UNIP UNIVERSIDADE PAULISTA

Interativa

# Estatística e Probabilidade

**Autora:** Profa. Christiane Mazur Doi
**Colaboradora:** Profa. Vanessa Santos Lessa

Professora conteudista: Christiane Mazur Doi

Graduada em Engenharia Química, licenciada em Matemática, mestra em Ciências (Tecnologia Nuclear) e doutora em Engenharia Metalúrgica e de Materiais, com aperfeiçoamento em Estatística. Coordenadora da Comissão de Qualificação e Avaliação (CQA) da UNIP e professora titular da UNIP.

**Dados Internacionais de Catalogação na Publicação (CIP)**

D657e Doi, Christiane Mazur.

Estatística e Probabilidade / Christiane Mazur Doi. – São Paulo: Editora Sol, 2022.

184 p., il.

Nota: este volume está publicado nos Cadernos de Estudos e Pesquisas da UNIP, Série Didática, ISSN 1517-9230.

1. Estatística descritiva. 2. Probabilidade. 3. Estatística indutiva. I. Título.

CDU 519.2(076.1)

U514.20 – 22

© Todos os direitos reservados. Nenhuma parte desta obra pode ser reproduzida ou transmitida por qualquer forma e/ou quaisquer meios (eletrônico, incluindo fotocópia e gravação) ou arquivada em qualquer sistema ou banco de dados sem permissão escrita da Universidade Paulista.

Prof. Dr. João Carlos Di Genio
**Reitor**

Profa. Sandra Miessa
**Reitora em Exercício**

Profa. Dra. Marilia Ancona Lopez
**Vice-Reitora de Graduação**

Profa. Dra. Marina Ancona Lopez Soligo
**Vice-Reitora de Pós-Graduação e Pesquisa**

Profa. Dra. Claudia Meucci Andreatini
**Vice-Reitora de Administração**

Prof. Dr. Paschoal Laercio Armonia
**Vice-Reitor de Extensão**

Prof. Fábio Romeu de Carvalho
**Vice-Reitor de Planejamento e Finanças**

Profa. Melânia Dalla Torre
**Vice-Reitora de Unidades do Interior**

**Unip Interativa**

Profa. Elisabete Brihy
Prof. Marcelo Vannini
Prof. Dr. Luiz Felipe Scabar
Prof. Ivan Daliberto Frugoli

**Material Didático**

Comissão editorial:
Profa. Dra. Christiane Mazur Doi
Profa. Dra. Angélica L. Carlini
Profa. Dra. Ronilda Ribeiro

Apoio:
Profa. Cláudia Regina Baptista
Profa. Deise Alcantara Carreiro

Projeto gráfico:
Prof. Alexandre Ponzetto

Revisão:
Ricardo Duarte
Ana Fazzio

# Sumário

## Estatística e Probabilidade

**Unidade II**

## APRESENTAÇÃO

Caro aluno,

Você, como futuro cientista da computação, terá à sua disposição máquinas e dispositivos capazes de operar com quantidades elevadíssimas de dados, que irão requerer organização adequada e análise correta para que decisões consistentes sejam tomadas. O que dissemos faz parte da estatística.

E não é só isso. Por exemplo, o esboço, os testes e a implementação de algoritmos e ferramentas computacionais muitas vezes demandam a realização de cálculos probabilísticos.

Resultados da aplicação de métodos estatísticos podem ser mais bem compreendidos se forem expressos por meio de gráficos, tabelas ou animações. Aqui, temos mais aplicações que permitem vermos como a estatística está intimamente ligada à ciência da computação.

No campo da ciência, por exemplo, os testes preliminares de validação de uma nova vacina a ser usada mundialmente são executados pela observação da sua efetividade em determinados grupos. A definição de tais grupos e a análise dos dados experimentais são feitas com o apoio de softwares estatísticos.

Neste livro-texto, você terá a oportunidade de compreender conceitos fundamentais da estatística pelo estudo de situações do dia a dia e de situações voltadas ao contexto específico da ciência da computação. Além disso, conhecerá as ferramentas matemáticas indicadas para a realização dos cálculos apropriados a determinados modelos e segmentos da estatística.

Vale acrescentar que este material é escrito em linguagem simples e direta, como se houvesse uma conversa entre a autora e o leitor. Adicionalmente, são inseridas muitas figuras, que auxiliam no entendimento dos tópicos desenvolvidos. Os itens chamados de **observação** e de **lembrete** são oportunidades para que você tire eventuais dúvidas. Os itens chamados de **saiba mais** possibilitam que você amplie seus conhecimentos. Há, ainda, muitos exemplos de aplicação, resolvidos em detalhes, o que implica a fixação dos assuntos abordados.

Boa leitura!

## INTRODUÇÃO

Uma das maneiras de classificarmos a estatística é "dividi-la" em:

- estatística descritiva;
- probabilidades;
- estatística indutiva.

