Medidas de tendência central

Média, mediana e moda

Uma **medida da tendência central** é um valor que representa uma entrada típica ou central do conjunto de dados. As três medidas da tendência central mais comumente usadas são a média, a mediana e a moda.

Definição

A **média** de um conjunto de dados é a soma das entradas de dados dividida pelo número de entradas. Para encontrar a média de um conjunto de dados, use uma das fórmulas a seguir.

Média populacional: $\mu = \frac{\sum x}{N}$

Média amostral: $\bar{x} = \frac{\sum x}{n}$

Encontrando a média amostral

Os preços (em dólares) para uma amostra de voos de ida e volta partindo de Chicago, Illinois, para Cancun, México, são listados a seguir. Qual é a média dos preços dos voos?

872 432 397 427 388 782 397

A soma dos preços dos voos é:

$$\sum x = 872 + 432 + 397 + 427 + 388 + 782 + 397 = 3.695.$$

$$\bar{x} = \frac{\sum x}{n} = \frac{3.695}{7} \approx 527,9.$$

a média dos preços dos voos é aproximadamente \$ 527,90

Definição

A **mediana** de um conjunto de dados é um valor que está no meio dos dados quando o conjunto de dados é ordenado. A mediana mede o centro de um conjunto de dados ordenado dividindo-se em duas partes iguais. Se o conjunto de dados tem um número ímpar de entradas, a mediana é a entrada de dados do meio. Se o conjunto de dados tem um número par de entradas, a mediana é a média das duas entradas do meio.

Encontrando a mediana

Encontre a mediana dos preços dados no Exemplo 1.

Solução

Para encontrar a mediana dos preços, primeiro ordene os dados.

388 397 397 427 432 782 872

a mediana dos preços dos voos é $ 427.

Encontrando a mediana

No Exemplo 2, o preço de voo de \$ 432 não está mais disponível. Qual é a mediana dos preços restantes?

Solução

Os preços restantes, em ordem, são 388, 397, 397, 427, 782 e 872.

$$\text{Mediana} = \frac{397 + 427}{2} = 412$$

a mediana dos preços dos voos restantes é \$ 412.

Dica de estudo

Em um conjunto de dados, há o mesmo número de valores de dados acima da mediana bem como abaixo da mediana. Por exemplo, no Exemplo 2, três dos preços estão abaixo de \$ 427 e três estão acima de \$ 427.

Definição

A moda de um conjunto de dados é uma entrada do conjunto de dados que ocorre com a maior frequência. Se nenhuma entrada é repetida, o conjunto de dados não tem moda. Se duas entradas ocorrem com a mesma frequência, cada entrada é uma moda e o conjunto é chamado de **bimodal**.

Encontrando a moda

Em um debate político, pede-se que uma amostra dos membros da plateia nomeie o partido político ao qual pertencem. Suas respostas são mostradas na tabela. Qual é a moda da respostas?

Partido político	Frequência *f*
Democratas	34
Republicanos	56
Outro	21
Não responderam	9

Solução

A resposta que ocorre com maior frequência é Partido Republicano. Então, a moda é Partido Republicano.

Interpretação

Nessa amostra, havia mais republicanos do que pessoas de qualquer outra afiliação.

Embora a média, a mediana e a moda descrevam, cada uma, determinada entrada típica dos dados, há vantagens e desvantagens no uso de cada uma delas. A média é uma medição confiável, pois leva em conta cada entrada dos dados, mas pode ser muito afetada quando o conjunto de dados tem valores discrepantes.

Definição

Um **valor discrepante** (outlier) é uma entrada de dados que está muito afastada das outras entradas em um conjunto de dados.

Um conjunto de dados pode ter um ou mais valores discrepantes, causando **lacunas** em uma distribuição. As conclusões que são tomadas de um conjunto de dados que contém valores discrepantes podem conter falhas.

