

## CONHEÇA O VOLUME

### Objetivos a serem desenvolvidos neste Volume

Ao realizar os estudos propostos neste livro da Coleção, esperamos que os objetivos apresentados a seguir sejam alcançados.

- Compreender e calcular medidas de tendência central para dados agrupados ou não agrupados e identificar qual dessas medidas melhor representa o conjunto dos dados obtidos em uma pesquisa.
- Compreender e calcular medidas de dispersão para analisar um conjunto de dados em relação à sua média aritmética.
- Reconhecer quando uma pesquisa é censitária e quando é amostral.
- Compreender ideias de algumas técnicas de amostragem e definir qual é a mais adequada para obter uma amostra de acordo com a natureza da população e das variáveis investigadas.
- Realizar pesquisas estatísticas amostrais, com dados primários ou secundários, para investigar questões relevantes e utilizar gráficos, tabelas e medidas de tendência central e de dispersão para interpretar e divulgar os resultados obtidos, com ou sem auxílio de planilha eletrônica.
- Compreender a ideia dos princípios multiplicativo e aditivo da contagem.
- Construir árvores de possibilidades e tabelas de dupla entrada para representar as possibilidades de resultados possíveis em um experimento.
- Compreender os conceitos de arranjo, combinação e permutação e identificar qual deles utilizar para resolver um problema.
- Reconhecer o uso de cálculos de probabilidade em atividades e estudos de diversas áreas do conhecimento e situações do cotidiano.
- Determinar o espaço amostral de um experimento aleatório.
- Calcular a probabilidade de ocorrência ou não de um evento, considerando um espaço amostral equiprovável.
- Calcular a probabilidade da união de dois eventos e a probabilidade condicional.
- Identificar se eventos de um mesmo espaço amostral são dependentes ou independentes entre si.
- Calcular probabilidades em experimentos binomiais.
- Estimar a probabilidade de ocorrência de um evento a partir de dados estatísticos.
- Resolver e elaborar problemas, individualmente ou em grupo, envolvendo medidas de tendência central, medidas de dispersão, realização de pesquisas estatísticas amostrais, princípio multiplicativo e aditivo da contagem, arranjo, combinação, permutação e cálculo de probabilidades, relacionados ou não a situações do cotidiano.

## Justificativa da pertinência dos objetivos

Os objetivos citados anteriormente são fundamentais para o desenvolvimento do raciocínio lógico, para entender criticamente a realidade e contribuir na construção de uma sociedade justa e inclusiva, para acessar e divulgar informações por diferentes meios e recursos digitais, e para sua formação como cidadão crítico, reflexivo e atuante, que investiga, argumenta, promove o respeito e propõe soluções, contribuindo em suas práticas sociais, individual ou coletiva.

O planejamento e a realização de pesquisas para investigar questões relevantes, sejam elas voltadas a demandas da região em que mora, ao mercado de trabalho etc., possibilitam a mobilização e a articulação de conceitos matemáticos, como a técnica de amostragem mais adequada, o cálculo e a interpretação de medidas de tendência e de dispersão e o tipo de gráfico mais adequado para representar os resultados obtidos. Com isso, você tem aporte para tomar decisões responsáveis relacionadas à questão investigada. O processo de tratamento das informações e elaboração de relatórios possibilita também a utilização de recursos tecnológicos e de diferentes linguagens.

O trabalho com técnicas de contagem contribui para a construção de significados de ideias próprias da Matemática que podem ser utilizadas para interpretar e resolver problemas em diferentes contextos, por exemplo, ao fazer o levantamento das possibilidades de escolha existentes em determinada situação. Já o estudo de probabilidades pode ser identificado e aplicado em diferentes situações cotidianas, em que seja necessário analisar os riscos envolvidos e/ou escolher uma opção, entre outras disponíveis, levando em consideração a probabilidade atrelada a determinado evento.

Os conceitos estudados poderão ser usados para elaborar e resolver problemas e analisar e verificar os resultados obtidos, aqui ou em situações com as quais você possa se deparar futuramente.



