

# Medidas de posição

## Quartis

Nesta seção, você aprenderá como usar os fractis para especificar a posição de uma entrada de dados dentro de um conjunto de dados. Fractis são números que separam, ou dividem, um conjunto de dados ordenado em partes iguais. Por exemplo, a mediana é um fractil porque divide um conjunto de dados ordenados em duas partes iguais.

## Definição

Os três quartis,  $Q_1$ ,  $Q_2$  e  $Q_3$ , dividem aproximadamente um conjunto de dados ordenado em quatro partes iguais. Aproximadamente  $1/4$  dos dados está acima ou abaixo do primeiro quartil  $Q_1$ . Aproximadamente metade dos dados está acima ou abaixo do segundo quartil  $Q_2$  (o segundo quartil é o mesmo que a mediana do conjunto de dados). Aproximadamente  $3/4$  dos dados estão acima ou abaixo do terceiro quartil  $Q_3$ .

## Encontrando os quartis de um conjunto de dados

As notas dos testes de 15 funcionários matriculados em um curso de treinamento de CPR são listadas a seguir. Encontre o primeiro, o segundo e o terceiro quartis das notas dos testes.

13 9 18 15 14 21 7 10 11 20 5 18 37 16 17

### Solução

Primeiro, ordene o conjunto de dados e encontre a mediana  $Q_2$ . Depois de encontrar  $Q_2$ , divida o conjunto de dados em duas metades. O primeiro e o terceiro quartil são as medianas das metades inferior e superior do conjunto de dados.

<i>Metade inferior</i>							<i>Metade superior</i>							
5	7	9	10	11	13	14	15	16	17	18	18	20	21	37
			↑				↑				↑			
			$Q_1$				$Q_2$				$Q_3$			

### Interpretação

Aproximadamente  $1/4$  dos funcionários obteve 10 ou menos; aproximadamente metade obteve 15 ou menos; e aproximadamente  $3/4$  obtiveram 18 ou menos.

Depois de encontrar os quartis dos conjuntos de dados, você pode encontrar a amplitude interquartil.

## Definição

---

A amplitude interquartil (IQR) de um conjunto de dados é a diferença entre o terceiro e o primeiro quartil.

$$\text{Amplitude interquartil (IQR)} = Q_3 - Q_1$$

## Encontrando a amplitude interquartil

Encontre a amplitude interquartil das 15 notas de testes do Exemplo 1. O que podemos concluir com o resultado?

### Solução

Pelo Exemplo 1, sabemos que  $Q_1 = 10$  e  $Q_3 = 18$ . Então, a amplitude interquartil é:

<i>Metade inferior</i>							<i>Metade superior</i>							
5	7	9	10	11	13	14	15	16	17	18	18	20	21	37
			↑				↑				↑			
			$Q_1$				$Q_2$				$Q_3$			

$$\text{IQR} = Q_3 - Q_1 = 18 - 10 = 8.$$

### Interpretação

As notas do teste na porção da metade do conjunto de dados variam no máximo 8 em pontos.

O IQR é uma medida de variação que fornece uma ideia de quanto 50% dos dados varia. Também pode ser usado para identificar valores discrepantes. Qualquer valor de dado que seja maior que 1,5 IQRs à esquerda de  $Q_1$  ou à direita de  $Q_3$  é um valor discrepante. Por exemplo, o IQR no Exemplo 1 é  $18 - 10 = 8$ . Então, 1,5 IQRs à direita de  $Q_3$  é  $Q_3 + 1,5(8) = 18 + 12 = 30$ . Em razão de  $37 > 30$ , 37 é um valor discrepante.

Outra importante aplicação dos quartis é representar conjuntos de dados usando o gráfico caixa-e-bigodes. Um gráfico caixa-e-bigodes é uma ferramenta de análise de dados exploratória que enfatiza as características mais importantes de um conjunto de dados. Para representar graficamente um gráfico caixa-e-bigodes, você deve saber os valores a seguir.

