ESTATÍSTICA

Ana Laura Bertelli Grams

Medidas de posição: média, mediana e moda

Objetivos de aprendizagem

Ao final deste texto, você deve apresentar os seguintes aprendizados:

- Calcular as medidas de posição: média, mediana e moda.
- Escolher a medida de posição mais adequada.
- Aplicar as medidas estatísticas a partir das definições.

Introdução

Após a coleta e organização dos dados de uma pesquisa, é fundamental que se faça a análise para futura tomada de decisão. A análise mais trivial de um conjunto de dados é feita por meio de medidas de posição.

Neste capítulo, você reconhecerá as medidas de posição central, chamadas média, mediana e moda, identificando suas definições, características e aplicações em conjuntos numéricos agrupados e não agrupados.

Medidas de posição: média, mediana e moda

Para análise das variáveis qualitativas, precisamos nos restringir apenas à sua distribuição de frequências, enquanto que, em sua análise, as variáveis quantitativas permitem que algumas medidas que descrevem suas características sejam manipuladas e praticadas (BARBETTA; REIS; BORNIA, 2008). As medidas que estudaremos agora serão **medidas de posição central**.

As medidas estatísticas informam características importantes da amostra, que geralmente são um rol com muitos dados difíceis de serem analisados quando apresentados todos juntos. Por isso, buscamos algumas medidas que os descrevem. As medidas de posição mais utilizadas são as de tendência central: **média**, **mediana** e **moda**.

Essas medidas são chamadas de medidas de tendência central, pois cada uma delas tende a se dispor em torno dos valores que ocupam as posições centrais de um rol de dados. Além delas, temos as medidas de posição chamadas separatrizes, que são: **quartil**, **decil** e **percentil**.

Média

A média é definida como o centro de massa, ou o ponto de equilíbrio, do conjunto (MILONE, 2006). Entre as principais médias, destacamos a **média aritmética**.

A média aritmética é calculada por meio da soma dos dados (quantitativos) do conjunto e da divisão da soma pela quantidade de dados do conjunto:

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

onde x_i representa os dados em questão (na posição 1 até *n-ésima*), e n a quantidade de dados do conjunto.

Características da média

1. A média é afetada por **todos** os elementos do conjunto (para o seu cálculo, é preciso somar todos eles). Como consequência, ela se altera a cada mudança dos elementos do conjunto, e, ainda, valores de extremos, muito altos ou muito baixos, tendem a aumentá-la ou diminuí-la, respectivamente, de maneira bastante significativa.

Exemplo

Sendo 30, 32, 44, 82 e 97 dados de uma amostra qualquer, sua média é obtida com $\bar{x} = \frac{30+32+44+82+97}{5} = 57$. Se qualquer dado for afetado por alguma mudança, a média também será afetada, especialmente se os extremos se alterarem:

2, 32, 44, 82, 97 $\rightarrow \bar{x} = \frac{2+32+44+82+97}{5} = 51{,}4$ ou ainda: 30, 32, 44, 82 e 250

$\rightarrow \bar{x} = \frac{30+32+44+82+250}{5} = 87{,}6$.

2. A média apresenta propriedades algébricas de manipulação, que são: somando-se uma constante a todos os dados da amostra, a média é aumentada da mesma constante.

Exemplo

A média dos valores 41, 75 e 64 é $\frac{41 + 75 + 64}{3} = 60$. Ao somarmos a constante 5 aos dados, temos 46, 80, 69, e a média dos novos valores é $\frac{46 + 80 + 69}{3} = 65$.

3. O valor da média estará sempre entre o maior e o menor valor do conjunto de dados e pode não corresponder a algum valor do próprio conjunto.

Exemplo

Como, no conjunto anterior (41, 75, 64), a média é igual a 60, sendo assim, $41 < \bar{x} < 75$ e, ainda, não é igual a nenhum dado do conjunto.