# Estatística: pesquisa e medidas de posição e de dispersão

## Conectividade

O mundo virtual é um espaço de expressão e de descobertas. Por meio das redes sociais, que fazem parte das tecnologias da informação e da comunicação, podemos nos conectar a pessoas em um ambiente virtual. Além de entretenimento, esses ambientes virtuais possibilitam o acesso a informações por meio de noticiários e de *sites* que ajudam na formação educacional, por exemplo. O modo como lidamos com a quantidade de informação disponível, como nos posicionamos e como nos expressamos nesses ambientes ajudam a construir a nossa identidade.

É importante ter clareza que, apesar dos benefícios, são necessários diversos cuidados. É fundamental que tenhamos bom senso quando estamos utilizando a internet, além de ficarmos atentos ao uso excessivo, ou seja, ao tempo que ficamos conectados e ao tipo de informação que estamos acessando. Ficar muitas horas na internet pode comprometer a qualidade de vida e desencadear alguns problemas de saúde.

» Grupo de estudantes acessando diferentes tipos de mídia (digital e impressa). Jovens de 16 a 24 anos são os usuários que passam mais tempo nas redes sociais.

RAWPIXEL.COM/SHUTTERS TOOK.COM

### COMPETÊNCIAS E HABILIDADES DA BNCC:

Competências gerais: 5 e 6

Matemática e suas Tecnologias

Competências específicas: 2 e 3

Habilidades: EM13MAT202 e EM13MAT316

Competências da Natureza e suas Tecnologias

Competência específica: 3

O texto integral das competências e das habilidades citadas encontra-se no final deste livro do estudante.

## Após ler as informações, converse com os colegas e o professor sobre os itens abaixo.

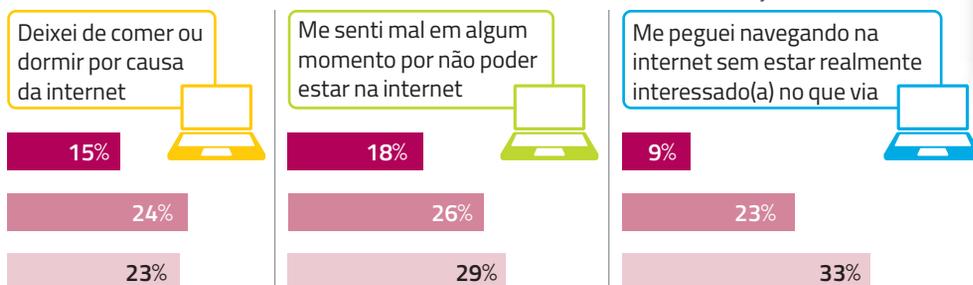
- Você costuma utilizar recursos tecnológicos digitais em seu dia a dia? Com que frequência?
- Em sua opinião, quais cuidados devem ser tomados ao navegar em redes sociais?
- Qual é a média de minutos que os estudantes brasileiros gastam por dia na internet, fora da escola? Em sua opinião, o que essa medida representa no contexto apresentado? *Veja os comentários sobre a abundância desses itens nas Orientações para o professor.*

Leia o texto a seguir, extraído do manual de orientação **Saúde de crianças e adolescentes na era digital**, publicado pela Sociedade Brasileira de Pediatria (SBP).

Estudos científicos comprovam que a tecnologia influencia comportamentos através do mundo digital, modificando hábitos desde a infância, que podem causar prejuízos e danos à saúde. O uso precoce e de longa duração de jogos *online*, redes sociais ou diversos aplicativos com filmes e vídeos na Internet pode causar dificuldades de socialização e conexão com outras pessoas e dificuldades escolares; a dependência ou o uso problemático e interativo das mídias causa problemas mentais, aumento da ansiedade, violência, *cyberbullying*, transtornos de sono e alimentação, sedentarismo, problemas auditivos por uso de *headphones*, problemas visuais, problemas posturais e lesões de esforço repetitivo (LER); problemas que envolvem a sexualidade [...]; compra e uso de drogas, pensamentos ou gestos de autoagressão e suicídio; além das “brincadeiras” ou “desafios” *online* que podem ocasionar consequências graves [...].