Exemplo 6

Comparando a média, a mediana e a moda

Encontre a média, a mediana e a moda da amostra das idades da turma à esquerda. Qual medida da tendência central melhor descreve uma entrada típica desse conjunto de dados? Há valores discrepantes?

Idades da turma
20 20 20 20 20 20 21
21 21 21 22 22 22 23
23 23 23 24 24 65

Média: $\bar{x} = \frac{\sum x}{n} = \frac{475}{20} \approx 23{,}8$ anos

Mediana: $\text{Mediana} = \frac{21+22}{2} = 21{,}5$ anos

Moda: A entrada que ocorre com maior frequência é 20 anos.

Interpretação

A média leva todas as entradas em consideração, mas é influenciada pelo valor discrepante de 65. A mediana também leva em consideração todas as entradas e não é afetada pelo valor discrepante. Nesse caso, a moda existe, mas não parece representar uma entrada típica. Algumas vezes, a comparação gráfica pode ajudar a decidir qual medida de tendência central melhor representa o conjunto de dados. O histograma mostra a distribuição dos dados e a localização da média, da mediana e da moda. Nesse caso, parece que é a mediana que melhor descreve o conjunto de dados.

Média ponderada e média de dados agrupados

Às vezes, os dados contêm entradas que têm um maior efeito na média do que outras. Para encontrar a média de tais conjuntos de dados, você deve encontrar a média ponderada.

Definição

Uma **média ponderada** é a média de um conjunto de dados cujas entradas têm pesos variados. Uma média ponderada é dada por:

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

onde w é o peso de cada entrada x.

Encontrando a média ponderada

Você está frequentando uma aula na qual sua nota é determinada com base em 5 fontes: 50% da média de seu exame, 15% do seu exame bimestral, 20% de seu exame final, 10% de seu trabalho no laboratório de informática e 5% de seus deveres de casa. Suas notas são: 86 (média do exame), 96 (exame bimestral), 82 (exame final), 98 (laboratório) e 100 (dever de casa). Qual é a média ponderada de suas notas? Se a média mínima para um A é 90, você obteve uma nota A?

Fonte	Nota, x	Peso, w	xw
Média do exame	86	0,50	43,0
Exame bimestral	96	0,15	14,4
Exame final	82	0,20	16,4
Laboratório	98	0,10	9,8
Dever de casa	100	0,05	5,0
		$\Sigma w = 1$	$\Sigma (x \cdot w) = 88{,}6$

A média ponderada para o curso é 88,6. Portanto, você não obteve um A.

Definição

A média de uma distribuição de frequência para uma amostra é aproximada por:

$$\bar{x} = \frac{\sum(x \cdot f)}{n}, \qquad \text{Note que } n = \sum f$$

em que *x* e *f* são os pontos médios e as frequências de uma classe, respectivamente.

Instruções

Encontre a média da distribuição de frequência

Em palavras	*Em símbolos*
1. Encontre o ponto médio de cada classe.	**1.** $x = \frac{(\text{limite inferior}) + (\text{limite superior})}{2}$
2. Encontre a soma dos produtos dos pontos médios e as frequências.	**2.** $\Sigma\,(x \cdot f)$
3. Encontre a soma das frequências.	**3.** $n = \Sigma f$
4. Encontre a média da distribuição de frequência.	**4.** $\bar{x} = \frac{\sum(x \cdot f)}{n}$

Encontrando a média da distribuição de frequência

Use a distribuição de frequência à direita para aproximar o número médio de minutos que uma amostra de usuários de Internet gasta on-line durante sua mais recente sessão.

Ponto médio das classes x	Frequência f	$(x \cdot f)$
12,5	6	75,0
24,5	10	245,0
36,5	13	474,0
48,5	8	388,0
60,5	5	302,0
72,5	6	435,0
84,5	2	169,0
	$n = 50$	$\Sigma = 2.089{,}0$

Solução

$$\bar{x} = \frac{\sum(x \cdot f)}{n} = \frac{2.089}{50} \approx 41{,}8$$

média do tempo gasto on-line foi de aproximadamente 41,8 minutos.