Média de dados agrupados

O conceito de **média** e suas características mantém-se para qualquer conjunto de dados. Contudo, o processo do cálculo pode variar, dependendo de como esses dados estão apresentados. O caso mais simples para encontrar o valor da média é em um rol de dados simplesmente ordenados (ou não), em que basta aplicarmos a equação que a define. Já em dados que são apresentados em uma distribuição de frequência, precisamos de uma etapa anterior, para então aplicarmos a mesma fórmula.

Considere a tabela de distribuição de frequência no Quadro 1, relativa ao número de acidentes ocorridos com 30 motociclistas em uma empresa de entrega rápida.

Quadro 1. Número de acidentes com 10 motoristas de mototáxi

Número de acidentes (variável)	Número de motociclistas (frequência)
1	13
2	5
3	9
4	1
5	2

As frequências dos acidentes indicam a intensidade deles, facilitando a apresentação das variáveis. Contudo, para o cálculo da média, precisamos ficar atentos a elas e não nos esquecer de que cada variável tem a sua quantidade indicada na coluna ao lado. O cálculo da média de acidentes por motociclista deve ser feito da seguinte maneira:

$$\bar{x} = \frac{(13 \cdot 1) + (5 \cdot 2) + (9 \cdot 3) + (1 \cdot 4) + (2 \cdot 5)}{13 + 5 + 9 + 1 + 2} = 2{,}133$$

onde cada acidente é multiplicado pela frequência em que ocorreram e a soma deles dividida pelo total de motociclistas na empresa.

De maneira geral, a média em uma distribuição de frequência é calculada pela lei:

$$\bar{x} = \frac{\sum (x_i \cdot f_i)}{\sum f_i}$$

Ou seja, o somatório do produto entre a variável (x_i) e a sua frequência correspondente a (f_i), divido pelo somatório das frequências ($\sum f_i$).

Média de dados agrupados com intervalos de classe

Além do formato do Quadro 1 para apresentação dos dados, podemos, ainda, expressá-los por meio de intervalos de classe, que se trata do agrupamento dos valores em intervalos. Essa prática é comumente utilizada em variáveis contínuas e quando cada valor tem uma baixa frequência, resultando, assim, em uma tabela com muitas linhas, que se torna inconveniente para análise. O

Quadro 2 mostra um exemplo de distribuição de frequência com intervalos de classe.

Quadro 2. Estatura (em cm) de 50 alunos de uma classe

Estatura (variável)	Número de alunos (frequência)
160 ⊢ 165	5
165 ⊢ 170	20
170 ⊢ 175	11
175 ⊢ 180	1
180 ⊢ 185	3

Por característica das distribuições de frequência com dados agrupados, ocultamos algumas informações anteriormente tidas nos dados brutos. Perceba que a tabela nos indica que cinco estudantes apresentam estatura entre 160 cm e 165 cm, porém não nos orienta para a altura exata de cada um deles.

Para cálculo da média de dados apresentados dessa forma, precisamos assumir um único valor para esses intervalos de classe. Fizemos isso por meio do cálculo da própria média das classes. Para o exemplo anterior, teremos o Quadro 3.

Quadro 3. Estatura (em cm) de 50 alunos de uma classe — inserção das colunas x_i e $x_i \cdot f_i$ para cálculo da média

Estatura (variável)	x_i (média das classes)	Número de alunos (f_i)	$x_i \cdot f_i$
160 ⊢ 165	$\frac{160 + 165}{2} = 162{,}5$	5	812,5
165 ⊢ 170	167,5	20	3350
170 ⊢ 175	172,5	11	1897,5
175 ⊢ 180	177,5	1	177,5
180 ⊢ 185	182,5	3	547,5
			6785

Note que, no Quadro 3, inserimos, além da média das classes, uma coluna com a multiplicação entre a variável e a frequência. Isso pode facilitar no cálculo da média. Contudo, é o mesmo que aplicarmos a seguinte lei:

$$\bar{x} = \frac{\sum (x_i \cdot f_i)}{\sum f_i}$$

$$\bar{x} = \frac{812{,}5 + 3350 + 1897{,}5 + 177{,}5 + 547{,}5}{40} = \frac{6785}{40} = 169{,}63 \text{ cm}$$

Concluímos, assim, que a média das estaturas entre os 40 alunos pesquisados é 169,63 cm.