SBP. **Saúde de crianças e adolescentes na era digital**. Rio de Janeiro, 2016. Disponível em: [www.sbp.com.br/fileadmin/user\\_upload/2016/11/19166d-MOrient-Saude-Crian-e-Adolesc.pdf](http://www.sbp.com.br/fileadmin/user_upload/2016/11/19166d-MOrient-Saude-Crian-e-Adolesc.pdf). Acesso em: 27 maio 2020.

### » Porcentual de crianças e adolescentes, por situações vivenciadas ao uso excessivo da internet no Brasil, em 2017



#### Legenda:

- De 11 a 12 anos
- De 13 a 14 anos
- De 15 a 17 anos

Fonte dos dados: CETIC.BR. **TIC kids online Brasil 2018**. Disponível em: [https://cetic.br/media/docs/publicacoes/216370220191105/tic\\_kids\\_online\\_2018\\_livro\\_eletronico.pdf](https://cetic.br/media/docs/publicacoes/216370220191105/tic_kids_online_2018_livro_eletronico.pdf). Acesso em: 20 abr. 2020.

### » Média de minutos que os estudantes gastam por dia na internet, fora da escola



Fonte: FERREIRA, P. Brasil é o segundo país onde alunos passam mais tempo na internet nas horas vagas. **O Globo**, Rio de Janeiro, 19 abr. 2017. Disponível em: <https://oglobo.globo.com/sociedade/educacao/brasil-o-segundo-pais-onde-alunos-passam-mais-tempo-na-internet-nas-horas-vagas-21227360>. Acesso em: 28 maio 2020.

#### Conexões

Acesse este *site* para obter mais informações e dicas sobre o uso adequado de recursos tecnológicos para crianças e adolescentes:

- SBP. **Saúde de crianças e adolescentes na era digital**. Rio de Janeiro, 2016. Disponível em: [www.sbp.com.br/fileadmin/user\\_upload/2016/11/19166d-MOrient-Saude-Crian-e-Adolesc.pdf](http://www.sbp.com.br/fileadmin/user_upload/2016/11/19166d-MOrient-Saude-Crian-e-Adolesc.pdf). Acesso em: 27 maio 2020.



O trabalho com este tópico favorece, com maior ênfase, o desenvolvimento da competência específica 3 e da habilidade EM13MAT316 da área de Matemática e suas Tecnologias.

# Medidas de dispersão

Vimos que as medidas de posição, como a média, a moda e a mediana, podem ser utilizadas para representar um conjunto de dados por um único valor ou por poucos valores. No entanto, em algumas situações, é necessário utilizar medidas que indicam o quanto um conjunto de dados está próximo ou afastado em relação à média aritmética. Considere, por exemplo, a situação descrita a seguir.

Uma microempresa que fabrica calçados funciona de segunda-feira a sábado em dois turnos de trabalho com o mesmo tempo de jornada e com a mesma quantidade de funcionários cada um: turno **A** das 8 h às 14 h e turno **B** das 14 h às 20 h. Com o objetivo de entender melhor a produção e elaborar um plano que ajude a aperfeiçoar os processos, o proprietário organizou a quantidade de pares de calçados produzidos por dia em uma certa semana. Analise a tabela.

» Quantidade de pares de calçados produzidos em certa semana, por turno

Turno \ Dia	seg.	ter.	qua.	qui.	sex.	sáb.
A	16	14	15	14	15	16
B	10	18	12	21	10	19

Fonte: Proprietário da microempresa.



» Linha de produção em indústria de calçados no município de Franca (SP). Fotografia de 2016.

Para analisar essa produção, podemos, inicialmente, calcular a média de pares de calçados produzidos por dia em cada turno.

$$\begin{aligned} \bar{x}_A &= \frac{16 + 14 + 15 + 14 + 15 + 16}{6} = \frac{90}{6} = 15 \\ \bar{x}_B &= \frac{10 + 18 + 12 + 21 + 10 + 19}{6} = \frac{90}{6} = 15 \end{aligned}$$