1. A entrada mínima

2. O primeiro quartil  $Q_1$

3. A mediana  $Q_2$

4. O terceiro quartil  $Q_3$

5. A entrada máxima

Regra dos cinco itens de um conjunto de dados.

# Instruções

## Desenhando um gráfico caixa-e-bigodes

1. Encontre a regra dos cinco itens do conjunto de dados.
2. Construa uma escala horizontal que transpasse a amplitude dos dados.
3. Represente os cinco números sobre a escala horizontal.
4. Desenhe uma caixa acima da escala horizontal a partir de  $Q_1$  para  $Q_3$  e desenhe uma linha vertical na caixa em  $Q_2$
5. Desenhe os bigodes a partir da caixa para as entradas mínimas e máximas.



## Desenhando um gráfico caixa-e-bigodes

Desenhe um gráfico caixa-e-bigodes que represente as 15 notas de testes dadas no Exemplo 1. O que podemos concluir com o gráfico?

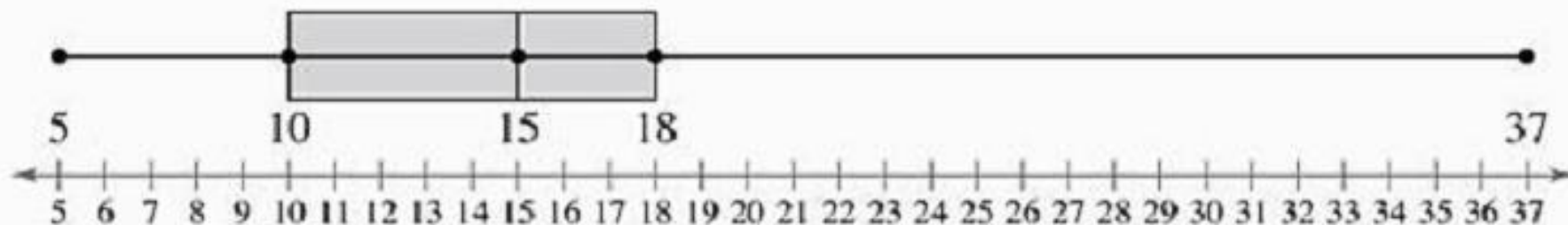
<i>Metade inferior</i>							<i>Metade superior</i>							
5	7	9	10	11	13	14	15	16	17	18	18	20	21	37
			↑				↑				↑			
			$Q_1$				$Q_2$				$Q_3$			

### Solução

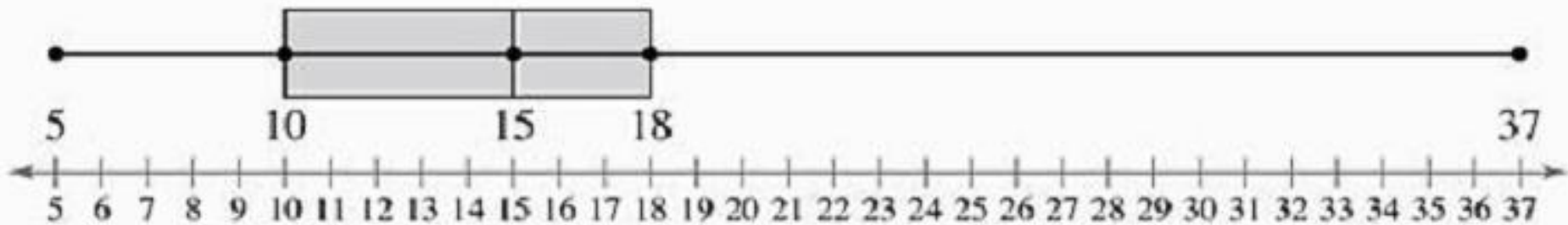
A regra dos cinco itens é dada a seguir. Usando-a, pode-se construir o gráfico caixa-e-bigodes mostrado.

$$\text{Mín} = 5 \quad Q_1 = 10 \quad Q_2 = 15 \quad Q_3 = 18 \quad \text{Máx} = 37.$$

Notas dos testes na aula de CPR



## Notas dos testes na aula de CPR



### *Interpretação*

Você pode tirar diversas conclusões com o gráfico. Uma delas é que aproximadamente metade das notas está entre 10 e 18. Olhando para o comprimento do bigode de direito, podemos concluir também que a nota 37 é um possível valor discrepante.