Mediana

Outra medida de centro bastante utilizada é a mediana. Seu conceito é dado por: **o valor que se encontra no centro de uma série ordenada de números**. Ou seja, é o dado que divide o conjunto **ordenado** em dois subconjuntos de mesmo número de elementos (CRESPO, 2002).

A posição da mediana é encontrada por $\frac{n+1}{2}$. Em um conjunto de dados não agrupados, como 8, 5, 14, 9, 56, 32, 23, no qual temos n = 7 dados, a posição da mediana é dada por $\frac{8}{2} = 4$, ou seja, na quarta posição. Contudo, antes de localizarmos o dado que se encontra na quarta posição, é preciso ordená-los segundo um critério preestabelecido, de ordem crescente, por exemplo. Sendo assim, temos 5, 8, 9, 14, 23, 32, 56, onde constatamos que a mediana é igual a 14.

Em casos em que a quantidade de dados é par, teremos dois termos no centro da série. Assim, precisamos encontrar o ponto médio dos dois valores para determinarmos a mediana. Na série 2, 5, 8, 9, 14, 23, 32, 56, o quarto e o quinto termos são que dividem a série em dois subconjuntos com o mesmo número de elementos. Dessa forma, a mediana dessa é dada por $\frac{9 + 14}{2} = 11{,}5$.

Fique atento

Perceba que a mediana, além de uma medida de tendência central, também é considerada **separatriz**, pois divide o conjunto de dados em duas partes com iguais quantidades de elementos.

Saiba mais

As separatrizes separam o conjunto de dados em grupos com o mesmo número de valores, os quartis dividem o conjunto em 4 (quatro) partes iguais, os decis em 10 (dez) e os percentis em 100 (cem).

Moda

A moda é geralmente a medida de tendência central mais simples de ser informada, pois exige apenas a observação dos dados existentes. Definimos moda como o valor que ocorre com maior frequência em um conjunto de dados. Ou seja, é o valor mais comum dentre todos do conjunto.

No exemplo 2, 5, 8, 9, 14, 23, 32, 56, temos um conjunto em que todos os elementos têm a mesma frequência. Isso implica em um conjunto **amodal**, ou sem moda. Já a série de dados 2, 5, 8, 8, 8, 9, 9 14, 23, 32, 56 tem moda igual a 8, e a série 2, 5, 8, 8, 8, 9, 9 14, 23, 32, 56, 56, 56 tem duas modas: 8 e 56. Neste último caso, chamamos o conjunto de **bimodal**.

Escolha da medida de posição mais adequada

A escolha entre a média, a mediana e a moda depende dos fatores que elas afetam. É necessário conhecer suas propriedades com a finalidade de adequar a melhor medida a cada caso em estudo.

Uma das características da média é sua sensibilidade a valores muito altos ou muito baixos do conjunto de dados, pois é uma medida que reflete cada valor do conjunto. Sendo assim, uma análise possível é: **quando os valores extremos do conjunto de dados são consideravelmente dispersos dos demais, a média não é uma medida de posição indicada para análise**, pois ela não representa adequadamente a maioria dos dados do conjunto.

Por outro lado, a mediana é, de fato, insensível aos valores extremos do conjunto, podendo estes se alterarem, e, mesmo assim, a mediana se manter. Portanto, no caso citado, a indicação é a utilização da mediana como medida de posição mais adequada.

Em contrapartida, a média é mais prática de ser calculada, visto que, para encontrar a mediana, é imprescindível a ordenação dos dados, o que acarreta em grande dificuldade quando o conjunto apresenta grande quantidade de dados, sobretudo quando não se utiliza de recursos tecnológicos para tal.