Forma das distribuições

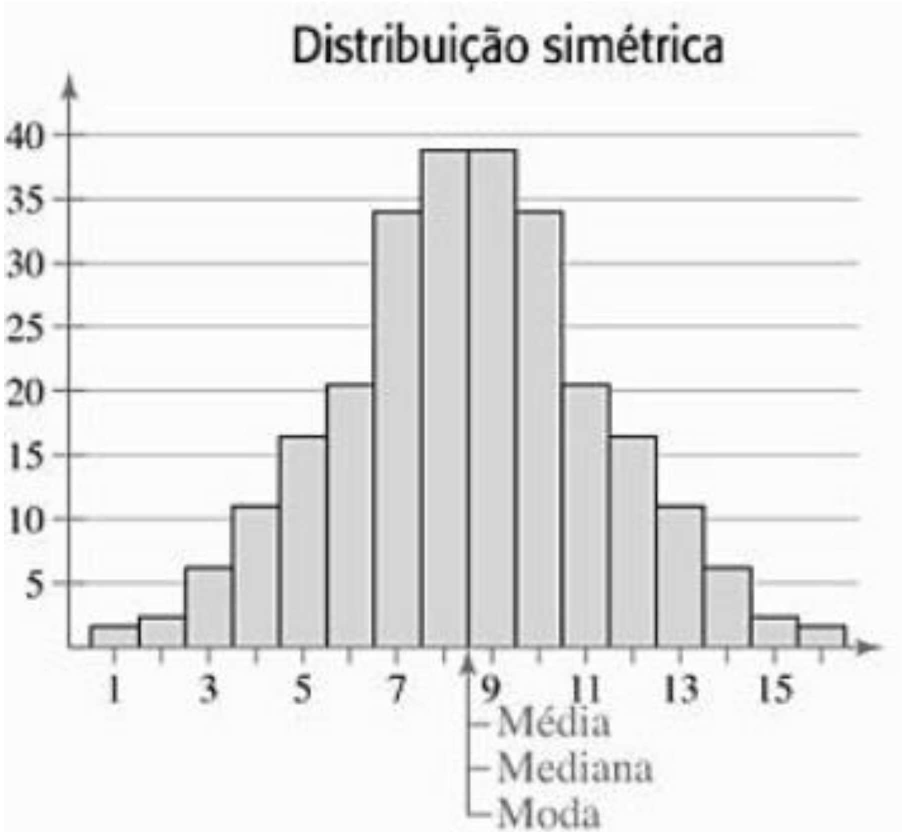
Um gráfico revela diversas características de uma distribuição de frequência. Uma delas é a forma das distribuições.