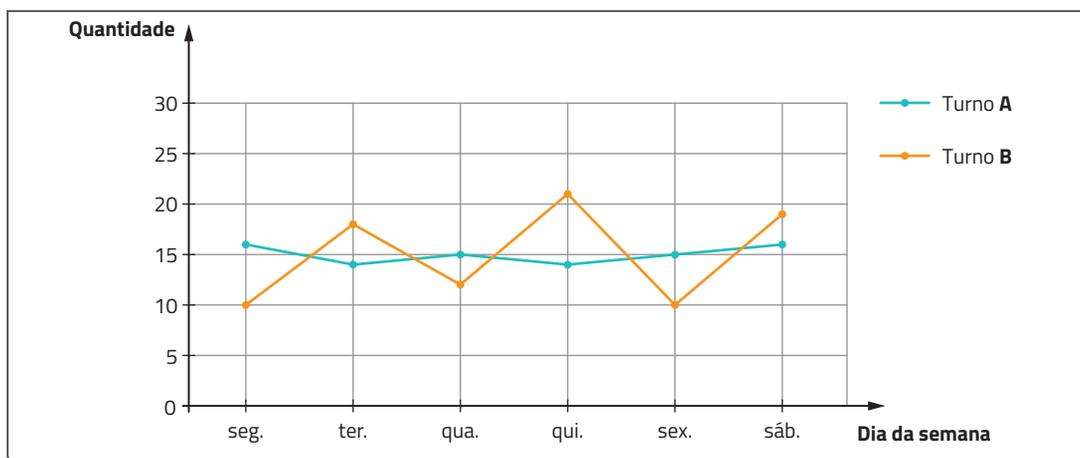
### Para pensar

O que você pode afirmar sobre as produções diárias de cada turno em relação à média obtida na semana? Explique.

Resposta esperada: As produções diárias no turno A correspondem, de maneira geral, a valores mais próximos da média do que as produções no turno B.

Agora, analise o gráfico.

### » Quantidade de pares de calçados produzidos em certa semana, por turno



Fonte: Proprietário da microempresa.

Note que, para os dois turnos de trabalho, a média diária de produção é igual a 15 pares de calçados. No entanto, ao analisarmos os dados da produção, podemos verificar que a produção no turno **A**, em cada dia, foi mais uniforme, ou seja, mais próxima à média, do que a produção no turno **B**. Esse fato sugere a necessidade de uma investigação mais minuciosa do processo produtivo.

Em situações como essa, é possível calcular as denominadas **medidas de dispersão**, como a amplitude, a variância e o desvio padrão, que estudaremos a seguir.

## Amplitude

Quando pesquisamos em um dicionário o significado da palavra “amplitude”, um sinônimo que podemos encontrar é “extensão”, ou seja, aquilo que dá ideia da dimensão ou tamanho de algo. Leia, a seguir, o conceito de amplitude em Estatística.

Dados  $n$  números reais  $x_1, x_2, x_3, \dots, x_{n-1}, x_n$ , com  $x_1 \leq x_2 \leq x_3 \leq \dots \leq x_{n-1} \leq x_n$ , a **amplitude** desse conjunto de dados, indicada por  $a$ , corresponde à diferença entre o número maior e o menor.

$$a = x_n - x_1$$

Em relação à situação apresentada na página anterior, podemos determinar a amplitude da produção diária de pares de calçados na semana, no turno **A** ( $a_A$ ) e no turno **B** ( $a_B$ ).

- $a_A = 16 - 14 = 2$
- $a_B = 21 - 10 = 11$

Assim, podemos dizer que, nessa semana, a amplitude na produção diária do turno **A** foi de dois pares de calçados, enquanto no turno **B** a amplitude foi de 11 pares de calçados.

### Para pensar

Escreva quatro números cuja amplitude entre eles seja igual a 5. Depois, escreva quatro números cuja amplitude entre eles seja igual a zero.

Uma possível resposta: 5, 7, 8, 10  
(10 - 5 = 5). 3, 3, 3, 3 (3 - 3 = 0).

## Variância

Agora, vamos definir a medida de dispersão denominada variância.

