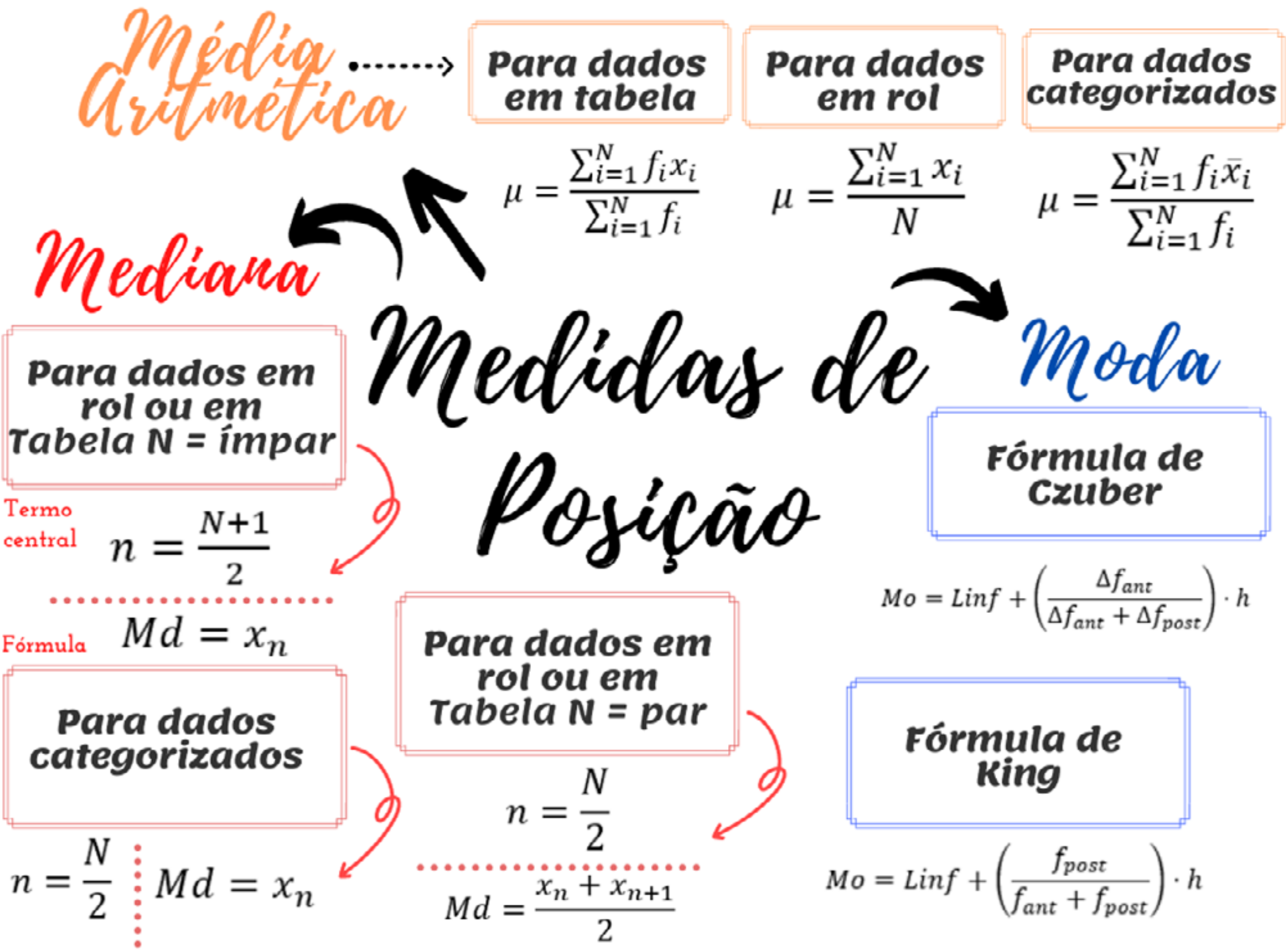
$$Mo = Linf + \left(\frac{f_{post}}{f_{ant} + f_{post}}\right) \cdot h$$

Perceba que inverte. Na Moda de Czuber, temos Δf_{ant} no numerador. Na Moda de King, temos f_{post} no numerador.

Robustez x Sensibilidade

A média aritmética é sensível a valores atípicos (outliers). Já a mediana e a moda são estimativas robustas.

MAPA MENTAL


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 categorizados
$n = \frac{N}{2}$
$Md = x_n$
Para dados em rol ou em Tabela N = par
$n = \frac{N}{2}$
$Md = \frac{x_n + x_{n+1}}{2}$
Moda
Fórmula de Czuber
$Mo = Linf + \left(\frac{\Delta f_{ant}}{\Delta f_{ant} + \Delta f_{post}}\right) \cdot h$
Fórmula de King
$Mo = Linf + \left(\frac{f_{post}}{f_{ant} + f_{post}}\right) \cdot h$

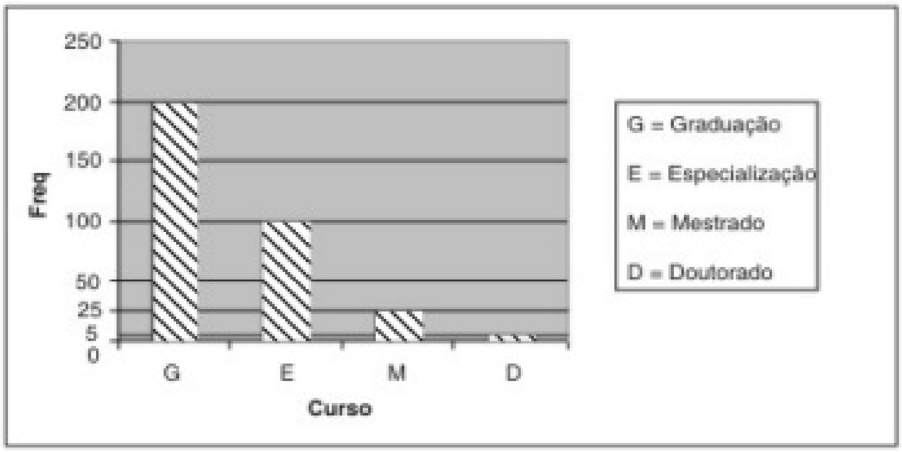
QUESTÕES COMENTADAS EM AULA

001. (CESPE/SEDUC-AM/2011/ESTATÍSTICO)

RG(*)	gênero	grau de instrução	hiperatividade
35805684	M	3	N
21355706	F	5	N
43674475	M	2	S
2305386	F	3	N
⋮	⋮	⋮	⋮
98814652	M	4	N

(*) números fictícios gerados por computador.

A tabela acima contém um conjunto de dados formado por quatro variáveis: RG; gênero (M = masculino; F = feminino); grau de instrução (1 = analfabeto; 2 = fundamental incompleto; 3 = fundamental completo; 4 = médio incompleto; 5 = médio completo ou superior); e hiperatividade (S = sim; N = não). Com base nessa tabela, julgue o item.
As variáveis mostradas na tabela são qualitativas.



002. (CESPE/SEDUC-AM/2011/ESTATÍSTICO) A qualificação dos professores é de grande importância para a qualidade da formação dos estudantes. Considerando que a figura acima apresenta a distribuição do número de professores em uma faculdade, segundo a formação acadêmica (curso), julgue o item.
A variável curso é qualitativa nominal.

003. (UFU-MG/2019/TÉCNICO EM ESTATÍSTICA) Considere as seguintes variáveis.
I – Tamanho de um objeto (pequeno, médio ou grande)
II – Volume de água em um rio
III – Número de clientes numa fila

IV – Número da seção de votação
V – Comprimento de um inseto
VI – Classe Social
Com relação à classificação dos dados requeridos como variáveis de pesquisa, é correto afirmar que:
a) as variáveis I, IV e VI são qualitativas.
b) as variáveis III e V são quantitativas contínuas.
c) as variáveis II e III são quantitativas discretas.
d) a variável IV é qualitativa ordinal.

004. (CESPE/TJ-PA/2020/ANALISTA JUDICIÁRIO – ESTATÍSTICA) Uma amostra aleatória dos registros de furto no município de Abaetetuba, no ano de 2017, apresenta os valores 245, 247, 238, 282 e 261. Uma estimativa não tendenciosa e eficiente para a média de furtos ocorridos em Abaetetuba no ano de 2017, considerando os dados apresentados na amostra, é
a) 238,0.
b) 254,6.
c) 260,0.
d) 282,7.
e) 308,5.

005. (FCC/CÂMARA DE FORTALEZA – CE/2019/AGENTE ADMINISTRATIVO) Em um teatro com 200 lugares, houve quatro apresentações de uma peça. Na primeira apresentação foram vendidos todos os ingressos; na segunda apresentação foram vendidos 88% dos ingressos; na terceira, 56% dos ingressos e, na quarta, 44% dos ingressos. Em média, a quantidade de ingressos vendidos por apresentação foi de
a) 72
b) 144
c) 56
d) 76
e) 140

006. (FCC/SABESP/2018/ESTAGIÁRIO DE NÍVEL MÉDIO) Para que a média aritmética dos números: 8, 8, 1, 10, 11, 12, 7, 2, 10, 6, **x** e 5 seja 7, o valor de **x** deverá ser:
a) 2
b) 4
c) 8
d) 5
e) 3

007. (CESPE/PREFEITURA DE SÃO CRISTÓVÃO – SE/PROFESSOR DE MATEMÁTICA) Segundo o IBGE, a massa da renda média mensal real domiciliar per capita em 2016 foi de aproximadamente R$ 264 bilhões; a população brasileira nesse ano era de aproximadamente 190 milhões de pessoas.
A partir dessas informações, julgue o item a seguir.
A renda média mensal dos brasileiros em 2016 foi superior a R$ 1.300.

008. (FCC/PREFEITURA DE TERESINA – PI) Uma carteira aplica 25% na ação A, 40% na ação B e o restante na ação C. Os retornos das ações A, B e C são, respectivamente, 10%, 12% e 20%. O retorno médio da carteira será
a) 14,5%.
b) 14,8%.
c) 14,6%.
d) 14,0%.
e) 14,3%.

009. (FCC/SPAG-PE/2019/ANALISTA DE PLANEJAMENTO/ADAPTADA) Durante 40 dias, foi registrado o número de pessoas atendidas por dia em um guichê de uma repartição. A tabela abaixo apresentou os dados observados sendo que não foram fornecidas as quantidades de dias em que foram atendidas uma pessoa por dia e duas pessoas por dia, indicadas na tabela por $q_1 = 11$ e $q_2 = 14$, respectivamente.

Número de pessoas atendidas	Quantidade de dias
0	9
1	q_1
2	q_2
3	5
4	1
Total	40

Tem-se que a média aritmética do número de pessoas atendidas por dia é:
a) 1,40
b) 1,50
c) 1,25
d) 1,60
e) 1,45

010. (FCC/TCE-PR/2011/ANALISTA DE CONTROLE) A média dos salários, calculada supondo-se que todos os valores dentro de uma faixa salarial tenham seus valores iguais ao ponto médio desta faixa, em número de salários mínimos, é igual a:

Faixa salarial (em número de salários mínimos)	Frequência Absoluta
1 |— 3	200
3 |— 5	400
5 |— 7	200
7 |— 9	200

a) 4,2
b) 4,5
c) 4,6
d) 4,8
e) 5,0

011. (IFPA/2019/ESTATÍSTICO) Uma variável aleatória discreta Z tem função de probabilidade dada por:

Z	-2	-1	0	1	2
P(Z)	0,3	0,1	0,2	0,3	0,5

Pode-se afirmar que a média da variável aleatória Z é igual a:
a) 0,9
b) 0,6
c) -0,6
d) -0,9
e) 1

012. (CESPE/SEDUC-AL/2018/PROFESSOR DE MATEMÁTICA) **Situação hipotética**: A média aritmética dos pesos dos 60 alunos de uma sala de aulas é igual a 51,8 kg. Nessa sala, a média aritmética do peso dos meninos é de 62 kg e das meninas, 45 kg. **Assertiva**: Nesse caso, essa sala de aulas tem 24 meninos e 36 meninas.

013. (CESPE/ANP/2013/ESPECIALISTA EM REGULAÇÃO) Considere duas amostras aleatórias de tamanhos diferentes, em que a soma dos valores observados sejam S_1 e S_2 e as respectivas médias $\overline{x_1}$ e $\overline{x_2}$. Nesse caso, a média aritmética global, ou seja, a média dos valores observados nas duas amostras é igual à média das médias $\overline{x_1}$ e $\overline{x_2}$.