Podemos pensar que a estatística descritiva se destina a organizar, a descrever, a explorar, a expressar e a sintetizar as informações brutas vindas da aplicação de um questionário, da observação de algum evento ou da contagem de ocorrências, por exemplo. Em suma, na estatística descritiva, trabalhamos com conjuntos de dados oriundos de algo "certo", que já aconteceu, que "pertence ao passado". Por exemplo, se soubermos a idade de 3 pessoas e quisermos calcular a idade média desse grupo de 3 pessoas, não há incerteza associada a tal cálculo.

Na parte das probabilidades, lidamos com a chance de algo acontecer, ou seja, estamos no campo da incerteza, dos eventos aleatórios, dos acontecimentos que não podem ser previstos com 100% de exatidão. Por exemplo, sabemos ter 50% de probabilidade de obtermos cara ao lançar uma moeda. Logo, não há como dizermos, com certeza, se obteremos cara ou coroa nesse lançamento.

Na estatística indutiva, trabalhamos com uma amostra, a fim de que, com o uso de técnicas e métodos adequados, obtenhamos informações a respeito da população que tal amostra representa. Nesse caso, para dado intervalo de confiança, temos um erro associado. Por exemplo, se uma pesquisa eleitoral diz que certo candidato tem 60% dos votos, com margem de erro de 3% e confiança de 95%, significa que esse candidato tem 95% de chance de ter entre 57% e 63% dos votos na data da pesquisa. Cabe destacar que nem sempre podemos utilizar a relação "amostra boa é amostra grande". O que vale é "amostra boa é amostra que traz consigo todas as características presentes na população e na proporção em que ocorrem na população".

Este livro-texto é dividido didaticamente em duas unidades (unidade I e unidade II).

Na unidade I, apresentamos:

- noções gerais sobre conjunto de dados (estatística descritiva);
- organização de um conjunto de dados brutos;
- frequência absoluta e frequência relativa das medidas de um conjunto de dados;
- medidas de tendência central (média, moda e mediana) e medidas de posição (amplitude, variância e desvio padrão) de um conjunto de dados;
- noções gerais sobre probabilidades;
- probabilidades da união e da intersecção de eventos;
- probabilidade condicional;
- eventos independentes;
- distribuição discreta de probabilidade (função discreta de probabilidade);

- principais modelos de distribuição discreta de probabilidade (modelo uniforme discreto, modelo de Bernoulli e modelo binomial);
- distribuição contínua de probabilidade (função densidade de probabilidade);
- principais modelos de distribuição contínua de probabilidade (modelo uniforme contínuo e modelo normal).

Na unidade II, apresentamos:

- noções gerais sobre estatística indutiva, incluindo as definições de parâmetro, estimador e estimativa;
- correlação e regressão linear;
- intervalos de confiança para média, variância e desvio padrão;
- introdução ao teste de hipótese;
- testes qui-quadrado (teste de aderência e teste de independência).

Bom estudo!

# Unidade I

## 1 ESTATÍSTICA DESCRITIVA (CONJUNTO DE DADOS): PARTE 1

### 1.1 Noções gerais sobre conjunto de dados

Suponha que, em determinada localidade, tenha ocorrido uma Olimpíada de Informática na qual os concorrentes deveriam propor a solução de um problema por meio de um algoritmo. O vencedor foi o participante que sugeriu o algoritmo que apresentou o menor número possível de falhas.

Na tabela 1, temos os nomes dos 17 finalistas dessa Olimpíada de Informática fictícia e as quantidades de falhas presentes nos algoritmos por eles propostos. Nessa tabela, o número de chamada não é a posição do finalista em um ranking de classificação.

**Tabela 1 – Nomes dos 17 finalistas de uma Olimpíada de Informática e quantidades de falhas nos algoritmos por eles propostos**

| N. de chamada | Nome do finalista | Quantidade de falhas no algoritmo |
|---|---|---|
| 1 | André | 3 |
| 2 | Bianca | 2 |
| 3 | Beatriz | 3 |
| 4 | Catarina | 2 |
| 5 | Diego | 1 |
| 6 | Elsa | 5 |
| 7 | Fábio | 1 |
| 8 | Gabriela | 2 |
| 9 | Júlia | 3 |
| 10 | Laila | 2 |
| 11 | Marcelo | 0 |
| 12 | Mariana | 1 |
| 13 | Patrícia | 2 |
| 14 | Paulo | 2 |
| 15 | Rafael | 3 |
| 16 | Sofia | 2 |
| 17 | Tobias | 2 |

A tabela 1 é um exemplo do que chamamos de **tabela de dados brutos**, em que tabulamos o conjunto das informações obtidas a respeito de certa característica – no caso, a quantidade de falhas presentes em um algoritmo.

Veja que estamos lidando com algo "certo", que já ocorreu, que "pertence ao passado", ou seja, estamos lidando com um conjunto de dados vindos, por exemplo, da observação de algum evento já acontecido, do relato de um questionário já aplicado ou de uma contagem de casos já vistos.

Os resultados observados em uma pesquisa, um questionário ou uma brincadeira formam o conjunto de dados que, usualmente, mostra os possíveis valores da característica estudada e suas respectivas contagens.

Podemos organizar as informações contidas em uma tabela de dados brutos de modo a obtermos tabelas de frequência absoluta e de frequência relativa, que serão estudadas nos próximos tópicos, ou seja, de modo a obtermos tabelas de dados organizados.