A moda é geralmente um ponto isolado, mas de maior peso no conjunto de elementos. Sua característica é vantajosa sobre as demais, pois é sempre um valor típico, o qual tem maior quantidade de valores concentrados no mesmo ponto.

Fique atento

Quando temos dados qualitativos, não podemos aplicar as medidas de posição média e mediana, por motivos óbvios. Em contrapartida, a moda é uma medida de posição que pode ser obtida mesmo em conjuntos de dados qualitativos.

Aplicação a partir das definições

Nesta etapa de estudo, aplicaremos os conceitos estudados anteriormente em alguns exemplos de atividades, a fim de utilizar as ferramentas estatísticas para o desenvolvimento do raciocínio lógico, enquanto descobrimos a melhor maneira para encontrar as soluções.

Exemplo

Em um conjunto com 15 dados, a média aritmética é igual a 9. Depois de uma vistoria detalhada nos dados, descobriu-se que alguns eram inconsistentes e precisavam ser desconsiderados. Assim, os números 34, 27, 14 foram retirados. Qual será a nova média do conjunto?

Solução:
Temos que o primeiro conjunto tinha média igual a:

$$\bar{x} = \frac{x_1 + \cdots x_{15}}{15} = 9$$

Assim, a soma de todos os 15 elementos do conjunto de dados é dada por:

$$x_1 + \cdots x_{15} = 9 \cdot 15 = 135$$

Com a retirada de três elementos, passamos a ter 12 dados, e sua soma representada por:

$$x_1 + \cdots x_{12} = 135 - 34 - 27 - 14 = 60$$

Aplicando a definição de média, temos:

$$\bar{x} = \frac{x_1 + \cdots x_{12}}{12} = \frac{60}{12} = 5$$

Exemplo

Aplicou-se uma prova para 80 alunos da turma da disciplina de Estatística. Porém, como o espaço físico era pequeno, dividiu-se a turma em duas partes, que realizaram a prova em dias diferentes. No primeiro dia, 35 alunos realizaram a avaliação, e a média desse grupo foi 9,0. No segundo dia, aplicou-se a prova para os demais, que obtiveram média igual a 7,0. Qual foi a média da turma toda?

Solução:
Podemos representar a média da turma do primeiro dia como:

$$\bar{x}_1 = \frac{x_1 + \cdots x_{35}}{35} = 9$$

bem como a média da segunda turma é:

$$\bar{x}_2 = \frac{x_1 + \cdots x_{45}}{45} = 7$$

$$x_1 + \cdots x_{35} = 9 \cdot 35 = 315$$

$$x_1 + \cdots x_{45} = 7 \cdot 45 = 315$$

$$x_1 + \cdots x_{80} = 315 + 315 = 630$$

Portanto, a média final é igual a:

$$\bar{x}_f = \frac{x_1 + \cdots x_{80}}{80} = \frac{630}{80} = 7{,}87$$

Exemplo

Uma loja de roupas está promovendo um bazar de suas peças e fez a seguinte promoção:

- 2 blusas custam R$ 89,00 cada;
- 4 blusas custam R$ 68,00 cada;
- 6 blusas custam R$ 57,00 cada.

Qual é o preço médio das blusas desta loja no seu bazar?

Solução:

Os valores expostos na promoção nos fornecem a seguinte relação:

$$\bar{x} = \frac{(2 \cdot 89{,}00) + (4 \cdot 68{,}00) + (6 \cdot 57{,}00)}{12} = \frac{792}{12} = 66{,}00$$

Concluímos, assim, que o preço médio de cada blusa é igual a R$ 66,00.

Os próximos exemplos da aplicação da média são exercícios adaptados de concursos de vestibular, que mostram variações no raciocínio utilizado para empregar o cálculo da média.