014. (FCC/SEFAZ-GO/2018/AUDITOR-FISCAL) Os matemática definem diferentes tipos de médias entre dois números positivos e, para cada aplicação, escolhem qual o tipo mais adequado a ser utilizado. A média harmônica H entre os números positivos a e b, por exemplo, é definida como o inverso da média aritmética dos inversos desses números, ou seja,

$$H = \frac{\frac{1}{a} + \frac{1}{b}}{2}$$

A média aritmética dos números 5 e 20 supera a média harmônica desses mesmos números em
a) 4 unidades.
b) 4,25 unidades.
c) 4,5 unidades.
d) 4,75 unidades.
e) 5 unidades.

015. (CEPERJ/SEFAZ-RJ/2013/ANALISTA DE CONTROLE INTERNO) Um filme foi exibido em um cinema em 8 diferentes sessões, ao longo de todo o dia. O número de presentes em cada sessão é mostrado na tabela abaixo:

Sessão	Número de presentes	Sessão	Número de presentes
1	88	5	94
2	102	6	82
3	90	7	80
4	76	8	68

O número médio de pessoas por sessão corresponde a:
a) 68
b) 72
c) 76
d) 81
e) 85

016. (VUNESP/PREFEITURA DE SÃO JOSÉ DOS CAMPOS/2012/AUDITOR TRIBUTÁRIO MUNICIPAL – GESTÃO TRIBUTÁRIA) A média aritmética de alturas de 10 alunos de um time de futebol é 175 cm. Dois novos alunos entram para o time, e a nova média de alturas passa a ser 178 cm. Se a diferença entre as alturas desses dois novos jogadores é 6 cm, o maior dos dois mede, em cm,
a) 188.
b) 190.
c) 192.
d) 194.
e) 196.

017. (FCC/ALESE/2018/ANALISTA LEGISLATIVO – ECONOMIA/ADAPTADA) Em um grupo de pessoas encontramos as seguintes idades: 20, 30, 50, 39, 20, 25, 41, 47, 36, 45, 41, 52, 18, 41. A moda é:

a) 42.
b) 45.
c) 41.
d) 20.
e) 39.

018. (FCC/BACEN/2006) O valor da moda, obtida com a utilização da Fórmula de Czuber, é igual a (desprezar os centavos na resposta):

Salários dos empregados da empresa XYZ em dezembro de 2005

Salários (R$)	Freqüências Simples Absolutas
1 000,00 ⊢ 2 000,00	2
2 000,00 ⊢ 3 000,00	8
3 000,00 ⊢ 4 000,00	16
4 000,00 ⊢ 5 000,00	10
5 000,00 ⊢ 6 000,00	4

a) R$3.201,00
b) R$3.307,00
c) R$3.404,00
d) R$3.483,00
e) R$3.571,00

019. (ESAF/ANAC/2016/ANALISTA ADMINISTRATIVO) Os valores a seguir representam a quantidade de aviões que decolaram por hora durante as 10 primeiras horas de certo dia.
33 34 27 30 28 26 34 23 14 31
Logo, levando em consideração somente essas 10 horas, pode-se afirmar corretamente que:
a) o número médio de aviões que decolaram por hora é igual a 27.
b) o número mediano de aviões que decolaram por hora é igual a 29.
c) em 50% das horas o número de aviões que decolaram por hora ficou abaixo da média.
d) o número mediano de aviões que decolaram por hora é igual a 27.
e) em 30% das horas o número de aviões que decolaram por hora foi superior a 30.

020. (FCC/CNMP/2015/ANALISTA DO CNMP – ESTATÍSTICO) A tabela de frequências absolutas abaixo corresponde à distribuição dos valores dos salários dos funcionários de nível médio lotados em um órgão público no mês de dezembro de 2014.

Classe de Salários (R$)	Frequências Absolutas
1.500 ⊣ 2.500	f_1
2.500 ⊣ 3.500	f_2
3.500 ⊣ 4.500	f_3
4.500 ⊣ 5.500	f_4
5.500 ⊣ 6.500	f_5
6.500 ⊣ 7.500	f_6

Observação: $f_i = -i^2 + 10i + 1,\ 1 \leq i \leq 6$.

O valor da mediana destes salários, obtido pelo método da interpolação linear, é, em R$, igual a:

a) 5.320,00

b) 5.040,00

c) 5.260,00

d) 4.900,00

e) 5.400,00

021. (FCC/SEFAZ-SC/2018/AUDITOR-FISCAL DA RECEITA ESTADUAL – AUDITORIA E FISCALIZAÇÃO (PROVA 1)) A tabela a seguir apresenta a distribuição de frequências dos salários, em número de salários mínimos (SM), dos funcionários de um órgão público:

Faixa salarial (SM)	Porcentagem
2 ⊢ 4	a
4 ⊢ 6	a + 20
6 ⊢ 8	b
8 ⊢ 12	b – 10

Sabe-se que: b – a = 5%,

$\overline{x}$ é a média salarial, obtida por meio dessa tabela, calculada como se todos os valores de cada faixa salarial coincidissem com o ponto médio da referida faixa, md é a mediana salarial, calculada por meio dessa tabela pelo método da interpolação linear.

Nessas condições, $\overline{x}$ + md, em anos, é igual a:

a) 9,85

b) 11,35

c) 11,05

d) 10,95

e) 11,65

022. (FCC/INFRAERO/2011/ESTATÍSTICO) A tabela de frequências relativas abaixo corresponde à distribuição da renda mensal das pessoas que adquiriram pacotes de excursão de uma empresa de turismo em 2010. O valor da média aritmética da renda (Me) foi obtido considerando que todos os valores incluídos num certo intervalo de classe são coincidentes com o ponto médio deste intervalo. O valor da mediana (Md) foi obtido pelo método da interpolação linear.

Renda (R$)	Frequência Relativa
2.500 ⊢ 3.500	K
3.500 ⊢ 4.500	2K + 0,125
4.500 ⊢ 5.500	3K + 0,150
5.500 ⊢ 6.500	4K + 0,075
6.500 ⊢ 7.500	5K − 0,100
Total	1,000

O valor da moda (Mo), obtido pela relação de Pearson: Mo = 3Md - 2Me, é igual a:

a) R$4.250,00

b) R$4.750,00

c) R$5.000,00

d) R$5.250,00

QUESTÕES DE CONCURSO

Questões Adicionais Nível 1

023. (FCC/TRF 3ª REGIÃO/2019/ANALISTA JUDICIÁRIO – INFORMÁTICA) Havia cinco garrafas de vinhos em uma adega. O preço médio desses vinhos era R$ 120,00. Uma garrafa desapareceu e o preço médio das quatro garrafas que sobraram passou para R$ 110,00. O valor, em reais, da garrafa que desapareceu é

a) 130,00.
b) 120,00.
c) 150,00.
d) 140,00.
e) 160,00.

O preço médio pode ser obtido como a razão entre a soma dos preços dos vinhos e a quantidade de vinhos presentes na adega.

$$\mu = \frac{S}{N} \therefore S = \mu N$$

Quando havia 5 garrafas, o preço médio era R$ 120. Portanto, a soma dos preços das 5 garrafas é:

$$S = \mu N = 5.120 = 600$$

Após retirar um único vinho de preço P, a soma total dos preços das garrafas fica 600 – P. E essa nova soma será igual ao produto da nova média pela nova quantidade de garrafas – que é 4, haja vista que uma garrafa foi retirada.

$$S' = \mu' N'$$

$$600 - P = 110.4$$

$$600 - P = 440$$

$$\therefore P = 600 - 440 = 160$$

Letra e.

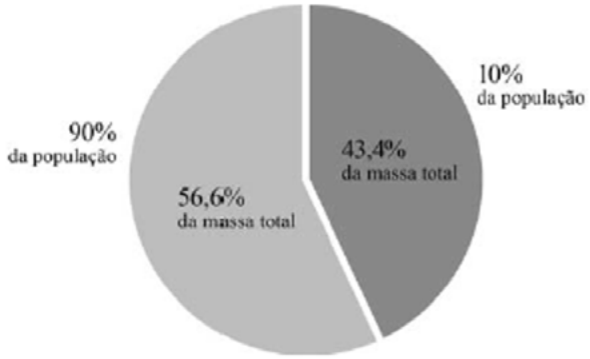
---

024. (CESPE/PREFEITURA DE SÃO CRISTÓVÃO/2019/PROFESSOR DE MATEMÁTICA) Segundo o IBGE, a massa da renda média mensal real domiciliar *per capita* em 2016 foi de aproximadamente R$ 264 bilhões; a população brasileira nesse ano era de aproximadamente 190 milhões de pessoas.

A partir dessas informações, julgue o item a seguir.

O gráfico a seguir mostra que, em 2016, mais de 40% da massa de renda mensal real domiciliar *per capita* coube a 10% da população; ao restante coube menos de 60% dessa massa de renda. A partir do gráfico, é correto inferir que, naquele ano, em média, a renda mensal desses 10% da população era superior a R$ 10.000.

PNAD-C | distribuição da massa de rendimento mensal real domiciliar *per capita*
Brasil - 2016



De acordo com as informações do enunciado, uma fatia da população correspondia a 10% do total de pessoas e a 43,4% da massa total de renda.

Com base nisso, vamos calcular a massa total da renda desse grupo.

$$S = 0{,}434.264 = 114{,}58\ bi$$

O número de pessoas pertencentes a essa classe pode ser obtido como 10% do total da população.

$$N = 0{,}10.190\ mi = 19\ mi$$

Portanto, a massa de renda de R$ 114,58 bilhões está distribuída entre um grupo de 19 milhões de pessoas. Logo, a renda média desse grupo é obtida como a razão entre a soma total das rendas e o número de pessoas pertencentes ao grupo.

Vamos escrever a renda de R$ 114,58 bilhões como R$ 114.580 milhões.

$$\mu = \frac{S}{N} = \frac{114580}{19} \cong 6030 < 10000$$

Portanto, a renda média do grupo é inferior a R$ 10.000.

Errado.

025. (FCC/SABESP/2019/ESTAGIÁRIO DE ENSINO MÉDIO REGULAR) A média dos salários dos 25 trabalhadores de uma pequena empresa é de R$ 2.320,00. Um desses trabalhadores, e apenas ele, terá um aumento de 10% em seu salário e, com isso, a média dos salários passará a ser R$ 2.360,00. O salário desse trabalhador, sem o aumento, é

R$ 10.400,00
b) R$ 9.800,00
c) R$ 10.000,00
d) R$ 8.000,00
e) R$ 11.000,00

A soma inicial dos salários dos 25 trabalhadores pode ser obtida como o produto entre a média salarial e o número de trabalhadores.

$$\mu = \frac{S}{N} \therefore S = \mu N$$

Fazendo as contas, temos:

$$S = \mu N = 2320.25 = 58000$$

Depois de o funcionário ter recebido aumento, a soma passou a ser superior – denotaremos por S’ – e ela será igual à nova média multiplicada pela quantidade de trabalhadores.

$$S' = \mu' N$$

$$S' = 2360.25 = 59000$$

Como o aumento salarial foi recebido por um único funcionário, a variação da folha de pagamentos corresponde exatamente ao aumento por ele recebido.

$$Aumento = S' - S = 59000 - 58000 = 1000$$

Sabemos, ainda, que esse aumento corresponde a 10% do salário inicial do trabalhador.

$$0{,}10.Salário = 1000$$

$$\therefore Salário = \frac{1000}{0{,}10} = 10000$$

Portanto, o seu salário inicial era de R$ 10.000 e ele recebeu um aumento de 10% (ou seja, R$ 1.000), passando a receber R$ 11.000 de salário. Como a questão pediu o salário sem o aumento, a resposta é R$ 10.000.
Letra c.

---

026. (CESPE/PREFEITURA DE SÃO CRISTÓVÃO – SE/2019/PROFESSOR DE MATEMÁTICA) Há cinco anos, João, Paulo e Miguel se associaram para montar uma lanchonete. João entrou com R$ 80.000; Paulo, com R$ 120.000; e Miguel, com R$ 200.000. A lanchonete foi vendida, hoje, por R$ 3.200.000 e essa quantia foi dividida entre os três de forma diretamente proporcional aos valores que cada um investiu.

A partir dessa situação hipotética, julgue o item a seguir.
Considerando o lucro obtido com a venda, é correto inferir que, enquanto na propriedade dos três, a lanchonete teve uma valorização média anual inferior a R$ 600.000.