Com base nas tabelas de dados organizados, podemos, por exemplo, expressar as informações por meio de diferentes tipos de gráfico e fazer cálculos das medidas de tendência central (média, moda e mediana) e das medidas de dispersão (variância e desvio padrão), como veremos adiante.

## 1.2 Organização de um conjunto de dados

Vamos entender como se faz a organização de um conjunto de dados voltando ao caso da Olimpíada de Informática, cujas informações foram apresentadas na tabela 1.

Começamos respondendo às perguntas a seguir.

- Há finalistas que propuseram algoritmos sem falhas?

  Sim, apenas um finalista, o Marcelo.

- Há finalistas que propuseram algoritmos com 1 falha?

  Sim, 3 finalistas, o Diego, o Fábio e a Mariana.

- Há finalistas que propuseram algoritmos com 2 falhas?

  Sim, 8 finalistas, a Bianca, a Catarina, a Gabriela, a Laila, a Patrícia, o Paulo, a Sofia e o Tobias.

- Há finalistas que propuseram algoritmos com 3 falhas?

  Sim, 4 finalistas, o André, a Beatriz, a Júlia e o Rafael.

- Há finalistas que propuseram algoritmos com 4 falhas?

  Não, nenhum (0 finalistas).

- Há finalistas que propuseram algoritmos com 5 falhas?

  Sim, 1 finalista, a Elsa.

Com essas respostas, podemos elaborar a tabela a seguir, que mostra as quantidades de finalistas que fizeram algoritmos com 0, 1, 2, 3, 4 ou 5 falhas. Além disso, adicionamos os nomes dos finalistas.

**Tabela 2 – Quantidades de falhas no algoritmo e quantidades de finalistas**

| Quantidade de falhas | Quantidade de finalistas | Nome do finalista |
|---|---|---|
| 0 | 1 | Marcelo |
| 1 | 3 | Diego, Fábio e Mariana |
| 2 | 8 | Bianca, Catarina, Gabriela, Laila, Patrícia, Paulo, Sofia e Tobias |
| 3 | 4 | André, Beatriz, Júlia e Rafael |
| 4 | 0 | - |
| 5 | 1 | Elsa |
| Total | 1 + 3 + 8 + 4 + 0 + 1 = 17 | |

Pela tabela, vemos, por exemplo, que dos 17 nomes dos 17 finalistas da Olimpíada de Informática, 3 fizeram algoritmos com 1 falha e nenhum fez algoritmos com 4 falhas. Na tabela, temos o que chamamos de **medidas** ou **observações**.

### 1.2.1 Frequência absoluta de cada medida de um conjunto de dados

Vamos chamar de **frequência absoluta** de cada medida, indicada por FA, a quantidade de finalistas da Olimpíada de Informática que fez determinado número de falhas no algoritmo, indicado por x.

Assim, na situação em estudo, temos o seguinte.

- A FA de 0 falhas é igual a 1 (se x = 0, FA = 1).
- A FA de 1 falha é igual a 3 (se x = 1, FA = 3).
- A FA de 2 falhas é igual a 8 (se x = 2, FA = 8).
- A FA de 3 falhas é igual a 4 (se x = 3, FA = 4).
- A FA de 4 falhas é igual a 0 (se x = 4, FA = 0).
- A FA de 5 falhas é igual a 1 (se x = 5, FA = 1).

Na tabela a seguir, apresentamos as frequências absolutas antes apontadas. Chamamos a quantidade total de medidas (ou de observações) de N. No caso em estudo, temos N = 17.

**Tabela 3 – Frequências absolutas das quantidades de falhas nos algoritmos**

| Quantidade de falhas (x) | Frequência absoluta (FA) |
|---|---|
| 0 | 1 |
| 1 | 3 |
| 2 | 8 |
| 3 | 4 |
| 4 | 0 |
| 5 | 1 |
| Total | N = 1 + 3 + 8 + 4 + 0 + 1 = 17 |

As frequências absolutas da tabela 3 podem ser representadas, por exemplo, em um gráfico de barras, conforme ilustrado na figura 1.

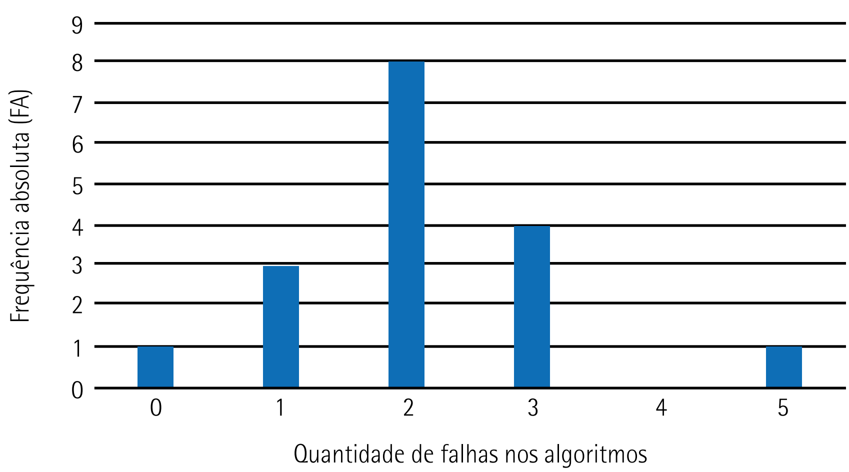



Figura 1 – Gráfico de barras: frequências absolutas das quantidades de falhas nos algoritmos feitos pelos finalistas de uma Olimpíada de Informática