Exemplo

(FUVEST) Sabe-se que a média aritmética de 5 dados, sendo esses números inteiros distintos, estritamente positivos, é igual a 16. O maior valor existente entre esses dados é igual a:

a) 16
b) 20
c) 50
d) 70
e) 100

Solução:

Como indicado, o conjunto tem cinco elementos. Assim, da mesma maneira das soluções anteriores, temos:

$$\bar{x} = \frac{x_1 + \cdots x_5}{5} = 16$$

Portanto, a soma de todos os 5 elementos do conjunto de dados é dada por:

$$x_1 + \cdots x_5 = 16 \cdot 5 = 80$$

Então, para descobrirmos o maior valor possível entre os 5 dados, assumiremos os 4 outros valores como os menores possíveis, ou seja:

$$1 + 2 + 3 + 4 + x = 80$$

Sendo assim, o maior valor possível do conjunto de dados é:

$$x = 80 - 1 - 2 - 3 - 4$$

$$x = 70$$

Resposta: letra D.

Exemplo

(FUVEST) Numa classe com vinte alunos, as notas do exame final podiam variar de 0 a 100, e a nota mínima para aprovação era 70. Realizado o exame, verificou-se que 8 alunos foram reprovados. A média aritmética das notas desses oito alunos foi 65, enquanto que a média dos aprovados foi 77. Após a divulgação dos resultados, o professor verificou que uma questão havia sido mal formulada e decidiu atribuir 5 pontos a mais para todos os alunos. Com essa decisão, a média dos aprovados passou a ser 80, e a dos reprovados, 68,8.

a) Calcule a média aritmética das notas da classe toda antes da atribuição dos cinco pontos extras.

b) Com a atribuição dos cinco pontos extras, quantos alunos, inicialmente reprovados, atingiram nota para a aprovação?

Solução:

a) Com os dados informados no problema, temos:

$$\bar{x} \text{ reprovados} = \frac{x_1 + \cdots x_8}{8} = 65$$

$$\bar{x} \text{ aprovados} = \frac{x_1 + \cdots x_{12}}{12} = 77$$

$$\bar{x} \text{ total} = \frac{(x_1 + \cdots x_8) + (x_1 + \cdots x_{12})}{20} = \frac{520 + 924}{20} = 72{,}2$$

A média das notas da classe antes da atribuição dos cinco pontos extras era de 72,2.

b) A nova média de toda a turma, após a atribuição dos cinco pontos por aluno, é:

$$x_1 + \cdots x_5 = 16 \cdot 5 = 80$$

$$\bar{x} = \frac{520 + 924 + (5 \cdot 20)}{20} = \frac{1544}{20} = 77{,}2$$

Com a atribuição dos cinco pontos, é possível que alguma quantidade de alunos tenha sido aprovada — chamemos essa quantidade de A. Sendo assim, a nova quantidade de alunos aprovados é 12 + A, e de alunos reprovados, 8 – A.

Temos, do enunciado, que a nova média dos aprovados é 80, e dos reprovados, 68,8. Então:

$$77{,}2 = \frac{(12 + A)\,80 + (8 - A)\,68{,}8}{20}$$

Resolvendo a equação, temos que A = 3.

Assim, 3 alunos foram aprovados após a atribuição dos 5 pontos.

Referências

BARBETTA, P. A.; REIS, M. M.; BORNIA, A. C. *Estatística para cursos de engenharia e informática*. 2. ed. São Paulo: Atlas, 2008.

CRESPO, A. A. *Estatística fácil*. 17. ed. São Paulo: Saraiva, 2002.

MILONE, G. *Estatística*: geral e aplicada. São Paulo: Thomson Learning, 2006.

Leitura recomendada

BECKER, J. L. *Estatística básica*: transformando dados em informação. Porto Alegre: Bookman, 2015.

Conteúdo:

SAGAH
SOLUÇÕES EDUCACIONAIS INTEGRADAS