Tomemos o valor inicial investido pelos quatro sócios, que corresponde à soma:

$$Inv = 80\ 000 + 120\ 000 + 200\ 000 = 400\ 000$$

A valorização pode ser obtida como a diferença entre o quanto os sócios receberam pela venda da lanchonete e o investimento inicial.

$$Val = 3\ 200\ 000 - 400\ 000 = 2\ 800\ 000$$

Agora, podemos calcular a valorização média anual, que pode ser obtida como a razão entre a valorização observada nos 5 anos e o total dos 5 anos.

$$Val = \frac{2\ 800\ 000}{5} = 560\ 000$$

Certo.

027. (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.
A mediana das quantidades *X* observadas na amostra em questão foi igual a 18 kg.

Como a amostra tem cinco unidades, para obter a mediana, devemos procurar pela posição intermediária entre a inicial e a final, ou seja:

$$\frac{1+5}{2} = \frac{6}{2} = 3$$

Portanto, a mediana equivale ao terceiro termo da sequência, quando organizada por ordem crescente. Vamos organizar as unidades amostrais de X.

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

Portanto, a mediana é igual a x_3 = 22.

Errado.

028. (CESPE/IPHAN/2018/ANALISTA I – ÁREA 2) Uma pesquisa a respeito das quantidades de teatros em cada uma de 11 cidades brasileiras selecionadas apresentou o seguinte resultado: {1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4}.

Com referência a esses dados, julgue o item seguinte.

A mediana do conjunto é igual a 3.

Como a amostra tem 11 elementos, devemos tomar como mediana o elemento que ocupa a posição:

$$\frac{1+11}{2} = \frac{12}{2} = 6$$

Agora, vamos associar cada elemento à sua posição na sequência.

1	2	2	3	3	3	4	4	4	4	4
$\mathbf{x_1}$	x_2	x_3	x_4	x_5	$\mathbf{x_6}$	x_7	x_8	x_9	x_{10}	x_{11}

Portanto, a mediana é exatamente igual ao $x_{6,}$ que é igual a 3.

Certo.

029. (CESPE/IPHAN/2018/ANALISTA I – ÁREA 2) A moda é o valor que apresenta a maior frequência da variável entre os valores observados.

É exatamente essa a definição de moda. É o valor mais frequente observado para a variável em estudo.

Certo.

030. (CESPE/IPHAN/2018/ANALISTA I – ÁREA 2) A mediana é o valor que ocupa a posição central da série de observações de uma variável, dividindo-se o conjunto de valores ordenados em partes assimétricas desiguais.

A mediana realmente é o valor central de uma amostra. Mas ela tem a propriedade de que metade dos termos da sequência são superiores a ela, e metade dos termos são inferiores. Portanto, a mediana divide o conjunto de valores ordenados em partes **simétricas e iguais.**
Errado.

031. (CESPE/BNB/2018/ANALISTA BANCÁRIO) Em uma faculdade, para avaliar o aprendizado dos alunos em determinada disciplina, o professor aplica as provas A, B e C e a nota final do aluno é a média ponderada das notas obtidas em cada prova. Na prova A, o peso é 1; na prova B, o peso é 10% maior que o peso na prova A; na prova C, o peso é 20% maior que o peso na prova B. Nesse caso, se P_A, P_B e P_C forem as notas obtidas por um aluno nas provas A, B e C, respectivamente, então a nota final desse aluno é expressa por

$$\frac{P_A+1{,}2P_B+1{,}32P_C}{3{,}52}$$

O peso da nota B é 10% maior que o peso da nota A. Portanto, sendo 1 o peso da nota A, concluímos que o peso da nota B é:

$$p_B = (1+0{,}10).1 = 1{,}10$$

Já o peso da nota C é 20% maior que o peso da nota B, que é 1,10. Portanto, podemos escrever:

$$p_c = (1+0{,}20).1{,}10 = 1{,}20.1{,}10 = 1{,}32$$

Agora, reunindo os pesos na expressão da média aritmética ponderada, teríamos:

$$\mu = \frac{P_A + 1{,}1P_B + 1{,}32P_C}{1+1{,}1+1{,}32} = \frac{P_A + 1{,}1P_B + 1{,}32P_C}{3{,}42}$$

Essa expressão está em desacordo com o enunciado. Portanto, a afirmação está errada.
Errado.

032. (CESPE/FUB/2018/ADMINISTRAÇÃO) A tabela seguinte mostra as quantidades de livros de uma biblioteca que foram emprestados em cada um dos seis primeiros meses de 2017.

	mês					
	1	2	3	4	5	6
quantidade	50	150	250	250	300	200

A mediana dos números correspondentes às quantidades de livros emprestados no primeiro semestre de 2017 é igual a 200.

Para calcular a mediana, devemos organizar os termos em ordem crescente.

	mês					
	1	2	3	4	5	6
quantidade	50	150	250	250	300	200
	x_1	x_2	x_4	x_5	x_6	x_3

Como a amostra tem um número par de termos, a mediana será obtida como a média aritmética entre os dois termos centrais. Vejamos. A forma mais simples de obter os termos centrais é tomar a posição central.

$$n = \frac{1+6}{2} = \frac{7}{2} = 3{,}5$$

Como chegamos a um número fracionário, precisamos das posições x_3 e x_4.

$$Md = \frac{x_3 + x_4}{2} = \frac{200+250}{2} = \frac{450}{2} = 225$$

Portanto, a mediana não é igual a 200, mas sim igual a 225.

Errado.

033. (CESPE/2018/IFF/CONHECIMENTOS GERAIS) Considere que o peso de 5 pessoas, juntas em um elevador, seja de 340 kg. Se, em determinado andar, mais um indivíduo entrar no elevador, sem que dele ninguém desça, e a média aritmética dos pesos dessas 6 pessoas passar a ser de 70 kg, esse sexto indivíduo pesa

a) 68,3 kg.
b) 69 kg.
c) 70 kg.
d) 80 kg.
e) 82 kg.

O peso das cinco pessoas juntas no elevador corresponde à soma de todos os pesos na situação inicial. Portanto, podemos anotar S_0 = 340 kg.

Quando um indivíduo é adicionado, a média aritmética passou a ser 70 kg. Sabemos que a soma final dos pesos pode ser obtida como o produto entre a média aritmética e a quantidade de pessoas no elevador. Portanto, a soma final dos pesos das seis pessoas é:

$$S = \mu N = 70.6 = 420$$

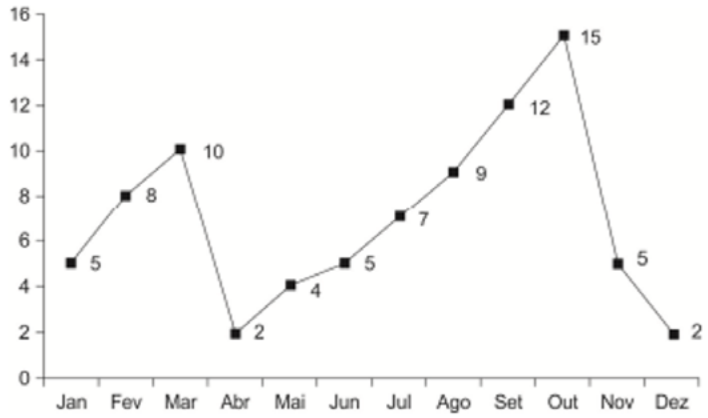
O peso desse último indivíduo a entrar no elevador pode ser obtido como a diferença entre as somas dos pesos após a sua entrada e antes dela.

$$P = S - S_0 = 420 - 340 = 80$$

Letra d.

---

034. (FCC/ARTESP/2017/ESPECIALISTA EM REGULAÇÃO DE TRANSPORTE – ADMINISTRAÇÃO DE EMPRESAS) O gráfico a seguir apresenta, hipoteticamente, a evolução do número de absenteísmo no departamento de operações da empresa que cuida da manutenção das rodovias no litoral do Estado.



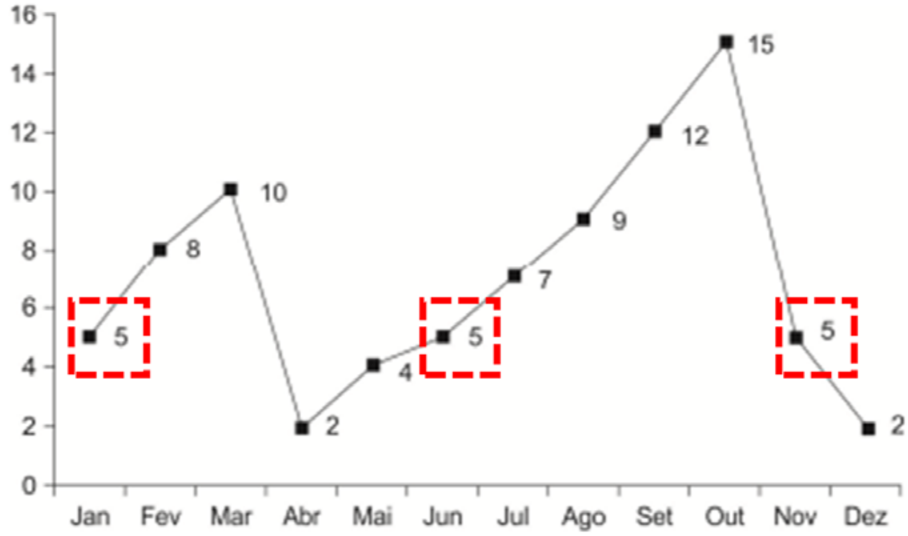
Como forma de resumir todas as características apresentadas por estes dados, utiliza-se as medidas de tendência central, portanto, a média, a moda e a mediana são expressas, respectivamente, por:

a) 7, 5 e 5.
b) 7, 5 e 6.
c) 7, 2 e 6.
d) 7, 2 e 5.
e) 7, 5 e 6,5.

A média pode ser calculada pela definição. Basta somar todos os valores observados e dividir pelo número de elementos (no caso, são 12).

$$\bar{x} = \frac{5+8+10+2+4+5+7+9+12+15+5+2}{12} = \frac{84}{12} = 7$$

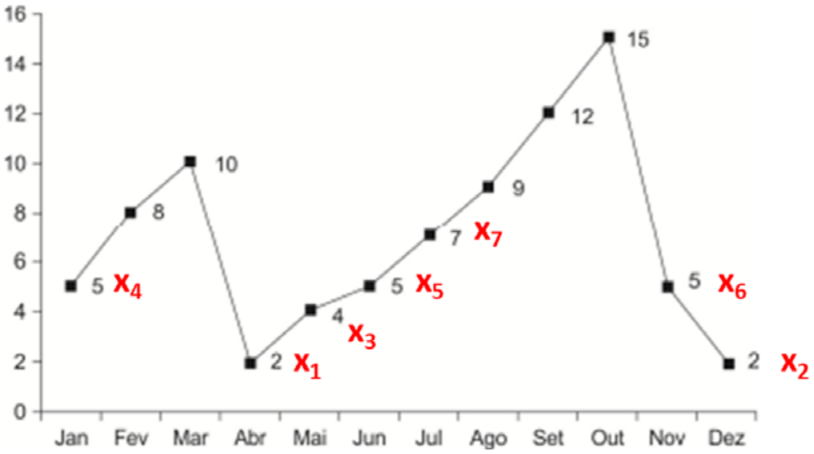
A moda da distribuição é igual a 5, tendo em vista que é o valor que se repete com mais frequência, sendo observado três vezes.



Para obter a mediana, primeiro, devemos organizar

$$n = \frac{1+12}{2} = \frac{13}{2} = 6{,}5$$

$$Md = \frac{x_6 + x_7}{2} = \frac{5+7}{2} = \frac{12}{2} = 6$$



Letra b.

---

035. (FGV/MPE-BA/2017) Um criminoso está avaliando se vale a pena ou não recorrer ao instituto da colaboração premiada. Caso não recorra, a sua probabilidade de ser condenado é igual a p, com 12 anos de reclusão. Se resolver delatar, pode pegar 6 anos de prisão, com probabilidade de 0,4, ou 10 anos, com a probabilidade complementar.

Supondo que a decisão será tomada com base na esperança matemática da pena, o criminoso deve:

a) não delatar se o valor de p for inferior a 0,75;

b) delatar se o valor de p for superior a 0,55;

c) não delatar caso o valor de p seja superior a 0,80;
d) mostrar-se indiferente caso o valor de p seja 0,70;
e) delatar caso o valor de p seja inferior a 0,60.