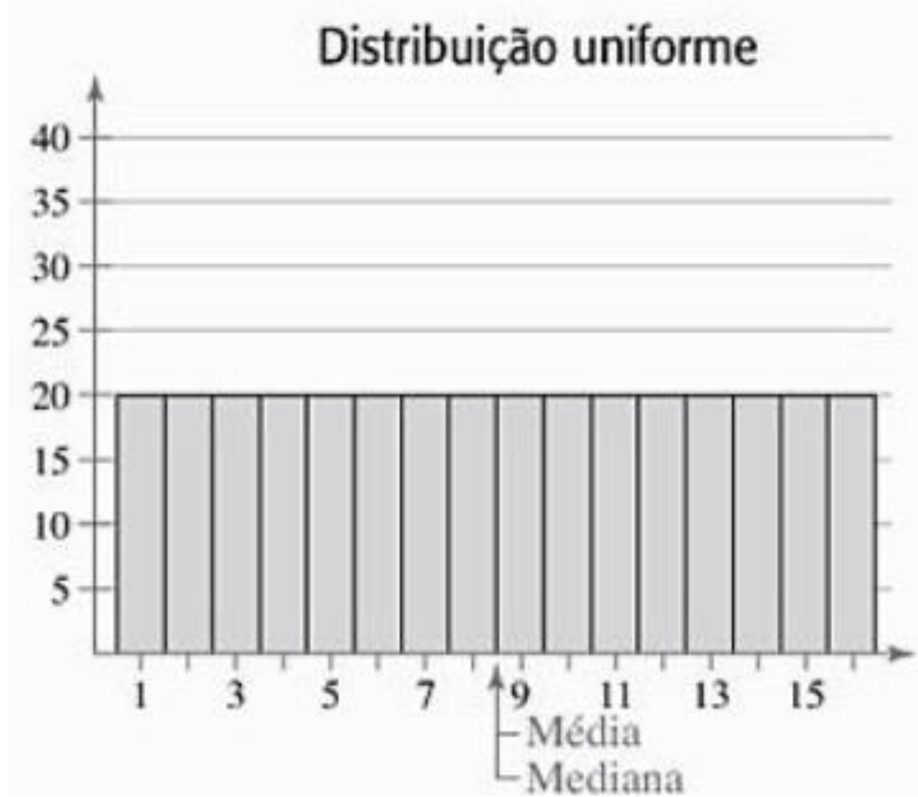
Definição

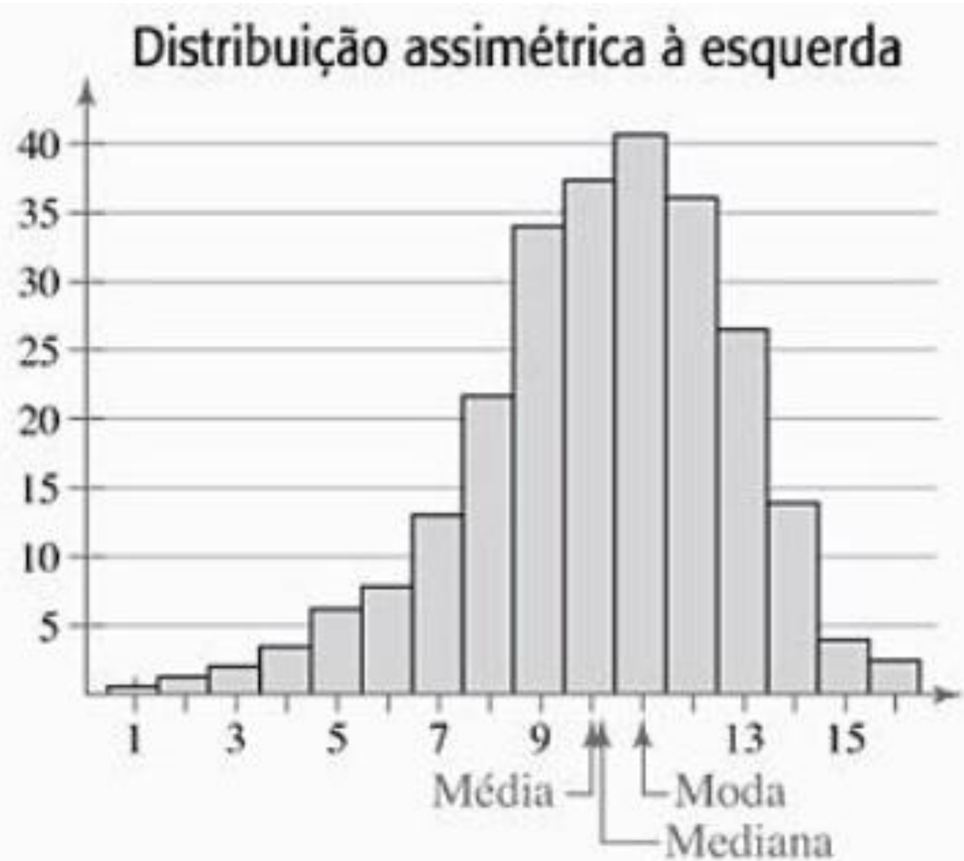
Uma distribuição de frequência é **simétrica** quando a linha vertical pode ser desenhada do meio do gráfico da distribuição e as metades resultantes são aproximadamente imagens espelhadas.