## Percentis e outros fractis

Além de usar quartis para especificar uma medida de posição, podemos também usar os percentis e os decis. Esses fractis comuns são resumidos a seguir.

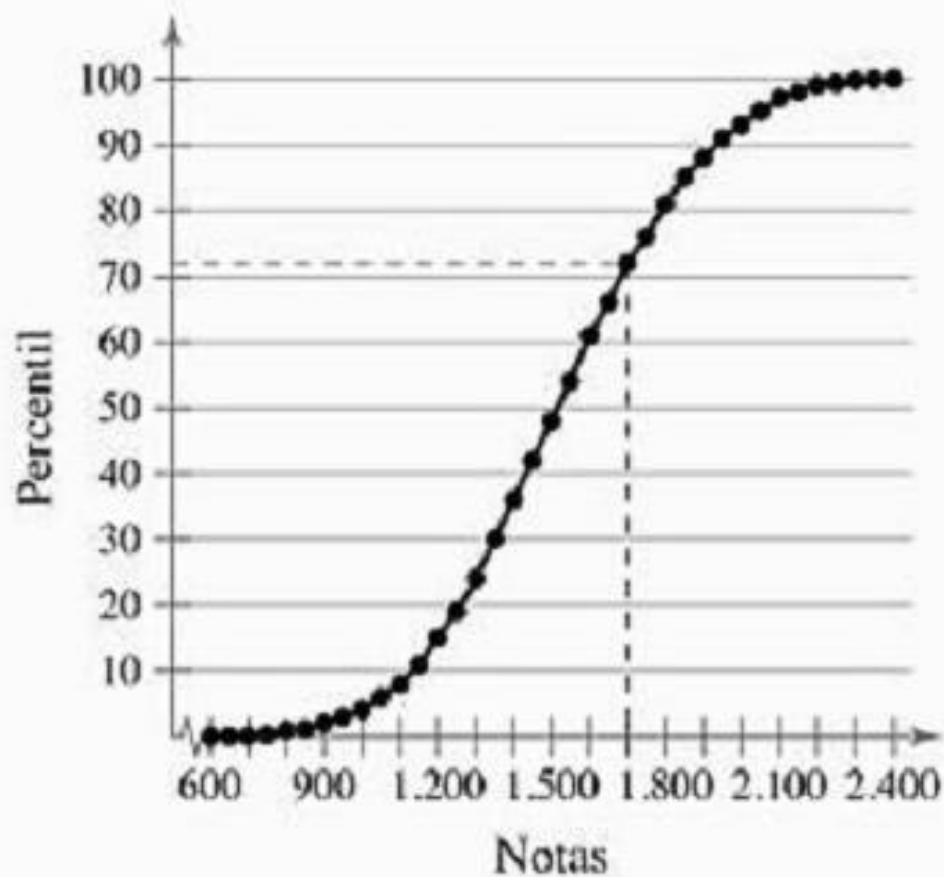
Fractil	Resumo	Símbolos
Quartis	Divide um conjunto de dados em 4 partes iguais	$Q_1, Q_2, Q_3$
Decis	Divide um conjunto de dados em 10 partes iguais	$D_1, D_2, D_3, \dots, D_9$
Percentis	Divide um conjunto de dados em 100 partes iguais	$P_1, P_2, P_3, \dots, P_{99}$

Os percentis são geralmente usados nas áreas relacionadas à saúde e educação para indicar como um indivíduo se compara a outros em um conjunto. Eles também podem ser usados para identificar valores excepcionalmente altos ou baixos. Por exemplo, as notas de um teste e as medidas de crescimento de crianças são normalmente expressas em percentis. Notas ou medidas no 95º percentil ou acima são excepcionalmente altas, enquanto aquelas no 5º percentil ou abaixo são excepcionalmente baixas.