### 1.2.2 Frequência relativa de cada medida de um conjunto de dados

Para o exemplo em estudo, podemos calcular a frequência relativa, indicada por FR, de cada quantidade de falhas presentes nos algoritmos feitos pelos finalistas da Olimpíada de Informática. Para isso, dividimos a frequência absoluta (FA) pelo número total N de finalistas, que é 17. Ou seja:

$$FR = \frac{FA}{N}$$

Na tabela a seguir, apresentamos as frequências absolutas de cada uma das medidas do exemplo em estudo e, também, calculamos suas frequências relativas.

**Tabela 4 – Frequências absolutas e frequências relativas das quantidades de falhas nos algoritmos**

| Quantidade de falhas (x) | Frequência absoluta (FA) | Frequência relativa (FR), sendo $FR = \frac{FA}{N}$ |
|---|---|---|
| 0 | 1 | 1/17 = 0,05882 |
| 1 | 3 | 3/17 = 0,17647 |
| 2 | 8 | 8/17 = 0,47059 |
| 3 | 4 | 4/17 = 0,23529 |
| 4 | 0 | 0/17 = 0 |
| 5 | 1 | 1/17 = 0,05882 |
| Total | N = 1 + 3 + 8 + 4 + 0 + 1 = 17 | $\text{Soma} = \frac{1}{17} + \frac{3}{17} + \frac{8}{17} + \frac{4}{17} + \frac{0}{17} + \frac{1}{17}$ <br> Soma = 1 |

Vale notar que, em qualquer conjunto de dados, a soma de todas as frequências relativas dá 1.

As frequências relativas da tabela podem ser representadas, por exemplo, em um gráfico de setores circulares, conforme ilustrado na figura 2.

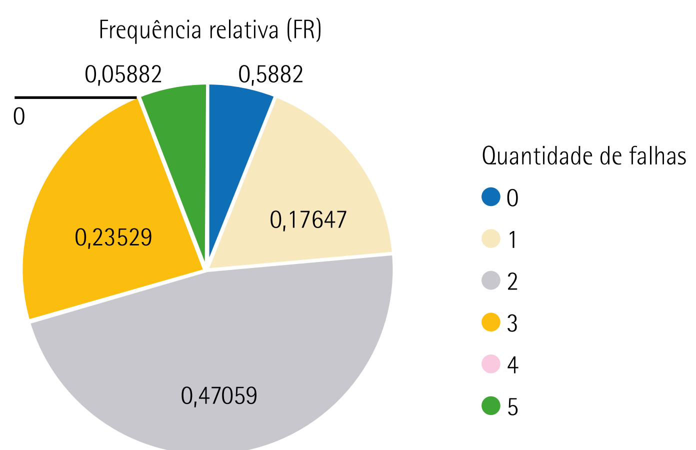



Figura 2 – Gráfico de setores circulares: frequências relativas das quantidades de falhas nos algoritmos feitos pelos finalistas de uma Olimpíada de Informática

### 1.2.3 Percentual de cada medida de um conjunto de dados

Podemos fazer um cálculo bastante semelhante ao feito para determinarmos a frequência relativa, multiplicando-a por 100%. Desse modo, obtemos os percentuais de cada quantidade de falhas relativas, indicados por Perc%. Ou seja:

Perc% = FR . 100%

Na tabela a seguir, temos as frequências absolutas, as frequências relativas e os percentuais da situação em estudo, relativa a uma Olimpíada de Informática fictícia.

**Tabela 5 – Frequências absolutas, frequências relativas e percentuais das quantidades de falhas nos algoritmos**

| Quantidade de falhas (x) | Frequência absoluta (FA) | Frequência relativa (FR), sendo $FR = \frac{FA}{N}$ | Percentual (Perc%), sendo Perc% = FR . 100 |
|---|---|---|---|
| 0 | 1 | 1/17 = 0,05882 | 5,882% |
| 1 | 3 | 3/17 = 0,17647 | 17,647% |
| 2 | 8 | 8/17 = 0,47059 | 47,059% |
| 3 | 4 | 4/17 = 0,23529 | 23,529% |
| 4 | 0 | 0/17 = 0 | 0% |
| 5 | 1 | 1/17 = 0,05882 | 5,882% |
| Total | N = 1 + 3 + 8 + 4 + 0 + 1 = 17 | $Soma = \frac{1}{17} + \frac{3}{17} + \frac{8}{17} + \frac{4}{17} + \frac{0}{17} + \frac{1}{17}$ Soma = 1 | Soma = 100% |

Vale notar que, em qualquer conjunto de dados, a soma de todos os percentuais de cada uma das medidas dá 100%.

Os percentuais da tabela podem ser representados, por exemplo, em um gráfico de setores circulares, conforme ilustrado na figura 3.