Dados  $n$  números reais  $x_1, x_2, x_3, \dots, x_{n-1}, x_n$  com média aritmética  $\bar{x}$ , a **variância** desse conjunto de dados, indicada por  $v$ , é dada por:

$$v = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

Acompanhe, a seguir, o cálculo da variância da produção diária de pares de calçados na semana, no turno **A** ( $v_A$ ) e no turno **B** ( $v_B$ ).

$$\begin{aligned} \blacksquare v_A &= \frac{(16 - 15)^2 + (14 - 15)^2 + (15 - 15)^2 + (14 - 15)^2 + (15 - 15)^2 + (16 - 15)^2}{6} \Rightarrow \\ \Rightarrow v_A &= \frac{1 + 1 + 0 + 1 + 0 + 1}{6} = \frac{4}{6} = \frac{2}{3} \end{aligned}$$

$$\begin{aligned} \blacksquare v_B &= \frac{(10 - 15)^2 + (18 - 15)^2 + (12 - 15)^2 + (21 - 15)^2 + (10 - 15)^2 + (19 - 15)^2}{6} \Rightarrow \\ \Rightarrow v_B &= \frac{25 + 9 + 9 + 36 + 25 + 16}{6} = \frac{120}{6} = 20 \end{aligned}$$

Como  $v_A < v_B$ , podemos dizer que a produção diária de pares de calçados na semana foi mais regular (ou homogênea) no turno **A** do que no turno **B**.

## Desvio padrão

Vimos que o cálculo da variância envolve o quadrado da diferença de cada valor do conjunto de dados e a média, o que pode dificultar a interpretação do resultado. Agora, vamos estudar o desvio padrão, uma medida de dispersão que corresponde à raiz quadrada da variância.

Dados  $n$  números reais cuja variância é  $v$ , temos que o **desvio padrão** desse conjunto de dados, indicado por  $d$ , é dado por:

$$d = \sqrt{v}$$

A variância, assim como o desvio padrão, procura medir a dispersão dos valores do conjunto de dados em torno de sua média.

Em relação à situação apresentada nas páginas anteriores, podemos calcular o desvio padrão da produção diária de pares de calçados na semana, no turno **A** ( $d_A$ ) e no turno **B** ( $d_B$ ).

$$\blacksquare d_A = \sqrt{\frac{2}{3}} \approx 0,816 \quad \blacksquare d_B = \sqrt{20} \approx 4,472$$

Portanto, podemos dizer que o desvio padrão na produção diária de pares de calçados na semana foi de aproximadamente 0,816 par de calçados no turno **A** e 4,472 pares de calçados no turno **B**.

### Para pensar

Qual medida de dispersão não considera, necessariamente, cada um dos valores do conjunto de dados da pesquisa em seu cálculo: amplitude, variância ou desvio padrão? Justifique.

Resposta esperada: Amplitude, pois essa medida de dispersão é calculada a partir de apenas dois valores do conjunto de dados: o maior e o menor valor.

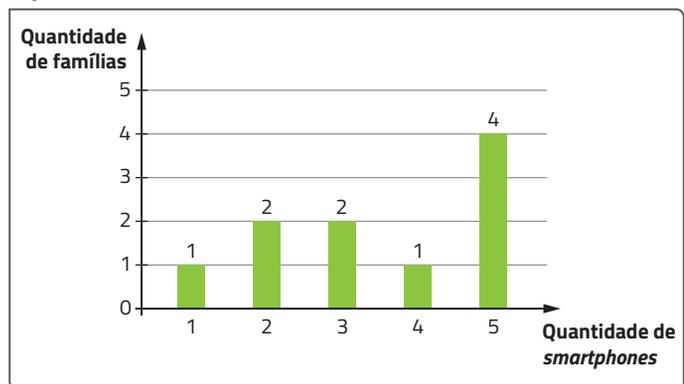
## Atividades resolvidas

**R4.** Uma loja especializada na venda de *smartphones* realizou uma pesquisa estatística com 10 famílias de certo bairro, escolhidas aleatoriamente, a fim de identificar hábitos de uso desses dispositivos. O gráfico ao lado consta no relatório em que foram registrados os resultados dessa pesquisa.

Em relação à quantidade de *smartphones* em uso, por família, calcule:

- a média aritmética;
- a amplitude;
- o desvio padrão.

» Resultado da pesquisa sobre *smartphones* em uso, por família, em 6/3/2020



Fonte: Loja de *smartphones*.

### Resolução

$$\text{a) } \bar{x} = \frac{1 \cdot 1 + 2 \cdot 2 + 3 \cdot 2 + 4 \cdot 1 + 5 \cdot 4}{1 + 2 + 2 + 1 + 4} = \frac{35}{10} = 3,5$$

Portanto, em média, havia 3,5 *smartphones* em uso, por família.