## Interpretando percentis

A ogiva representa a distribuição de frequência acumulada para as notas em um teste SAT de alunos pré-universitários. Qual nota representa o 72º percentil? Como devemos interpretar esse resultado? (Fonte: *College Board On-line.*)

### Notas do teste SAT



### Interpretação

Isso significa que 72% dos estudantes teve uma nota de 1.700 ou menos.

## Dica de estudo

É importante entender o que significa percentil. Por exemplo, se o peso de uma criança de 6 meses de idade está no 78º percentil, a criança pesa mais do que 78% de todas as crianças de mesma idade. Isso não significa que a criança pese 78% de algum peso ideal.

## Escore padrão

---

Quando sabemos a média e o desvio padrão de um conjunto de dados, podemos medir a posição do valor de um dado no conjunto de dados com um escore padrão ou z-escore.

## Definição

---

O **escore padrão** ou **z-escore** representa o número de desvios padrão que um valor dado  $x$  está a partir da média  $\mu$ . Para encontrar o **z-escore** para certo valor, use a fórmula a seguir:

$$z = \frac{\text{valor} - \text{média}}{\text{desvio padrão}} = \frac{x - \mu}{\sigma}$$

Um z-escore pode ser negativo, positivo ou zero. Se  $z$  for negativo, o valor  $x$  correspondente está abaixo da média. Se  $z$  for positivo, o valor  $x$  correspondente está acima da média. E se  $z = 0$ , o valor  $x$  correspondente é igual a média. Um z-escore pode ser usado para identificar valores incomuns do conjunto de dados que seja aproximadamente em forma de sino.

## Encontrando z-escores

A velocidade média de veículos em uma reta de uma rodovia é de 56 milhas por hora (mhp), com desvio padrão de 4 milhas por hora. Você mede a velocidade de três carros que estão passando pela reta da rodovia como 62 milhas por horas, 47 milhas por hora e 56 milhas por hora. Encontre o z-escore que corresponde a cada velocidade. O que podemos concluir?

### Solução

O z-escore que corresponde a cada velocidade é calculado a seguir.