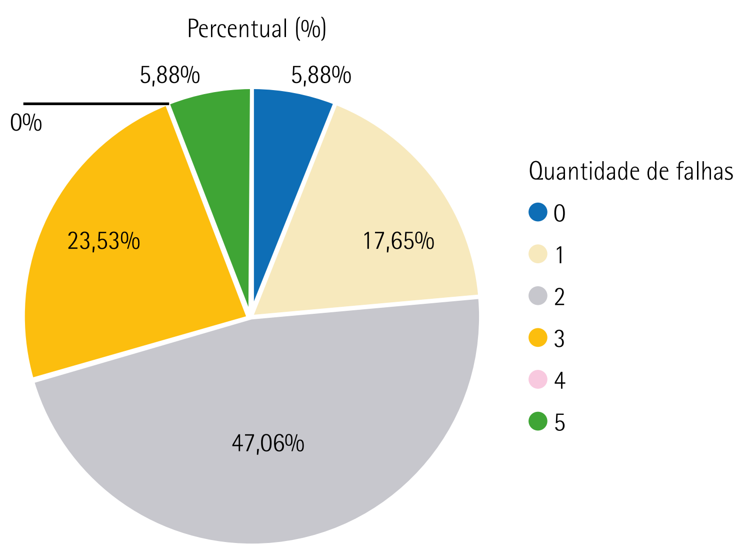



Figura 3 – Gráfico de setores circulares: percentuais das quantidades de falhas nos algoritmos feitos pelos finalistas de uma Olimpíada de Informática

Para fixar o que vimos, vamos imaginar que 132 estudantes do 2º ano do Ensino Médio das 5 escolas de um bairro tenham respondido à seguinte questão: você vê relações entre a matemática e situações do seu dia a dia?

Suponha que as respostas dadas foram as seguintes.

- 88 estudantes disseram que sim, que viam relações entre a matemática e situações do seu dia a dia em várias circunstâncias.
- 42 estudantes disseram que viam relações entre a matemática e situações do seu dia a dia apenas em algumas circunstâncias.
- 2 estudantes disseram que não, que não viam relações entre a matemática e situações do seu dia a dia.

Esses resultados estão sumarizados e organizados na tabela a seguir.

**Tabela 6 – Você vê relações entre a matemática e situações do seu dia a dia?**

| Você vê relações entre a matemática e situações do seu dia a dia? | Frequência absoluta | Frequência relativa | Percentual |
|---|---|---|---|
| Sim, em várias circunstâncias | 88 | 88/132 = 0,667 | 66,7% |
| Apenas em algumas circunstâncias | 42 | 42/132 = 0,318 | 31,8% |
| Não, em nenhuma circunstância | 2 | 2/132 = 0,015 | 1,5% |
| Soma | 132 | 1 | 100% |

Nas figuras 4 e 5, temos expressões gráficas dos dados apresentados na tabela.

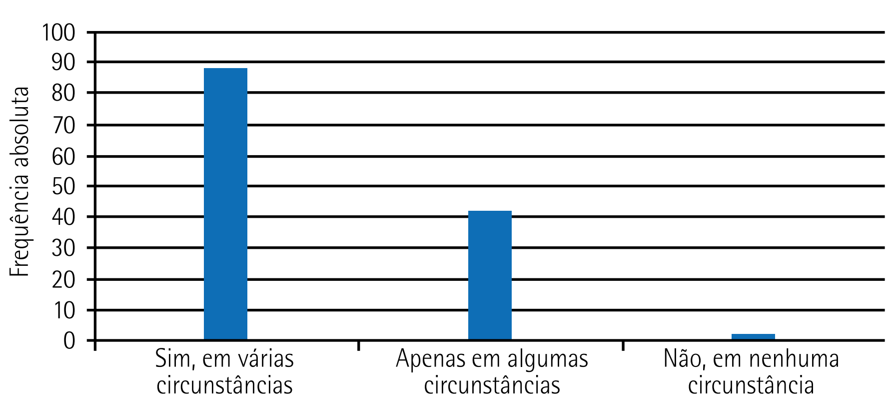



Figura 4 – Frequências absolutas: percepção de relações entre a matemática e situações do dia a dia

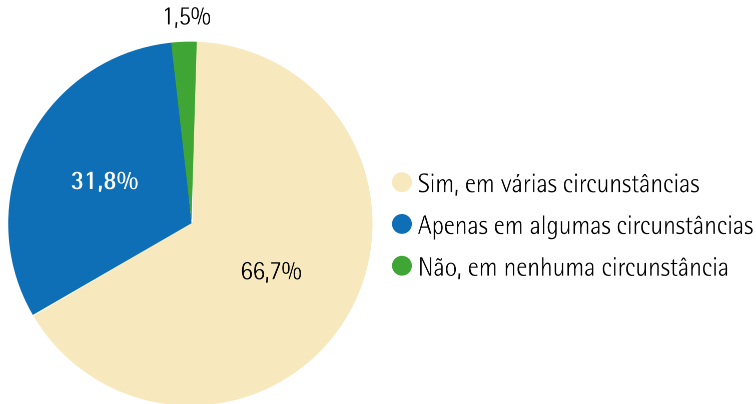



Figura 5 – Percentuais: percepção de relações entre a matemática e situações do dia a dia

## 2 ESTATÍSTICA DESCRITIVA (CONJUNTO DE DADOS): PARTE 2

### 2.1 Medidas de tendência central de um conjunto de dados: média, moda e mediana

Suponha que você queira mostrar, em um único número, a representação de todas as idades dos 40 alunos de uma turma que cursa Ciência da Computação. Para esse tipo de caso, frequentemente pensamos em algo como "Vamos calcular a idade média da turma" ou "Vamos calcular a média das idades da turma".