Caso o criminoso não opte pela delação premiada, o valor esperado de sua pena é:

$$E_1[X] = p.12 = 12p$$

Caso o criminoso opte pela delação premiada, ele tem 40% de chances de pegar seis anos de cadeia e 60% de chances de pegar 10 anos de cadeia. Portanto, o valor esperado de sua pena é:

$$E_2[X] = 0{,}4.6 + 0{,}6.10 = 2{,}4 + 6 = 8{,}4$$

Dessa forma, valerá a pena para ele tentar a delação premiada quando a esperança E_1 for maior que a esperança E_2.

$$12p > 8{,}4$$

$$\therefore p > \frac{8{,}4}{12} = 0{,}70 = 70\%$$

Portanto, valerá a pena delatar quando p > 70% e não vale a pena delatar quando p < 70%. Na situação de 70%, o criminoso será indiferente em relação às duas estratégias.
Letra d.

---

036. (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.
A mediana das quantidades *X* observadas na amostra em questão foi igual a 18 kg.

Como a amostra tem cinco unidades, para obter a mediana, devemos procurar pela posição intermediária entre a inicial e a final, ou seja:

$$\frac{1+5}{2} = \frac{6}{2} = 3$$

Portanto, a mediana equivale ao terceiro termo da sequência, quando organizada por ordem crescente. Vamos organizar as unidades amostrais de X.

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

Portanto, a mediana é igual a $x_3 = 22$.

Errado.

037. (CESPE/SEDUC-AL/2018/PROFESSOR DE MATEMÁTICA) Os babilônicos possuíam um método próprio para o cálculo da raiz quadrada de um número, utilizando aproximações sucessivas. Para determinar o valor aproximado de √P, se estimava, primeiramente, um valor p1 para essa raiz e calculava-se o quociente q1 = p/p1. Com esses dois números calculava-se um novo valor p2, a média aritmética de p1 e q1, isto é, p2 = (p1 + q1)/2. Repetindo esse processo sucessivamente, obtinha-se uma aproximação da raiz quadrada. No método babilônico, se a estimativa inicial para √2 for $p_1 = 1$, então a terceira aproximação de será $p_3 = 17/12$.

Trata-se de um processo iterativo para obter uma aproximação da raiz quadrada. Fazendo $p_1 = 1$, temos:

$$q_1 = \frac{2}{1} = 2$$

A próxima estimativa será a média aritmética entre p_1 e q_1.

$$p_2 = \frac{p_1 + q_1}{2} = \frac{1+2}{2} = \frac{3}{2}$$

Assim, calculemos o próximo quociente:

$$q_2 = \frac{2}{p_2} = \frac{2}{3/2} = \frac{4}{3}$$

E, agora, calculemos a próxima estimativa:

$$p_3 = \frac{p_2 + q_2}{2} = \frac{\frac{3}{2} + \frac{4}{3}}{2} = \frac{3}{4} + \frac{4}{6} = \frac{3}{4} + \frac{2}{3} = \frac{9+8}{12} = \frac{17}{12}$$

É interessante observar que esse valor já é muito próximo da raiz quadrada de 2. Para fins de curiosidade, temos:

$$p_3 = \frac{17}{12} = 1{,}41667$$

$$\sqrt{2} = 1{,}41421$$

Certo.

---

038. (QUADRIX/CRA-AC/2016/ASSISTENTE ADMINISTRATIVO) A Estatística é bastante utilizada em diversos ramos da sociedade, no intuito de realizar pesquisas, colher dados e processá-los, analisar informações, apresentar situações por meio de gráficos de fácil compreensão. O CRA-AC, por exemplo, ao elaborar um relatório ou ao apresentar um projeto, pode utilizar gráficos estatísticos que tornam as informações mais palpáveis e a leitura mais atraente. Um dos conceitos fundamentais da estatística é a mediana, que pode ser definida como:

a) valor representado através de porcentagem, divisão entre a frequência absoluta de cada variável e o somatório das frequências absolutas.
b) medida central em uma determinada sequência de dados numéricos.
c) medida de tendência central. Somatório dos valores dos elementos, dividido pelo número de elementos.
d) somatório dos valores dos elementos multiplicado por seus respectivos pesos, dividido pela soma dos pesos atribuídos.
e) valor de maior frequência em uma série de dados, o que mais se repete.

Embora não seja precisa, a definição que mais se aproxima de mediana é a da letra b.
As letras a, c e d se referem à média aritmética e a letra e se refere à moda.
Letra b.

---

039. (FGV/ISS RECIFE/2014/ANALISTA DE CONTROLE INTERNO) A seguinte amostra de idades foi obtida:
19; 25; 39; 20; 16; 27; 40; 38; 28; 32; 30.
Assinale a opção que indica a mediana dessas idades:
a) 27
b) 28
c) 29
d) 30
e) 31

Para resolver essa questão, devemos organizar a amostra de idades em ordem crescente.
16; 19; 20; 24; 27; 28; 30; 32; 38; 39; 40
Como a amostra é composta por onze elementos, temos que a mediana é calculada por:

$$\frac{1+11}{2}=\frac{12}{2}=6$$
$$Md=x_6=28$$

Apenas para fins de visualização, podemos visualizar que o 28 divide a amostra em duas partes, sendo que existem cinco números menores que 28 e cinco números maiores que 28 presentes na amostra.
16; 19; 20; 24; 27; **28**; 30; 32; 38; 39; 40
Letra b.

040. (FGV/TJ-BA/2015/TÉCNICO JUDICIÁRIO – ÁREA ADMINISTRATIVA) Marcos anotou o número de correspondências eletrônicas que ele recebeu diariamente, durante 13 dias. A tabela a seguir mostra os números anotados por ele:
3 4 18 16 15 16 22 5 2 20 16 15 17
A diferença entre a mediana e a média dos números anotados por Marcos é:
a) 5
b) 4
c) 3
d) 2
e) 1

Em primeiro lugar, notemos que a amostra tem 13 elementos. A seguir, podemos calcular a média somando todos os termos e dividindo por 13.

$$\mu=\frac{3+4+18+16+15+16+22+5+2+20+16+15+17}{13}$$
$$\mu=\frac{169}{13}=13$$

Por outro lado, para calcular a mediana, devemos organizar a amostra por ordem crescente, do primeiro ao décimo terceiro.

3	4	18	16	15	16	22	5	2	20	16	15	17
x_2	x_3	x_{11}	x_7	x_5	x_8	x_{13}	x_4	x_1	x_{12}	x_9	x_6	x_{10}

Dessa forma, a mediana deve ser calculada da seguinte forma.

$$\frac{1+13}{2} = \frac{14}{2} = 7$$

$$Md = x_7 = 16$$

Agora, basta calcular a diferença entre a média e a mediana.

$$Mediana - Média = 16 - 13 = 3$$

Letra c.

---

041. (CESPE/TCU/2008/ANALISTA DE CONTROLE EXTERNO) Considerando a tabela acima, que apresenta a distribuição do quadro de colaboradores da CAIXA, em mil pessoas, no final dos anos de 2006 e 2007, julgue os itens seguintes.
Se as idades médias dos funcionários, estagiários e prestadores de serviço em 2007 foram, respectivamente, iguais a 40 anos, 20 anos e 35 anos, então a média das idades dos colaboradores em 2007 foi inferior a 35 anos.

tipo	2006	2007
funcionários	70	74
estagiários	16	16
prestadores de serviço	14	10
total	**100**	**100**

Demonstrações contábeis da CAIXA – Exercício de 2007.
Internet: <www.caixa.gov.br> (com adaptações).

Como os dados foram fornecidos em categorias, a média aritmética deve ser ponderada.

$$\bar{x}_{2007} = \frac{74.40 + 16.20 + 10.35}{74 + 16 + 10} = \frac{3630}{100} = 36{,}3 > 35$$

Errado.

---

042. (CESPE/2009/MEC/AGENTE ADMINISTRATIVO) Os dados abaixo correspondem às quantidades diárias de merendas escolares demandadas em 10 diferentes escolas:
200, 250, 300, 250, 250, 200, 150, 200, 150, 200.
Com base nessas informações, julgue os próximos itens.
A mediana da distribuição do número diário de merendas escolares é igual a 225.

Para calcular a mediana, devemos colocar a amostra em ordem crescente.
150, 150, 200, 200, 200, 200, 250, 250, 250, 300

Como a amostra tem 10 elementos, devemos tomar como mediana a média aritmética dos dois termos centrais da amostra.

$$\frac{1+10}{2} = \frac{11}{2} = 5{,}5$$

$$Md = \frac{x_5 + x_6}{2} = \frac{200+200}{2} = \frac{400}{2} = 200$$

A seguir, separaremos os dois elementos centrais, evidenciando que existe quatro elementos abaixo deles e quatro elementos acima.

150, 150, 200, 200, **200, 200**, 250, 250, 250, 300

Errado.

---

043. (CESPE/2012/POLÍCIA FEDERAL/PAPILOSCOPISTA) Ao contrário da mediana amostral, a média aritmética é menos sensível à presença de valores extremos (ou valores atípicos ou outliers).

Como estudamos, a mediana é uma estimativa robusta, enquanto a média aritmética é sensível à presença de valores extremos.

Errado.

---

044. (CFC/2015/BACHAREL EM CIÊNCIAS CONTÁBEIS) Uma Sociedade Empresária obteve uma receita total, no ano de 2014, no valor de R$31.200,00, distribuída mensalmente como segue:

Mês	Receita
Janeiro	R$1.000,00
Fevereiro	R$15.000,00
Março	R$1.200,00
Abril	R$1.500,00
Maio	R$1.800,00
Junho	R$2.700,00
Julho	R$1.000,00
Agosto	R$1.600,00
Setembro	R$1.100,00
Outubro	R$1.600,00
Novembro	R$1.400,00
Dezembro	R$1.300,00

Analisando-se os dados e calculando-se as medidas de tendência central, pode-se afirmar que a mediana é de:

a) R$1.450,00

b) R$1.850,00

c) R$2.600,00
d) R$2.700,00

Para calcular a mediana, vamos organizar os dados fornecidos na tabela.

Mês	Receita	Ordem
Janeiro	R$ 1.000,00	1
Fevereiro	R$ 15.000,00	12
Março	R$ 1.200,00	4
Abril	R$ 1.500,00	7
Maio	R$ 1.800,00	10
Junho	R$ 2.700,00	11
Julho	R$ 1.000,00	2
Agosto	R$ 1.600,00	8
Setembro	R$ 1.100,00	3
Outubro	R$ 1.600,00	9
Novembro	R$ 1.400,00	6
Dezembro	R$ 1.300,00	5

Devemos pegar os termos centrais.

$$\frac{1+12}{2}=\frac{13}{2}=6{,}5\rightarrow 6{,}7$$

$$Md=\frac{x_6+x_7}{2}=\frac{1400+1500}{2}=\frac{2900}{2}=1450$$

É interessante perceber que a mediana descarta o valor extremo de R$ 15.000,00 que foi bastante diferente dos demais. É por isso que se diz que a mediana é uma estimativa robusta.
Letra a.

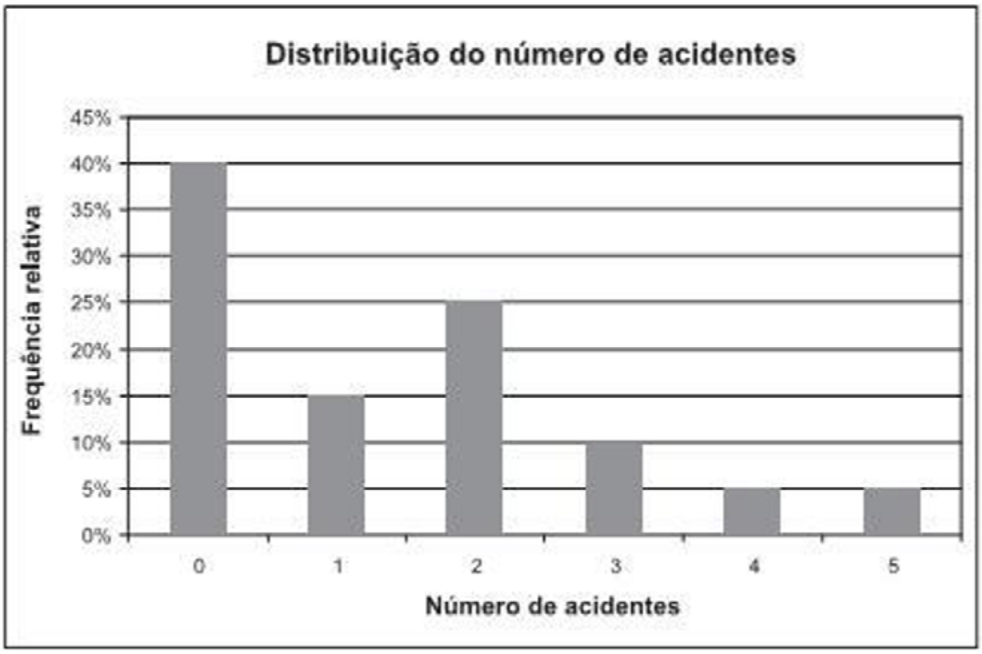
045. (FGV/FIOCRUZ/2010/TÉCNICO EM SAÚDE PÚBLICA) A medida de um conjunto ordenado de dados, que divide este conjunto em duas partes de igual número de observações, denomina-se:

a) Média
b) Moda
c) Mediana
d) Desvio-padrão
e) Variância

Perfeita definição de mediana. Vamos revisar as outras definições.
a) A média é a soma de todos os termos do conjunto dividido pela quantidade de elementos.
b) A moda é a observação mais frequente no conjunto de dados.
c) A mediana é exatamente a definição prevista no enunciado.
d) O desvio padrão é a raiz quadrada da variância.
e) A variância é a média aritmética do quadrado dos desvios de cada elemento da amostra em relação à média.
Letra c.