$$x = 62 \text{ mhp}$$

$$x = 47 \text{ mhp}$$

$$x = 56 \text{ mhp}$$

$$z = \frac{62 - 56}{4} = 1,5$$

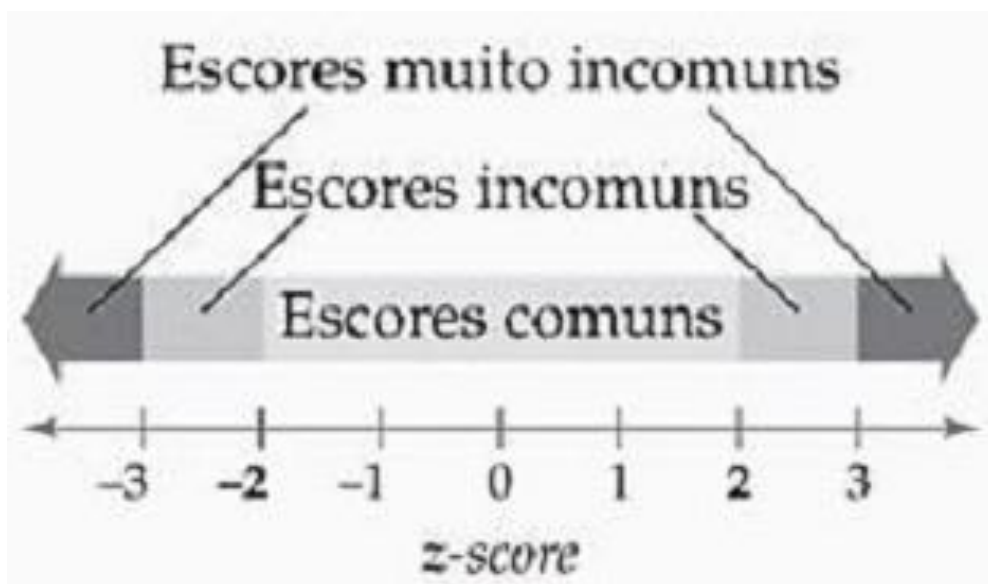
$$z = \frac{47 - 56}{4} = -2,25$$

$$z = \frac{56 - 56}{4} = 0$$

### Interpretação

A partir dos z-escores, podemos concluir que a velocidade de 62 mhp está 1,5 desvio padrão acima da média; a velocidade 47 mhp está 2,25 desvios padrão abaixo da média e a velocidade 56 mhp é igual a média. Se a distribuição das velocidades é aproximadamente em forma de sino, o carro que viaja a 47 mhp está excepcionalmente devagar, pois sua velocidade corresponde a um z-escore de  $-2,25$ .

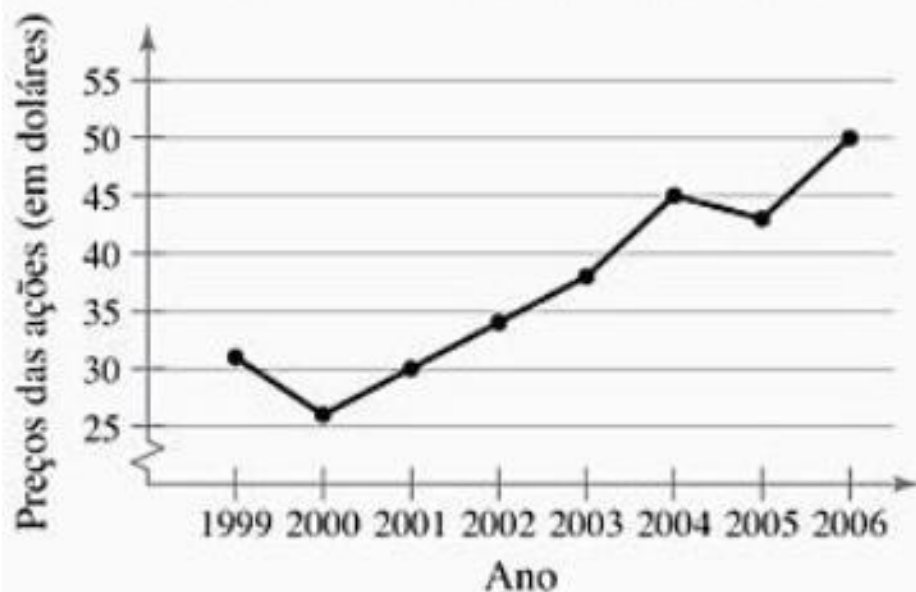
Quando a distribuição é aproximadamente em forma de sino, sabemos, pela regra empírica, que aproximadamente 95% dos dados está dentro de 2 desvios padrão da média. Então, quando esses valores de distribuição são transformados em z-escores, aproximadamente 95% dos z-escores deveriam estar entre  $-2$  e  $2$ . Um z-escore fora dessa amplitude ocorrerá aproximadamente 5% do tempo e será considerado incomum. Então, de acordo com a regra empírica, um z-escore menor do que  $-3$  ou maior que  $3$  será *muito* incomum, com tal score ocorrendo 0,3% do tempo.



# Abusos

Saber como as estatísticas são calculadas pode ajudá-lo a analisar estatísticas questionáveis. Por exemplo, suponha que você esteja entrevistando para a posição de vendas e a empresa reporta que a comissão média anual recebida por cinco pessoas de sua equipe de vendas é \$ 60.000. Esta é uma afirmação equivocada se for baseada em quatro comissões de \$ 25.000 e uma de \$ 200.000. A mediana descreveria mais corretamente a comissão anual, mas a empresa usou a média, pois é uma quantia maior.

## Preço das ações da Kellogg



## Preço das ações da Kellogg