**Observação**

Veja que estamos pensando em algum número que represente o conjunto de dados composto pelas idades de todos os alunos de determinada turma.

É comum querermos que um único número, um único valor, "sintetize" certa característica de todo o grupo. Na maioria das vezes, esse número é dado pela média das observações que fizemos. Mas, como veremos na sequência, a média não é a única opção que temos para resolver esse problema.

Nesse âmbito, temos as medidas de tendência central, como a média, a moda e a mediana, que serão estudadas a seguir no contexto de situações que envolvem um conjunto de dados.

**Observação**

As medidas de tendência central – média, moda e mediana – também são chamadas de medidas de posição.

### 2.1.1 Média de um conjunto de dados

A média de um conjunto de dados é feita da seguinte maneira.

- Somamos todos os valores observados.
- Dividimos o resultado dessa soma pela quantidade de elementos somados.

Observação

A média à qual estamos nos referindo é a média aritmética. Existem outros tipos de média, como a média ponderada, em que damos "pesos" diferentes para determinadas medidas.

Vamos ver um caso bem simples de cálculo de média. Observe, na tabela a seguir, as notas bimestrais de Lucas em matemática.

**Tabela 7 – Notas bimestrais de Lucas em matemática**

| Bimestre | Nota |
|---|---|
| 1º bimestre | 7,5 |
| 2º bimestre | 6 |
| 3º bimestre | 6 |
| 4º bimestre | 9 |

Para calcularmos a média anual de Lucas em matemática, somamos suas 4 notas bimestrais e dividimos esse total por 4:

$$\text{Média de Lucas em matemática} = \frac{7{,}5 + 6 + 6 + 9}{4} = \frac{28{,}5}{4} = 7{,}125$$

Veja que, em nenhum bimestre, Lucas obteve nota 7,125 em matemática, embora essa seja sua média anual na disciplina. Podemos pensar assim: se Lucas tivesse obtido "sempre" a mesma nota bimestral em matemática, essa nota seria 7,125.

Observação

Há uma brincadeira que ilustra bem o sentido da média: se você come uma barra de chocolate e eu não como chocolate algum, "na média", cada um de nós comeu meia barra de chocolate!

Imagine, agora, que estejamos registrando os pesos dos 12 bebês que nasceram em janeiro deste ano em determinada maternidade, conforme mostrado na tabela 8. Nesse caso, dizemos que a **característica observada** é a variável peso.

**Tabela 8 – Pesos dos recém-nascidos em janeiro em determinada maternidade**

| Nome do recém-nascido | Peso (kg) |
|---|---|
| Alice | 3,05 |
| Hugo | 3,12 |
| Laura | 2,97 |
| Silas | 3,21 |
| Tiago | 3,18 |
| Rodrigo | 3,09 |
| Carina | 2,98 |
| Gisele | 3,13 |
| Viviane | 3,04 |
| Ricardo | 3,13 |
| Ulisses | 3,26 |
| Marta | 3,22 |

Podemos usar um símbolo para representar, de modo "resumido", a característica estudada (no caso, o peso do recém-nascido). Em geral, utilizamos a letra x para isso. Assim, no caso da tabela 8, como o peso do primeiro recém-nascido é 3,05 kg, temos $x_1 = 3,05$, e assim por diante. Se usarmos essa notação, o peso do 12º recém-nascido será indicado por $x_{12}$.

Cada uma das **observações** – no caso, $x_1, x_2, x_3, ..., x_{12}$ – é chamada de **valor observado** ou **medida**. O número total de observações ou medidas, como já dissemos, geralmente é indicado por N (no caso, N = 12).

Assim, ficamos com o que se apresenta na tabela 9.

**Tabela 9 – Pesos dos recém-nascidos em janeiro em determinada maternidade**

| Medida | Peso (kg) |
|---|---|
| $x_1$ | 3,05 |
| $x_2$ | 3,12 |
| $x_3$ | 2,97 |
| $x_4$ | 3,21 |
| $x_5$ | 3,18 |
| $x_6$ | 3,09 |
| $x_7$ | 2,98 |

| Medida | Peso (kg) |
|---|---|
| $x_8$ | 3,13 |
| $x_9$ | 3,04 |
| $x_{10}$ | 3,13 |
| $x_{11}$ | 3,26 |
| $x_{12}$ | 3,22 |

A média do conjunto de dados apresentados na tabela anterior, indicada por $\overline{x}_{CD}$, é calculada pela expressão a seguir.

$$\overline{x}_{CD} = \frac{x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9 + x_{10} + x_{11} + x_{12}}{12}$$

$$\overline{x}_{CD} = \frac{3{,}05 + 3{,}12 + 2{,}97 + 3{,}21 + 3{,}18 + 3{,}09 + 2{,}98 + 3{,}13 + 3{,}04 + 3{,}13 + 3{,}26 + 3{,}22}{12}$$

$$\overline{x}_{CD} = \frac{37{,}38}{12} = 3{,}115$$

Concluímos que o peso médio dos bebês que nasceram em janeiro deste ano na maternidade em estudo é igual a 3,115 kg.