---

046. (CESGRANRIO/CHESF/2012) O gráfico a seguir apresenta o número de acidentes sofridos pelos empregados de uma empresa nos últimos 12 meses e a frequência relativa.



A mediana menos a média do número de acidentes é:
a) 1,4
b) 0,4
c) 0
d) -0,4
e) -1,4

Quando o gráfico fornece a frequência relativa, devemos tomar como mediana o valor registrado em 50% da amostra, que corresponderá ao termo central.

$$Md = x_{50\%} = 1$$

Dissemos que a mediana é igual a 1, porque, de 0 a 40%, temos a ocorrência de 0 acidentes e, de 40% a 55%, temos 1 acidente.

A média, por sua vez, deve ser calculada com o somatório:

$$\mu = \frac{0{,}40.0 + 0{,}15.1 + 0{,}25.2 + 0{,}10.3 + 0{,}05.4 + 0{,}05.5}{0{,}40 + 0{,}15 + 0{,}25 + 0{,}10 + 0{,}05 + 0{,}05} = \frac{1{,}4}{1} = 1{,}4$$

Sendo assim, a diferença entre a mediana e a média será:

$$Md - \mu = 1 - 1{,}4 = -0{,}4$$

Letra d.

Questões Adicionais Nível 2

047. (CESPE/SEFAZ-DF/2020/AUDITOR-FISCAL) 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.

A mediana amostral da variável X foi igual a 2,5.

É uma bela questão.

Note que a variável só pode assumir dois valores e que a média aritmética simples entre eles é:

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

Portanto, a média aritmética deve ser uma média ponderada para que cheguem ao valor 3. Inclusive, a média é mais próxima de 4 do que de -1, portanto, precisamos de um peso maior para o 4. Por exemplo:

$$\mu = \frac{-1+4+4}{3} = \frac{7}{3} \neq 3$$

Não conseguimos chegar à média desejada. Então, vamos tentar atribuir um novo peso para o 4.

$$\mu = \frac{-1+4+4+4}{4} = \frac{11}{4} \neq 3$$

Mais uma vez, não conseguimos. Mas podemos tentar novamente:

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

Olha só. Notamos que somos capazes de construir uma amostra: {1, 4, 4, 4, 4}, cuja média é igual a 3. Note que, nessa amostra, a mediana é igual a 4.

{1, 4, **4**, 4, 4}

Portanto, a afirmação está incorreta.

Se você não tivesse percebido de cara a proporção correta entre os números, poderíamos estabelecer: p_1 e p_2 como pesos para os números 1 e 4.

$$\mu = \frac{p_1.(-1) + p_2.4}{p_1 + p_2} = \frac{-p_1 + 4p_2}{p_1 + p_2} = 3$$

Fazendo meios pelos extremos, teríamos:

$$-p_1 + 4p_2 = 3.(p_1 + p_2)$$

$$-p_1 + 4p_2 = 3p_1 + 3p_2$$

$$4p_2 - 3p_2 = 3p_1 + p_1$$

$$p_2 = 4p_1$$

Agora, basta fazer p_1 = 1 e p_2 = 4. Poderíamos também fazer outras proporções que chegaríamos ao mesmo resultado.

{ −1, 4, 4, 4, 4}

Por exemplo, p_2 = 2 e p_2 = 8, teríamos:

{−1, −1, 4, 4, 4, 4, 4, 4, 4, 4}

Nesse caso, a mediana é obtida pelos dois termos destacados.

{−1, −1, 4, 4, **4, 4,** 4, 4, 4, 4}

Portanto, a mediana da amostra é:

$$\mu = \frac{4+4}{2} = \frac{8}{2} = 4$$

Mais uma prova de que a afirmação está errada.

Errado.

048. (FCC/PREFEITURA DE MANAUS – AM/2019/AUDITOR-FISCAL DE TRIBUTOS MUNICIPAIS) Conforme um levantamento realizado em um órgão público e analisando a distribuição dos salários, em R$ 1.000,00, de todos os seus funcionários, obteve-se a tabela de frequências absolutas abaixo, com k sendo um número inteiro positivo.

Salários (s)	Número de Funcionários
$2 < s \leq 4$	2k
$4 < s \leq 6$	20
$6 < s \leq 8$	50
$8 < s \leq 10$	80
$10 < s \leq 12$	8k
Total	**40 k**

Considere que a média aritmética (Me) foi calculada considerando que todos os valores incluídos num certo intervalo de classe são coincidentes com o ponto médio deste intervalo, que a mediana (Md) foi calculada pelo método da interpolação linear e que a moda (Mo) foi obtida pela relação de Pearson, ou seja, Mo = 3 Md – 2 Me. O valor encontrado para Mo, em R$ 1.000,00, foi igual a:

a) 1,76 k
b) 1,70 k
c) 1,64 k
d) 1,60 k
e) 1,82 k

Primeiramente, vamos calcular o valor de **k.** Sabemos que o número total de funcionários é 40k. Portanto, podemos escrever:

$$2k + 20 + 50 + 80 + 8k = 40k$$

$$10k + 150 = 40k$$

$$\therefore 150 = 40k - 10k = 30k$$

$$\therefore k = \frac{150}{30} = 5$$

Com base nisso, podemos completar a tabela, na qual incluiremos também o ponto médio de cada classe.

Número de funcionários	Salários (s)	Ponto médio
10	$2 < s \leq 4$	3
20	$4 < s \leq 6$	5
50	$6 < s \leq 8$	7
80	$8 < s \leq 10$	9
40	$10 < s \leq 12$	11

Seguindo as orientações do enunciado, podemos utilizar o ponto médio.

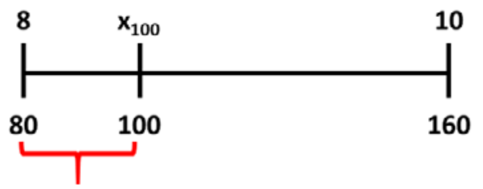
$$Me = \frac{10.3 + 20.5 + 50.7 + 80.9 + 40.11}{10 + 20 + 50 + 80 + 40}$$

$$Me = \frac{30 + 100 + 350 + 720 + 440}{200} = \frac{1640}{200} = 8{,}2$$

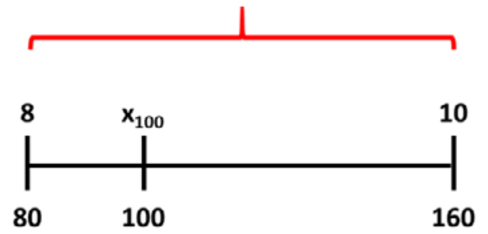
Para calcular a mediana, primeiramente, notemos que o total de funcionários na empresa é igual a 200. Portanto, a mediana equivale ao $\mathbf{x}_{100}$. Agora, vamos organizar a amostra por ordem crescente.

Número de funcionários	Salários (s)	Posição
10	$2 < s \leq 4$	x_1 a x_{10}
20	$4 < s \leq 6$	x_{11} a x_{30}
50	$6 < s \leq \mathbf{8}$	x_{31} a $\mathbf{x}_{80}$
80	**$8 < s < 10$**	**X_{81} a x_{160}**
40	$10 < s \leq 12$	X_{161} a x_{2000}

Agora, basta construir o esquema de interpolar linear. Associamos o final da classe anterior ao elemento máximo da classe.



$$\frac{x_{100} - 8}{100 - 80}$$



$$\frac{10 - 8}{160 - 80}$$

Pela interpolação linear, temos:

$$\frac{x_{100} - 8}{100 - 80} = \frac{10 - 8}{160 - 80}$$

$$\frac{x_{100} - 8}{20} = \frac{2}{80}$$

$$\therefore x_{100} - 8 = 20 \cdot \frac{2}{80} = 0{,}5$$

$$\therefore x_{100} = 8 + 0{,}5 = 8{,}5$$

Aplicando a expressão da moda de Pearson fornecida no enunciado, temos:

$$Moda = 3.Mediana - 2.Média$$

$$Moda = 3.8{,}5 - 2.8{,}2 = 25{,}5 - 16{,}4 = 9{,}1$$

Nas alternativas, os valores aparecem multiplicados por k. Portanto, é interessante dividirmos a moda por k.

$$\frac{Moda}{k} = \frac{9{,}1}{5} = 1{,}82$$

$$\therefore Moda = 1{,}82k$$

Letra e.

---

049. (FCC/SEFAZ-BA/2019/AUDITOR-FISCAL – ADMINISTRAÇÃO TRIBUTÁRIA/PROVA II – QUESTÃO DESAFIO) Considere a distribuição dos salários, em R$1.000,00, dos funcionários lotados em uma repartição pública, representada abaixo pela tabela de frequências relativas acumuladas, sendo **k** a frequência relativa acumulada do 4º intervalo de classe.

Classes de salários	Frequência relativa acumulada (%)
1 ⊢ 3	5
3 ⊢ 5	15
5 ⊢ 7	40
7 ⊢ 9	**k**
9 ⊢ 11	100

Sabe-se que a média aritmética (Me) foi calculada considerando que todos os valores incluídos num certo intervalo de classe são coincidentes com o ponto médio desse intervalo, que a mediana (Md) foi calculada pelo método da interpolação linear e que a moda (Mo) foi obtida pela relação de Pearson, ou seja, Mo = 3Md - 2Me. Dado que Me = R$7.200,00, então Mo é igual a

a) R$ 7.350,00.
b) R$ 8.500,00.
c) R$ 7.700,00.
d) R$ 8.100,00.
e) R$ 7.400,00.

Essa é uma questão bastante trabalhosa e cheia de etapas. Vamos fazê-las uma a uma.

Primeiramente, precisamos calcular o valor de **k**, mas observemos que a tabela foi dada em forma de distribuição acumulada, ou seja, foi dada a soma de todas as frequências relativas de cada classe.

Classes de salários	Ponto médio	Frequência relativa acumulada	Frequência relativa
1 a 3	2	5%	5%
3 a 5	4	15%	10%
5 a 7	6	40%	25%
7 a 9	8	K	**x**
9 a 11	10	100	60% – x

A média foi fornecida no enunciado: Me = R$ 7.200,00. Porém, como os dados foram fornecidos em múltiplos de R$ 1.000, na verdade, o valor calculado para a média foi igual a 7,2.

$$Me = 0{,}05.2 + 0{,}10.4 + 0{,}25.6 + x.8 + (0{,}60 - x).10 = 7{,}2$$

$$0{,}1 + 0{,}4 + 1{,}5 + 8x + 6 - 10x = 7{,}2$$

$$8 - 2x = 7{,}2$$

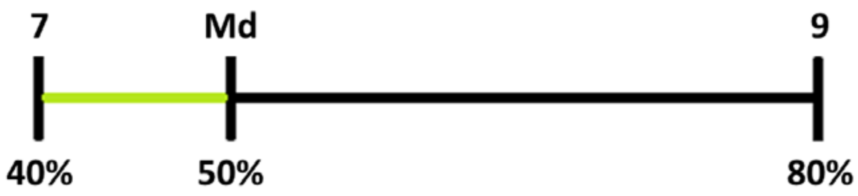
$$\therefore 2x = 8 - 7{,}2 = 0{,}8$$

$$\therefore x = \frac{0{,}8}{2} = 0{,}4 = 40\%$$

Com isso, chegamos à seguinte distribuição de frequências para as classes de salários:

Classes de salários	Frequência acumulada
1 a 3	0% a 5%
3 a 5	5% a 15%
5 a 7	15% a 40%
7 a 9	**40% a 80%**
9 a 11	80% a 100%

A classe mediana de salários é a que está destacada na tabela. A mediana será obtida pelo método da Interpolação Linear.