**b)** As quantidades máxima e mínima de *smartphones* em uso por essas famílias são, respectivamente, 5 e 1. Assim, temos:

$$a = 5 - 1 = 4$$

Portanto, a amplitude desse conjunto de dados corresponde a 4 *smartphones*.

$$\text{c) } v = \frac{1 \cdot (1 - 3,5)^2 + 2 \cdot (2 - 3,5)^2 + 2 \cdot (3 - 3,5)^2 + 1 \cdot (4 - 3,5)^2 + 4 \cdot (5 - 3,5)^2}{1 + 2 + 2 + 1 + 4} \Rightarrow$$

$$\Rightarrow v = \frac{1 \cdot 6,25 + 2 \cdot 2,25 + 2 \cdot 0,25 + 1 \cdot 0,25 + 4 \cdot 2,25}{10} = \frac{20,5}{10} = 2,05$$

Assim, segue que:

$$d = \sqrt{v} = \sqrt{2,05} \approx 1,43$$

Portanto, o desvio padrão desse conjunto de dados corresponde a aproximadamente 1,43 *smartphones*.

**R5.** (Enem/MEC) Um fiscal de certa empresa de ônibus registra o tempo, em minuto, que um motorista novato gasta para completar certo percurso. No Quadro 1 figuram os tempos gastos pelo motorista ao realizar o mesmo percurso sete vezes. O Quadro 2 apresenta uma classificação para a variabilidade do tempo, segundo o valor do desvio padrão. Com base nas informações apresentadas nos quadros, a variabilidade do tempo é

Quadro 1							
Tempos (em minuto)	48	54	50	46	44	52	49

Quadro 2	
Variabilidade	Desvio padrão do tempo (min)
Extremamente baixa	$0 < \sigma \leq 2$
Baixa	$2 < \sigma \leq 4$
Moderada	$4 < \sigma \leq 6$
Alta	$6 < \sigma \leq 8$
Extremamente alta	$\sigma > 8$

- a) extremamente baixa.                      c) moderada.                                      e) extremamente alta.  
b) baixa.    d) alta.

**Resolução**

Para resolver essa atividade, podemos realizar as seguintes etapas.

**1ª) Compreender o enunciado.**

Do enunciado, temos que:

- o tempo gasto pelo motorista para realizar um mesmo percurso foi registrado sete vezes;
- a variabilidade do tempo é interpretada de acordo com o desvio padrão;
- determinar como os tempos gastos no percurso por esse motorista podem ser interpretados.

**2ª) Elaborar um plano.**

Inicialmente, pode-se calcular a média e a variância dos tempos apresentados. Depois, com base no valor da variância, determinar o desvio padrão desses tempos e, com esse resultado, a classificação para a variabilidade do tempo.

**3ª) Executar o plano.**

Calculamos a média dos tempos:

$$\bar{x} = \frac{48 + 54 + 50 + 46 + 44 + 52 + 49}{7} = \frac{343}{7} = 49$$

Agora, calculamos a variância dos tempos:

$$v = \frac{(48 - 49)^2 + (54 - 49)^2 + (50 - 49)^2 + (46 - 49)^2 + (44 - 49)^2 + (52 - 49)^2 + (49 - 49)^2}{7} \Rightarrow$$

$$\Rightarrow v = \frac{70}{7} = 10$$

Por fim, determinamos o desvio padrão dos tempos:  $d = \sqrt{10}$ .

Como  $2 < \sqrt{10} \leq 4$ , temos que a variabilidade do tempo é classificada como baixa.

**4ª) Verificar os resultados.**

Para verificar os resultados, os cálculos podem ser refeitos com o auxílio de uma calculadora ou, ainda, por meio de uma planilha eletrônica ou outro programa de computador para determinar o desvio padrão do conjunto de dados apresentados.

Portanto, a alternativa **b** é a correta.

**Dica**

Na seção **Você conectado**, ao final desta Unidade, há informações sobre o cálculo de medidas de posição e de dispersão usando a planilha eletrônica **LibreOffice Calc**.