De modo geral, a média $\overline{x}_{CD}$ das N medidas de um conjunto de dados, em que $x_i$ representa uma medida qualquer (lida como "i-ésima medida"), é dada por:

$$\overline{x}_{CD} = \frac{x_1 + x_2 + ... + x_i + ... + x_N}{N}$$

A igualdade anterior também pode ser escrita como:

$$\overline{x}_{CD} = \frac{\sum_{i=1}^{N} x_i}{N}$$

Na expressão, $\sum_{i=1}^{N} x_i$ é lido como "somatória de xis com índice i, começando em 1 e indo até N".

Muitas vezes, queremos obter a média de um conjunto de dados que inclui valores negativos, como apresentado na tabela 10.

**Tabela 10 – Conjunto de dados que inclui valores negativos**

| Medida | Valor |
|---|---|
| $x_1$ | 21 |
| $x_2$ | –7,3 |
| $x_3$ | 12,5 |
| $x_4$ | 32 |
| $x_5$ | –47 |
| $x_6$ | 23 |
| $x_7$ | 3 |
| $x_8$ | –51,7 |
| $x_9$ | –0,8 |

A média do conjunto de 9 dados apresentados na tabela anterior, indicada por $\overline{x}_{CD}$, é –1,7, conforme calculado a seguir.

$$\overline{x}_{CD} = \frac{x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 + x_9}{9}$$

$$\overline{x}_{CD} = \frac{21 + (-7{,}3) + 12{,}5 + 32 + (-47) + 23 + 3 + (-51{,}7) + (-0{,}8)}{9} = \frac{-15{,}3}{9} = -1{,}7$$

Se tivermos uma tabela de frequências absolutas, como a tabela 3, poderemos calcular a média do conjunto de dados da seguinte maneira.

- Multiplicamos cada medida pela sua frequência absoluta de ocorrência.
- Somamos todas as multiplicações de medidas pelas suas respectivas frequências absolutas.
- Dividimos a soma anterior pelo número N de medidas.

Vamos aplicar esse procedimento aos dados da tabela 11, que foram reproduzidos da tabela 3.

**Tabela 11 – Frequências absolutas das quantidades de falhas nos algoritmos**

| Quantidade de falhas no algoritmo | Frequência absoluta |
|---|---|
| $x_1 = 0$ | $FA_1 = 1$ |
| $x_2 = 1$ | $FA_2 = 3$ |
| $x_3 = 2$ | $FA_3 = 8$ |
| $x_4 = 3$ | $FA_4 = 4$ |

| Quantidade de falhas no algoritmo | Frequência absoluta |
|---|---|
| $x_5 = 4$ | $FA_5 = 0$ |
| $x_6 = 5$ | $FA_6 = 1$ |
| | $N = 1 + 3 + 8 + 4 + 0 + 1 = 17$ |

$$\overline{x}_{CD} = \frac{x_1 . FA_1 + x_2 . FA_2 + x_3 . FA_3 + x_4 . FA_4 + x_5 . FA_5}{N}$$

$$\overline{x}_{CD} = \frac{0 . 1 + 1 . 3 + 2 . 8 + 3 . 4 + 4 . 0 + 5 . 1}{17} = \frac{0 + 3 + 16 + 12 + 0 + 5}{17} = \frac{36}{17} = 2{,}12$$

Veja que a média de 2,12 falhas por algoritmo é o que costumamos chamar de **valor teórico**, pois, para ter sentido concreto, tal quantidade precisa ser um número inteiro.

Uma aplicação prática em que fazemos o cálculo de valor médio é na determinação do chamado **tempo médio entre falhas**, simbolizado por MTBF (iniciais da expressão da língua inglesa *mean time between failures*). Trata-se de um indicador que mostra o tempo médio de ocorrência de erros em determinado equipamento e que é calculado por:

$$MTBF = \frac{\text{Tempo total em que o equipamento funcionou}}{\text{Número de falhas no tempo em que o equipamento funcionou}}$$

Nesse sentido, imagine que, em dada indústria, no mês de abril, um equipamento que opere continuamente tenha funcionado por:

- 340 horas até apresentar a primeira falha;
- 270 horas até apresentar a segunda falha;
- 110 horas até apresentar a terceira falha.

Vamos chamar de:

- $x_1$ o tempo até a primeira falha;
- $x_2$ o tempo até a segunda falha;
- $x_3$ o tempo até a terceira falha;
- N o número total de falhas ocorridas.