O primeiro pedaço do segmento é proporcional ao segmento inteiro.

$$\frac{Md - 7}{50\% - 40\%} = \frac{9 - 7}{80\% - 40\%}$$

$$\frac{Md - 7}{10\%} = \frac{2}{40\%}$$

$$\therefore Md - 7 = \frac{10\%}{40\%} \cdot 2 = \frac{1}{4} \cdot 2 = 0{,}5$$

$$\therefore Md = 7 + 0{,}5 = 7{,}5 = R\$7500$$

Por fim, a moda é obtida pela Relação de Pearson. Note que o enunciado foi muito gentil em nos fornecer a expressão dessa relação. Não espere que seja sempre assim.

$$Mo = 3Md - 2Me$$

$$Mo = 3.7500 - 2.7200 = 22500 - 14400 = 8100$$

Letra d.

---

050. (FCC/METRÔ – SP/2019/ENFERMEIRO DO TRABALHO) As massas dos objetos A, B e C satisfazem as seguintes relações:

- as massas de A e B, somadas, excedem em 13 kg a média das massas de B e C;
- subtraindo-se de 79 kg o quádruplo da massa de C, obtém-se a soma da massa de A com o dobro da massa de B.

Assim, a soma das massas de A, B e C, em kg, é igual a

a) 32
b) 34
c) 35
d) 31
e) 33

Vamos analisar e equacionar as informações do enunciado.

- as massas de A e B, somadas, excedem em 13 kg a média das massas de B e C;

$$A + B = \frac{B + C}{2} + 13$$

Multiplicando por 2, temos:

$$2A + 2B = B + C + 26$$

$$2A + B - C = 26$$

- subtraindo-se de 79 kg o quádruplo da massa de C, obtém-se a soma da massa de A com o dobro da massa de B.

$$79 - 4C = A + 2B$$

$$A + 2B + 4C = 79$$

Observe que podemos somar as duas equações:

$$(A + 2A) + (B + 2B) + (-C + 4C) = 26 + 79$$

$$3A + 3B + 3C = 105$$

$$\therefore A + B + C = \frac{105}{3} = 35$$

Letra c.

---

051. (FCC/CÂMARA LEGISLATIVA DO DISTRITO FEDERAL/2018/CONSULTOR TÉCNICO--LEGISLATIVO – ECONOMISTA) Os números de processos autuados em duas repartições públicas (R1 e R2) independentes, durante 40 dias, estão representados na tabela abaixo, sendo m e n inteiros positivos.

Número de processos	0	1	2	3	4	Total
Quantidade de dias (R1)	0	m	15	m	n	40
Quantidade de dias (R2)	2	(n+3)	m	16	4	40

Calculando a soma da média aritmética (número de processos por dia) com a moda e com a mediana de cada repartição, verifica-se que a soma obtida na repartição R2 supera a soma obtida na repartição R1 em:

a) 2,05
b) 0,55
c) 1,05
d) 1,30
e) 1,55

Devemos impor que a soma do número de elementos é igual ao total de elementos, que é 40 em ambos os dias. Para o dia R1, temos:

$$0 + m + 15 + m + n = 40$$

$$2m + n = 40 - 15$$

$$2m + n = 25$$

Para o dia R2, temos:

$$2 + (n + 3) + m + 16 + 4 = 40$$

$$25 + m + n = 40$$

$$\therefore m + n = 40 - 25 = 15$$

Podemos subtrair as duas equações.

$$\begin{array}{cc} (\boldsymbol{I}) & \boldsymbol{2m + n = 25} \\ (\boldsymbol{II}) & m + n = 15 \\ \hline (\boldsymbol{I}) - (\boldsymbol{II}) & m = 25 - 15 = 10 \end{array}$$

Agora, vamos calcular o valor de ***n*.**

$$2m + n = 25$$

$$2.10 + n = 25$$

$$20 + n = 25$$

$$\therefore n = 25 - 20 = 5$$

Com base nisso, vamos montar os diagramas das repartições 1 e 2.

Número de processos	Quantidade de dias (R1)	Quantidade de dias (R2)
0	0	2
1	10	8
2	15	10
3	10	16
4	5	4

Vamos calcular as médias.

$$Me_1 = \frac{0.0 + 10.1 + 15.2 + 10.3 + 5.4}{0 + 10 + 15 + 10 + 5} = \frac{10 + 30 + 30 + 20}{40} = \frac{90}{40} = 2{,}25$$

$$Me_2 = \frac{2.0 + 8.1 + 10.2 + 16.3 + 4.4}{2 + 8 + 10 + 16 + 4} = \frac{0 + 8 + 20 + 48 + 16}{40} = \frac{92}{40} = 2{,}3$$

Para calcular as medianas, devemos identificar as classes medianas. Considerando que a amostra tem 40 elementos, concluímos que a mediana será o elemento dado pela posição:

$$n = \frac{1 + 40}{2} = \frac{41}{2} = 20{,}5$$

$$Md = \frac{x_{20} + x_{21}}{2}$$

Então, vamos obter as classes em que se localizam os elementos x_{20} e x_{21}.

Número de processos	Quantidade de dias (R1)	Acumuladas	Quantidade de dias (R2)	Acumuladas
0	0	0	2	x_1 a x_2
1	10	x_1 a x_{10}	8	x_3 a x_{10}
2	**15**	**x_{11} a x_{25}**	**10**	**x_{11} a x_{20}**
3	10		**16**	**x_{21} a x_{36}**
4	5		4	

Dessa forma, as medianas das duas distribuições são:

$$Md_1 = \frac{x_{20} + x_{21}}{2} = \frac{2 + 2}{2} = \frac{4}{2} = 2$$

$$Md_2 = \frac{x_{20} + x_{21}}{2} = \frac{2 + 3}{2} = \frac{5}{2} = 2{,}5$$

Agora, vamos ao cálculo da moda, que deve ser feito buscando a classe que tem a maior quantidade de dias registrados.

Número de processos	Quantidade de dias (R1)	Quantidade de dias (R2)
0	0	2

Número de processos	Quantidade de dias (R1)	Quantidade de dias (R2)
1	10	8
2	**15**	10
3	10	**16**
4	5	4

Dessa forma, temos as modas:

$$Mo_1 = 2$$

$$Mo_2 = 3$$

Assim, a soma das médias, moda e mediana referentes a cada distribuição são:

$$S_1 = Me_1 + Md_1 + Mo_1 = 2{,}25 + 2 + 2 = 6{,}25$$

$$S_2 = Me_2 + Md_2 + Mo_2 = 2{,}3 + 2{,}5 + 3 = 7{,}8$$

Portanto, a diferença pedida é:

$$S_2 - S_1 = 7{,}8 - 6{,}25 = 1{,}55$$

Letra e.

---

052. (FCC/AFAP/2019/ANALISTA DE FOMENTO) Durante o ano de 2017, foi registrado mensalmente o número de projetos especiais analisados em um órgão público. Apurou-se que a sequência dos números registrados de projetos de janeiro a dezembro foram, respectivamente, {6, 6, 15, 12, 12, 15, 12, 9, 12, 9, 9, 6}, perfazendo então um total de 123 projetos analisados no ano. Com relação a esse período, obteve-se a média aritmética (Me), em número de projetos analisados por mês, a mediana (Md) e a moda correspondentes. Verifica-se que, nesse caso, a moda é igual a:

a) (3Md - 2Me).
b) (2Me + Md - 19).
c) (2Me - Md + 5).
d) (2Me + Md - 22).
e) (3Md - 2Me - 8).

A média dos projetos por mês é igual à razão entre o total de projetos e o número de meses.

$$Me = \frac{123}{12} = 10{,}25$$

A moda corresponde à observação mais frequente. Note que o número **12** é visualizado 4 vezes, enquanto os números **6** e **9** são visualizados apenas 3 vezes. Portanto, a moda é igual a 12.

{**6, 6**, 15, **12, 12,** 15, **12**, **9**, **12**, **9**, **9**, **6**}

$$Mo = 12$$

Primeiramente, vamos organizar o rol de elementos em ordem crescente.

6	6	15	12	12	15	12	9	12	9	9	6
x_1	x_2	x_{11}	x_7	x_8	x_{12}	x_9	x_4	x_{10}	x_5	x_6	x_3

Como o rol tem número par de elementos, a mediana será dada por uma média aritmética de dois termos da amostra.

$$n = \frac{N+1}{2} = \frac{12+1}{2} = \frac{13}{2} = 6{,}5$$

$$Md = \frac{x_6 + x_7}{2} = \frac{9+12}{2} = \frac{21}{2} = 10{,}5$$

Agora, vamos analisar as afirmações.

a) Mo = 3Md – 2Me?

$$3Md - 2Me = 3.10{,}5 - 2.10{,}25 = 31{,}5 - 21{,}5 = 11$$

Observe que a conta sugerida não deu igual à moda, que é igual a 12.

b) 2Me + Md – 19?

$$2Me + Md - 19 = 2.10{,}25 + 10{,}5 - 19$$

$$20{,}5 + 10{,}5 - 19 = 31 - 19 = 12$$

É isso mesmo. Chegamos ao valor da moda. Afirmação correta.

c) 2Me – Md + 5

$$2Me - Md + 5 = 2.10{,}25 - 10{,}5 + 5$$

$$20{,}5 - 10{,}5 + 5 = 10 + 5 = 15$$

Não encontramos o valor correto da moda.

d) 2Me + Md – 22

$$2Me + Md - 22 = 2.10{,}25 + 10{,}5 - 22$$

$$20{,}5 + 10{,}5 - 22 = 31 - 22 = 9$$

Não encontramos o valor da moda.

e) 3Md – 2Me – 8?

$$3Md - 2Me - 8 = 3.10{,}5 - 2.10{,}25 - 8 =$$

$$31{,}5 - 21{,}5 - 8 = 11 - 8 = 3$$

Letra b.

---

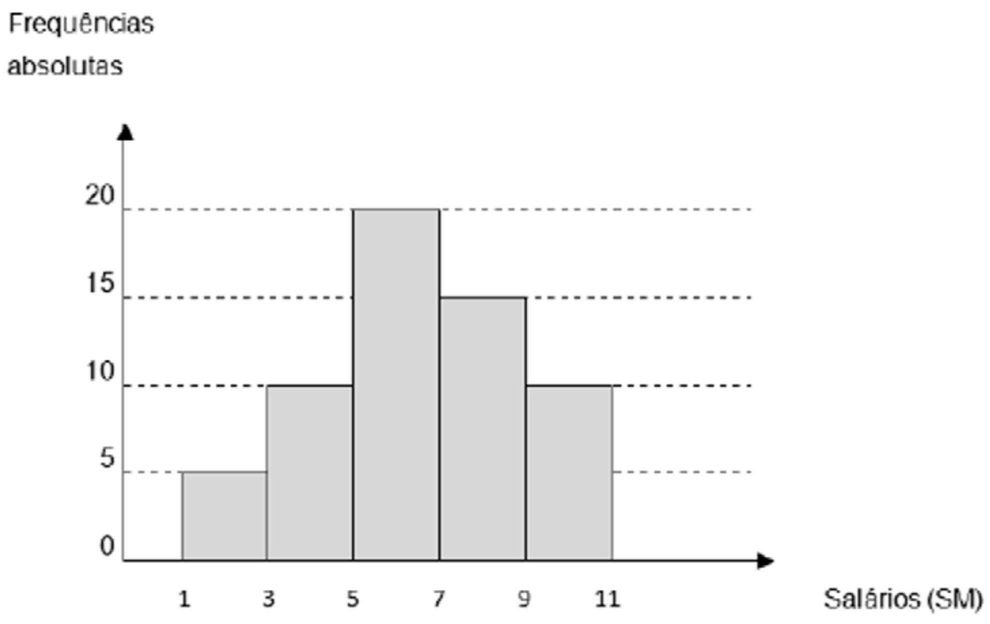
053. (FCC/TRT-11/2017/ANALISTA JUDICIÁRIO – ESTATÍSTICA) Analisando a distribuição dos salários dos empregados de uma empresa em número de salários mínimos (SM), obteve-se o histograma de frequências absolutas abaixo com os intervalos de classe fechados à esquerda e abertos à direita. Considere que:

I – Me é a média aritmética dos salários, calculada levando em conta que todos os valores incluídos num certo intervalo de classe são coincidentes com o ponto médio deste intervalo.