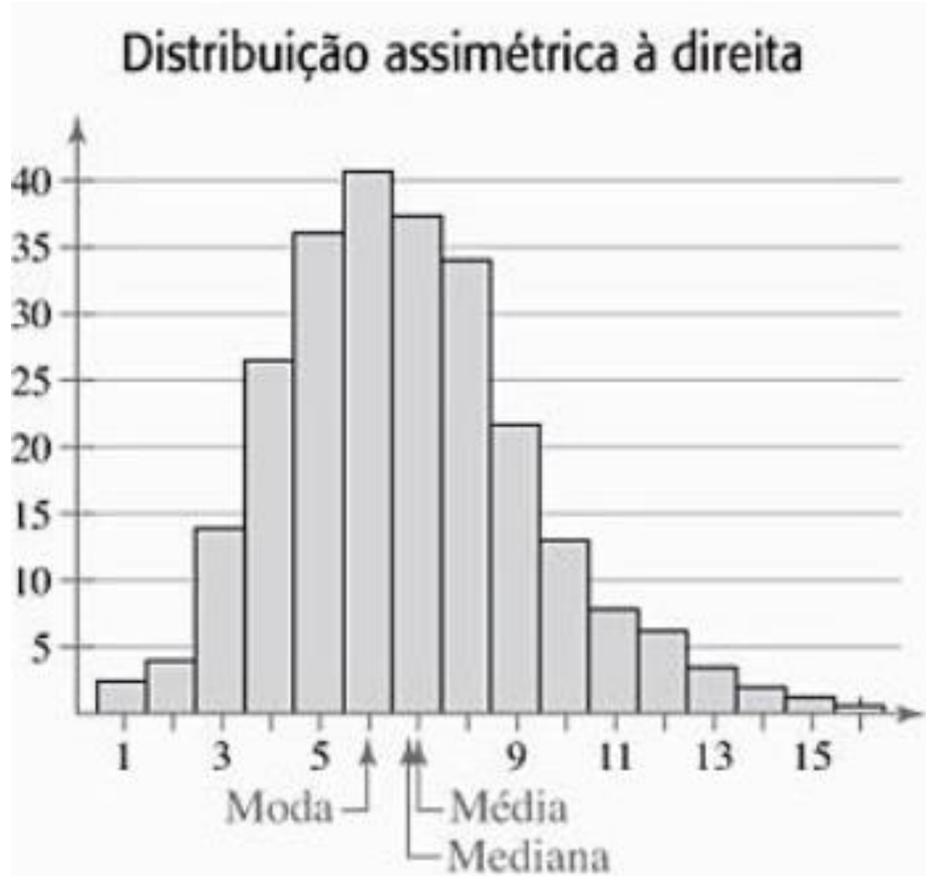
Uma distribuição de frequência é **uniforme** (ou **retangular**) quando todas as entradas, ou classes, na distribuição têm frequências iguais ou aproximadamente iguais. Uma distribuição uniforme também é simétrica.

Uma distribuição de frequências é assimétrica se a "cauda" do gráfico se alonga mais em um dos lados. Uma distribuição é **assimétrica à esquerda (negativamente assimétrica)** se a cauda se estende à esquerda, e **assimétrica à direita (positivamente assimétrica)** se a cauda se estende à direita.


Distribuição simétrica
40
35
30
25
20
15
10
5
1
3
5
7
9
11
13
15
Média
Mediana
Moda


Distribuição uniforme
40
35
30
25
20
15
10
5
1
3
5
7
9
11
13
15
Média
Mediana





A média sempre irá na direção em que a distribuição for assimétrica. Por exemplo, quando a distribuição é assimétrica à esquerda, a média está à esquerda da mediana.