II – Md é a mediana dos salários, calculada por meio do método da interpolação linear.

III – Mo é a moda dos salários, calculada com a utilização da fórmula de King.

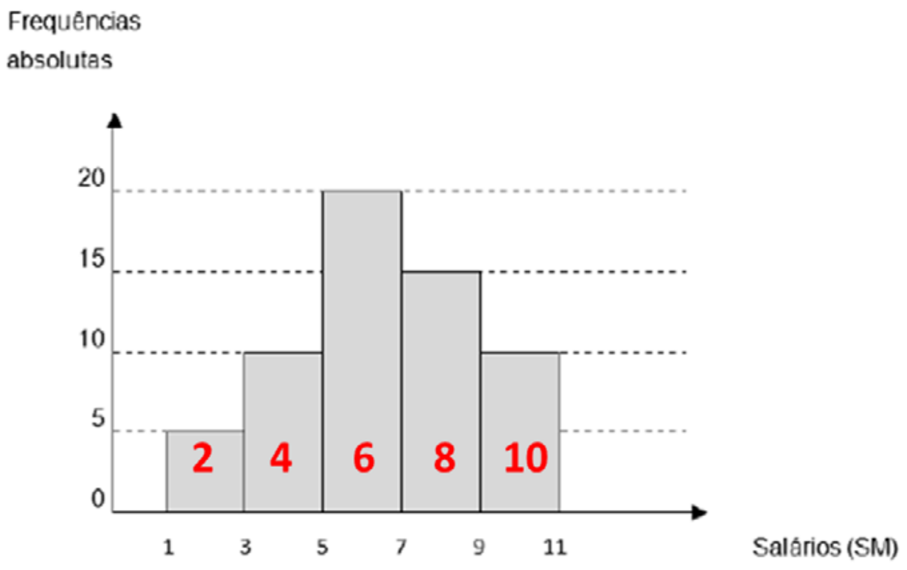
* $Mo = L + \frac{f^{**}}{f^{*} + f^{**}} \times h$, em que L é o limite inferior da classe modal (classe em que se verifica, no caso, a maior frequência), f* é a frequência da classe anterior à classe modal, f** é a frequência da classe posterior à classe modal e h é a amplitude do intervalo de classe correspondente.



O valor de (Me + Md + Mo) é, em SM, igual a

a) 18,6
b) 19,7
c) 19,2
d) 18,7
e) 18,5

A média deve ser calculada pela técnica do ponto médio, o qual será destacado no gráfico.



As frequências absolutas correspondem às alturas de cada uma das barras. E elas devem ser utilizadas como pesos.

$$Me = \frac{5.2 + 10.4 + 20.6 + 15.8 + 10.10}{5 + 10 + 20 + 15 + 10} = \frac{10 + 40 + 120 + 120 + 100}{60} = \frac{390}{60} = 6{,}5$$

Para a mediana, devemos calcular o x_{30}. Para isso, vamos localizar a classe mediana, que é exatamente aquele que envolve o x_{30}.

Classes	Frequência	Acúmulos
1 a 3	5	x_1 a x_5
3 a 5	10	x_6 a $\mathbf{x_{15}}$
5 a 7	**20**	$\mathbf{x_{16}}$ **a** $\mathbf{x_{35}}$
7 a 9	15	x_{36} a x_{50}
9 a 11	10	x_{51} a x_{60}

Agora, vamos utilizar a técnica da interpolação linear. Para isso, devemos considerar que o extremo da classe anterior (x_{15}) coincide com o extremo da classe (5). E que o extremo da classe mediana (X_{35}) coincide com o extremo da própria classe mediana (7).

Apliquemos a técnica da interpolação linear.

$$\frac{Md - 5}{30 - 15} = \frac{7 - 5}{35 - 15}$$

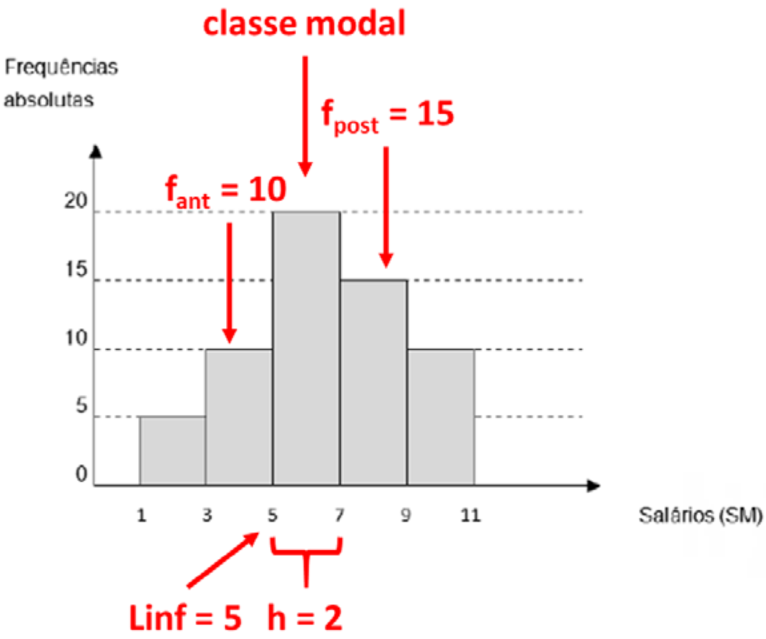
$$\frac{Md - 5}{15} = \frac{2}{20}$$

$$\frac{Md - 5}{15} = 0{,}1$$

$$\therefore Md - 5 = 15.0{,}1 = 1{,}5$$

$$\therefore Md = 5 + 1{,}5 = 6{,}5$$

Agora, vamos ao cálculo da Moda, utilizando a técnica de King. Para isso, precisamos determinar a classe modal, que é aquela que tem a maior frequência absoluta.



$$Mo = Linf + \frac{f_{post}}{f_{ant} + f_{post}} \cdot h$$

$$Mo = 5 + \frac{15}{15 + 10} \cdot 2$$

$$Mo = 5 + \frac{15}{25} \cdot 2 = 5 + 0{,}6.2 = 5 + 1{,}2 = 6{,}2$$

Por fim, a soma das três medidas de posição pedidas é:

$$Me + Md + Mo = 6{,}5 + 6{,}5 + 6{,}2 = 19{,}2$$

Letra c.

---

054. (FCC/SEFAZ-BA/2019/AUDITOR-FISCAL DE ADMINISTRAÇÃO, FINANÇAS E CONTROLE INTERNO – PROVA I) Os números de autos de infração lavrados pelos agentes de um setor de um órgão público, durante 10 meses, foram registrados mensalmente conforme a tabela abaixo.

Mês	I	II	III	IV	V	VI	VII	VIII	IX	X	Total
Número de autos	7	5	4	6	6	5	5	7	6	5	56

Verifica-se que, nesse período, o valor da soma da média aritmética (número de autos por mês) com a mediana é igual ao valor da moda multiplicado por

a) 2,42
b) 2,32

c) 2,12
d) 2,52
e) 2,22

Primeiramente, vamos calcular a média aritmética dos valores:

$$Me = \frac{7+5+4+6+6+5+5+7+6+5}{10} = \frac{56}{10} = 5{,}6$$

A mediana pode ser calculada organizando os números de autos de infração em ordem crescente.

	x_9	x_5	x_1	x_6	x_7	x_2	x_3	x_{10}	x_8	x_4	
Mês	I	II	III	IV	V	VI	VII	VIII	IX	X	**Total**
Número de autos	7	5	4	6	6	5	5	7	6	5	56

Como a amostra tem 10 elementos, a mediana pode ser calculada pelo termo da seguinte posição:

$$\frac{1+10}{2} = \frac{11}{2} = 5{,}5$$

Como a posição encontrada foi fracionária, basta pegar o x_5 e o x_6 e tirar a média aritmética.

$$Md = \frac{x_5 + x_6}{2} = \frac{5+6}{2} = \frac{11}{2} = 5{,}5$$

A moda, por sua vez, pode ser obtida como a observação mais frequente. O número de autos igual a 5 foi observado em 4 ocasiões. Portanto, é moda amostral.

$$Mo = 5$$

Agora, vamos calcular a soma da média com a mediana.

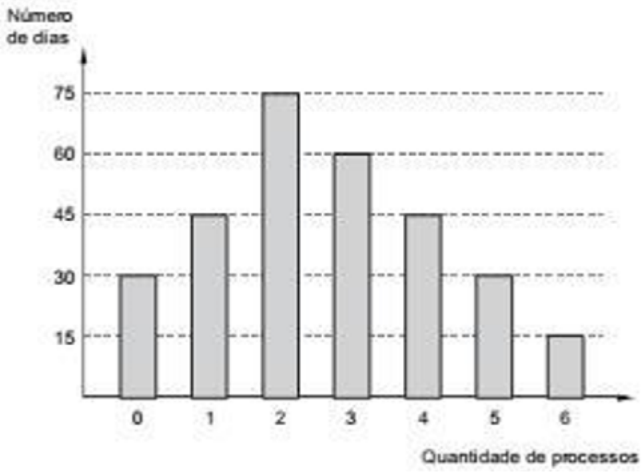
$$Me + Md = 5{,}6 + 5{,}5 = 11{,}1$$

Agora, vamos calcular a relação pedida entre a soma da média e da mediana e a moda.

$$\frac{Me + Md}{Mo} = \frac{11{,}1}{5} = 2{,}22$$

Letra e.

055. (FCC/CNMP/2015/ANALISTA DO CNMP – ESTATÍSTICA) Analisando a quantidade diária de processos autuados em uma repartição pública, durante um período, obteve-se o seguinte gráfico em que as colunas representam o número de dias em que foram autuadas as respectivas quantidades de processos constantes no eixo horizontal.



A soma dos valores respectivos da mediana e da moda supera o valor da média aritmética (quantidade de processos autuados por dia) em:

a) 1,85
b) 0,50
c) 1,00
d) 0,85
e) 1,35

A moda é a mais fácil de ser calculada. Basta pegar o valor mais frequente, que é 2.

Observe que temos um total de 30 + 45 + 75+ 60 + 45 + 30 + 15 = 300 dias que foram estudados. Sendo assim, a mediana deve ser calculada tomando-se:

$$\frac{1+300}{2} = \frac{301}{2} = 150{,}5$$

Portanto, precisamos dos termos x_{150} e x_{151}. Para isso, precisamos notar que:

x_i	Número de elementos	Elementos
0	30	1 a 30
1	45	31 a 75
2	75	76 a 150
3	60	151 a 210

Dessa maneira, temos que a mediana será:

$$Md = \frac{x_{150} + x_{151}}{2} = \frac{2+3}{2} = 2{,}5$$

A média, por sua vez, deve ser calculada como a média ponderada usando as frequências absolutas como pesos.

$$\mu = \frac{30.0 + 45.1 + 75.2 + 60.3 + 45.4 + 30.5 + 15.6}{300} = \frac{795}{300} = 2{,}65$$

Sendo assim, a soma pedida entre a média, moda e mediana é:

$$S = 2{,}5 + 2 - 2{,}65 = 4{,}5 - 2{,}65 = 1{,}85$$

Letra a.

---

056. (FCC/TRF 2ª REGIÃO/2012/ANALISTA JUDICIÁRIO – ESTATÍSTICA) Em dezembro de 2011 foi realizado um levantamento em uma empresa que proporcionou a tabela de frequências relativas abaixo, referente aos salários de seus empregados, observando que 3m + n = 25%.

CLASSE DE SALÁRIOS (R$)	FREQUÊNCIA RELATIVA (%)
1.500 ⊢ 2.500	3m
2.500 ⊢ 3.500	2n
3.500 ⊢ 4.500	5m
4.500 ⊢ 5.500	3n
5.500 ⊢ 6.500	n
TOTAL	100

O valor da média aritmética (Me) foi obtido considerando que todos os valores incluídos num intervalo de classe são coincidentes com o ponto médio deste intervalo. O valor da mediana (Md) foi obtido pelo método da interpolação linear. Então, tem-se que:

a) Md = 0,900Me.
b) Md = 0,950Me.
c) Md = 1,000Me.
d) Md = 1,025Me.
e) Md = 1,125Me.

Em primeiro lugar, devemos lembrar que a soma das frequências relativas deve ser igual a 100%.

$$3m + 2n + 5m + 3n + n = 100\%$$

$$\therefore 8m + 6n = 100\%$$

O enunciado nos forneceu uma informação adicional sobre as incógnitas m e n.

$$3m + n = 25\%$$

$$\therefore 8m + 6n = 100\%$$

Temos, portanto, um sistema de duas equações e duas incógnitas. Podemos, por exemplo, multiplicar a segunda por 6.

$$18m + 6n = 150\%$$

$$\therefore 8m + 6n = 100\%$$

Agora, subtraindo, temos:

$$18m - 8m = 150\% - 100\%$$

$$10m = 50\% \therefore m = \frac{50\%}{10} = 5\%$$

Agora, basta substituir o valor encontrado em uma das equações.

$$3m + n = 25\%$$

$$3.5\% + n = 25\% \therefore n = 25\% - 15\% = 10\%$$

Agora que encontramos os parâmetros m e n, podemos construir com mais facilidade as classes e seus percentuais.

Classe de salários	Frequência relativa
R$ 1.500 a R$ 2.500	15%
R$ 2.500 a R$ 3.500	20%
R$ 3.500 a R$ 4.500	25%
R$ 4.500 a R$ 5.500	30%
R$ 5.500 a R$ 6.500	10%

Como o próprio enunciado deu a dica, devemos calcular a média utilizando o ponto médio.

Ponto médio	Frequência relativa
R$ 2.000	15%
R$ 3.000	20%
R$ 4.000	25%
R$ 5.000	30%
R$ 6.000	10%

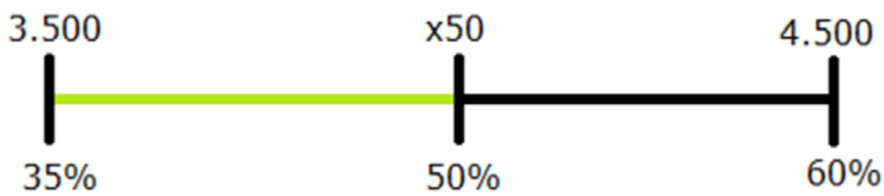
A média pode ser obtida usando como pesos as frequências relativas que foram informadas pelo enunciado. Além disso, quando usamos as frequências relativas como pesos, o somatório desses pesos é sempre igual a 1 (ou 100%).

$$\mu = Me = 0{,}15.2000 + 0{,}20.3000 + 0{,}25.4000 + 0{,}30.5000 + 0{,}10.6000$$
$$= 300 + 600 + 1000 + 1500 + 600 = 4000$$

Agora, para calcular a mediana, devemos encontrar o $x_{50\%}$. Para isso, devemos encontrar em que classe está esse termo.

Classe de salários	Frequência relativa	Frequência acumulada
R$ 1.500 a R$ 2.500	15%	15%
R$ 2.500 a R$ 3.500	20%	**35%**
R$ 3.500 a R$ 4.500	25%	**60%**
R$ 4.500 a R$ 5.500	30%	90%
R$ 5.500 a R$ 6.500	10%	100%

Destacamos na tabela os pontos de interesse para fazer a interpolação linear.



O primeiro pedaço, de 35% a 50%, é proporcional ao segmento completo. Com isso, temos:

$$\frac{x_{50} - 3500}{50\% - 35\%} = \frac{4500 - 3500}{60\% - 35\%}$$

$$\frac{x_{50} - 3500}{15\%} = \frac{1000}{25\%}$$

$$\therefore x_{50} - 3500 = \frac{1000.15\%}{25\%} = \frac{1000.15}{25} = 40.15 = 600$$

$$\therefore Md = x_{50} = 3500 + 600 = 4100$$

Agora, finalmente, podemos calcular a relação entre mediana e moda.

$$\frac{Md}{Me} = \frac{4100}{4000} = 1{,}025 \therefore Md = 1{,}025Me$$

Letra d.

057. (FCC/TCE-PR/2011/ANALISTA DE CONTROLE) [Considerando os dados do enunciado anterior,] O valor de X – md, em número de salários mínimos, é:

Dados: md = mediana dos salários calculada pelo método da interpolação linear;

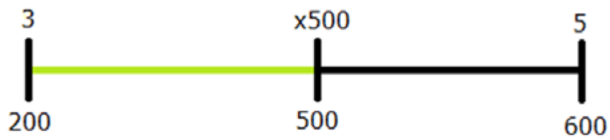
X – = valor que separa os 15% dos salários mais altos, calculado pelo método da interpolação linear.

a) 4,00
b) 3,25
c) 3,50
d) 3,75
e) 3,00

Como a amostra tem 1000 elementos, o cálculo da mediana requer o cálculo do x_{500}. Primeiramente, devemos encontrar a classe em que está esse elemento para fazermos a interpolação linear.

Classe de salários	Frequência absoluta	Frequência acumulada
1 a 3	200	**200**
3 a 5	400	**600**
5 a 7	200	800
7 a 9	200	1000

Destacamos os dados necessários para a interpolação linear.



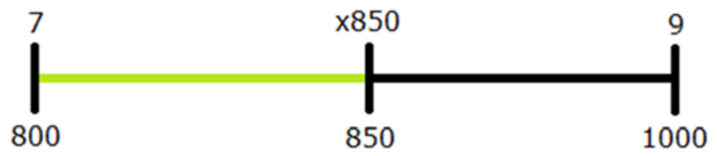
$$\frac{x_{500} - 3}{500 - 200} = \frac{5 - 3}{600 - 200} \therefore \frac{x_{500} - 3}{300} = \frac{2}{400}$$

$$\therefore x_{500} - 3 = \frac{2.300}{400} = 1{,}5 \therefore Md = x_{500} = 3 + 1{,}5 = 4{,}5$$

O enunciado também pediu o valor que representa os elementos 15% superiores. Dessa maneira, devemos tomar o x_{850}. Façamos:

Classe de salários	Frequência absoluta	Frequência acumulada
1 a 3	200	200
3 a 5	400	600
5 a 7	200	**800**
7 a 9	200	**1000**

Façamos o esquema da interpolação linear.



Aplicando que a parte menor deve ser proporcional ao todo.

$$\frac{x_{850} - 7}{850 - 800} = \frac{9 - 7}{1000 - 800} \therefore \frac{x_{850} - 7}{50} = \frac{2}{200}$$

$$\therefore x_{850} - 7 = \frac{2.50}{200} = \frac{100}{200} = 0{,}5 \therefore X = x_{850} = 7 + 0{,}5 = 7{,}5$$

Portanto, a diferença solicitada é:

$$X - Md = 7{,}5 - 4{,}5 = 3$$

Letra e.

---

058. (CESPE/EBC/2011/ANALISTA – ADMINISTRAÇÃO) Com base nos dados do quadro acima, em que se demonstra a distribuição de frequência das receitas de todas as empresas de uma cidade, julgue os itens a seguir.

classes	receitas (em R$)	quantidade de empresas
1	0 – 200.000	1.100
2	200.001 – 400.000	900
3	400.001 – 600.000	550
4	600.001 – 800.000	300
5	800.001 – 1.000.000	150

É correto inferir que a média das receitas das empresas da cidade em apreço é inferior a R$ 332 mil.

Como sabemos, a média deve ser calculada considerando-se os pontos médios das classes fornecidas.

Classe	Ponto médio	Quantidade de empresas
1	R$ 100.000	1100
2	R$ 300.000	900
3	R$ 500.000	550
4	R$ 700.000	300
5	R$ 900.000	150

$$Me = \frac{1100.100\ mil + 900.300\ mil + 550.500\ mil + 300.700\ mil + 150.900\ mil}{1100 + 900 + 550 + 300 + 150}$$

$$\frac{1000000\ mil}{3000} = 333333 > 332\ mil$$

Errado.

---

059. (ESAF/AFRF/2002) Frequências acumuladas de salários anuais, em milhares de reais, da Cia. Alfa.

Classes de Salários	Frequências Acumuladas
3 a 6	12
6 a 9	30
9 a 12	50
12 a 15	60
15 a 18	65
18 a 21	68

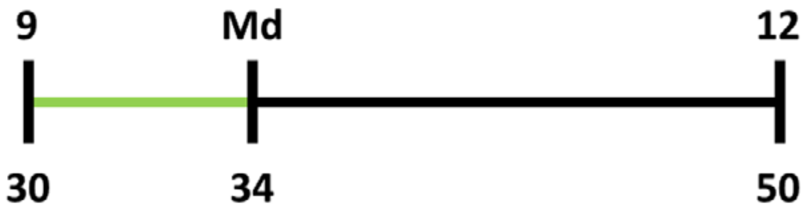
Quer-se estimar o salário mediano anual da Cia. Alfa. Assinale a opção que corresponde ao valor aproximado desta estatística, com base na distribuição de frequências.

a) 12,50.
b) 9,60.
c) 9,00.
d) 12,00.
e) 12,10.

Quando os dados são fornecidos na forma de categorias, deve-se utilizar a técnica da interpolação linear para calcular a mediana.
Como são 68 termos, a mediana será o termo x_{34}. Lembre-se de que a mediana deve ser calculada como 68/2 para dados categorizados. Deveria ser (68+1)/2 apenas para dados não categorizados.

Classes de salários	Frequências acumuladas
3 a 6	12
6 a 9	**30**
9 a 12	**50**
12 a 15	60
15 a 18	65
18 a 21	68

Vamos utilizar a técnica de interpolação linear. Para isso, dizemos que o extremo da classe mediana (x_{50}) de salários está associado ao valor máximo dessa classe (12). Da mesma forma, o extremo da classe anterior (x_{30}) está associado ao máximo dessa classe (9).



O valor da mediana corresponde à posição x_{34}, que pode ser obtido por interpolação linear.

$$\frac{Md - 9}{34 - 30} = \frac{12 - 9}{50 - 30}$$

$$\frac{Md - 9}{4} = \frac{3}{20}$$

$$\therefore Md - 9 = \frac{3}{20} \cdot 4 = \frac{12}{20} = 0{,}6$$

$$\therefore Md = 9 + 0{,}6 = 9{,}6$$

Letra b.

---

060. (CESPE/FUB/2016/ASSISTENTE EM ADMINISTRAÇÃO) Em um almoxarifado há, em estoque, 100 caixas na forma de paralelepípedos retângulos. Na tabela a seguir são mostra-dos alguns valores da frequência absoluta, da frequência relativa e da porcentagem da variável volume interno da caixa, em litros (L).

volume da caixa (L)	frequência absoluta	frequência relativa	porcentagem (%)
10	10	*	*
20	*	*	*
45	*	0,2	*
60	*	*	40
Total	100	1	100

Considerando essas informações, julgue o seguinte item
A média aritmética dos volumes dessas caixas é igual a 40 L.

As caixas que possuem 60 L correspondem a 40% do total de caixas. Como são 100 caixas, concluímos que 40% de 100 caixas é igual a 40 caixas.
Já as caixas de 45 L correspondem à frequência relativa de 0,2 ou 20%. Como são 100 caixas, 20% de 100 é igual a 20 caixas.
Como o total de caixas é igual a 100, o número de caixas de 20 L é o que falta para 100.

$$N_{20} = 100 - 10 - 40 - 20 = 30$$

Com base nisso, podemos fechar a tabela.

Volume da caixa	Frequência
10	10
20	30
45	20
60	40

Agora, vamos calcular o volume médio das caixas como a média aritmética dos volumes ponderados pela frequência.

$$\mu = \frac{10.10 + 30.20 + 20.45 + 40.60}{10 + 30 + 20 + 40} = \frac{100 + 600 + 900 + 2400}{100}$$

$$\mu = \frac{4000}{1000} = 40$$

Certo.

Chegamos ao final de mais uma aula.

Forte abraço!

Thiago Cardoso.

GABARITO

1. C
2. E
3. a
4. b
5. b
6. b
7. C
8. e
9. e
10. d
11. b
12. C
13. E
14. c
15. e
16. e
17. c
18. e
19. b
20. b
21. b
22. d
23. e
24. E
25. c
26. C
27. E
28. C
29. C
30. E
31. E
32. E
33. d
34. b
35. d
36. E
37. C
38. b
39. b
40. c
41. E
42. E
43. E
44. a
45. c
46. d
47. E
48. e
49. d
50. c
51. e
52. b
53. c
54. e
55. a
56. d
57. e
58. E
59. b
60. C

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.
NÃO SE ESQUEÇA DE AVALIAR ESTA AULA!
SUA OPINIÃO É MUITO IMPORTANTE PARA MELHORARMOS AINDA MAIS NOSSOS MATERIAIS.
ESPERAMOS QUE TENHA GOSTADO DESTA AULA!
PARA AVALIAR, BASTA CLICAR EM LER A AULA E, DEPOIS, EM AVALIAR AULA.
AVALIAR