No caso em estudo, temos:

- $x_1 = 340$
- $x_2 = 270$
- $x_3 = 110$
- $N = 3$

No caso em estudo, temos:

$$MTBF = \frac{x_1 + x_2 + x_3}{N} = \frac{x_1 + x_2 + x_3}{3}$$

$$MTBF = \frac{340 + 270 + 110}{3} = 240$$

Concluímos que, para o caso em estudo, o tempo médio entre falhas é de 240 horas. Esse resultado pode, por exemplo, orientar o gestor na decisão sobre manter o atual equipamento ou trocá-lo por outro que apresente maior MTBF.

### 2.1.2 Moda de um conjunto de dados

A moda Mo de um conjunto de dados é o valor mais frequente nesse conjunto, ou seja, é o valor que aparece mais vezes.

Um conjunto de dados pode não apresentar moda. Isso ocorre se todas as frequências são iguais ou se todas as frequências são diferentes.

Um conjunto de valores pode ter mais de uma moda. Isso ocorre se há mais de um valor predominante.

Por exemplo, vamos reescrever a tabela 3 na tabela 12, indicando em negrito o valor mais frequente.

**Tabela 12 – Frequências absolutas das quantidades de falhas nos algoritmos, com indicação em negrito do valor mais frequente**

| Quantidade de falhas (x) | Frequência absoluta (FA) |
|---|---|
| 0 | 1 |
| 1 | 3 |
| **2** | **8** |

| Quantidade de falhas (x) | Frequência absoluta (FA) |
|---|---|
| 3 | 4 |
| 4 | 0 |
| 5 | 1 |
| Total | N = 1 + 3 + 8 + 4 + 0 + 1 = 17 |

Vemos que o valor mais frequente da tabela 12, que aparece 8 vezes, é 2. Logo, a moda desse conjunto de dados é 2. Isso pode ser visualizado com clareza no gráfico da figura a seguir, em que a barra de maior altura, referente à frequência absoluta igual a 8, está associada a 2 falhas.

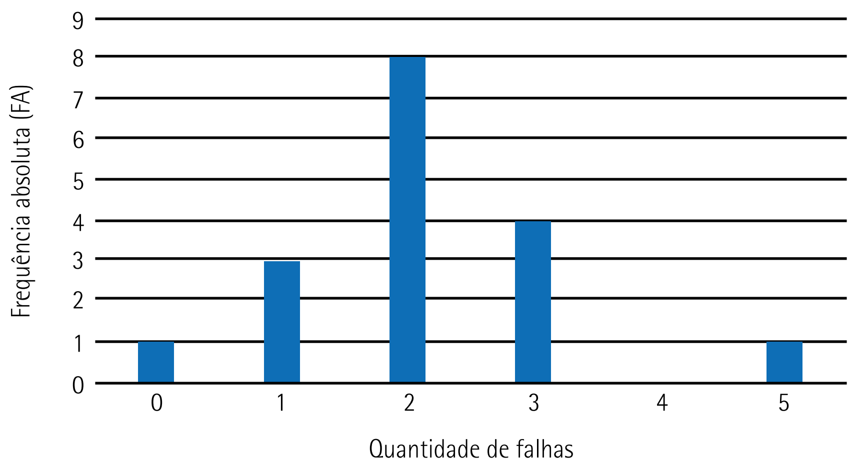



Figura 6 – Gráfico de barras: frequências absolutas das quantidade de falhas

Agora, vamos reescrever a tabela 7 na tabela 13, indicando em negrito o valor mais frequente.

**Tabela 13 – Notas bimestrais de Lucas em matemática, com indicação em negrito do valor mais frequente**

| Bimestre | Nota |
|---|---|
| 1º bimestre | 7,5 |
| **2º bimestre** | **6** |
| **3º bimestre** | **6** |
| 4º bimestre | 9 |

Vemos que o valor mais frequente da tabela 13, que aparece 2 vezes, é 6. Logo, a moda desse conjunto de dados é 6.

### 2.1.3 Mediana de um conjunto de dados

A mediana MD de um conjunto de dados escritos em ordem crescente é o valor central da distribuição de valores.

Assim, podemos concluir que a mediana é:

- o valor que ocupa a posição central da distribuição de dados se o número de elementos do conjunto for ímpar;

- a média aritmética dos dois valores centrais da distribuição de dados se o número de elementos do conjunto for par.

Voltemos à tabela 10. Vamos reescrevê-la ordenando os valores do menor para o maior. Isso foi feito na tabela 14.

**Tabela 14 – Conjunto de dados escrito em ordem crescente**

| Ordenamento | Valor |
|---|---|
| 1º valor | –51,7 |
| 2º valor | –47 |
| 3º valor | –7,3 |
| 4º valor | –0,8 |
| **5º valor** | **3** |
| 6º valor | 12,5 |
| 7º valor | 21 |
| 8º valor | 23 |
| 9º valor | 32 |

No caso da tabela 14, como temos 9 dados, a mediana é o 5º valor, ou seja, MD = 3.

Agora, vamos analisar um caso em que temos número par de valores no conjunto de dados, como ocorre na tabela 9. Para isso, reescrevemos essa tabela, com os valores colocados em ordem crescente, na tabela 15.

**Tabela 15 – Pesos dos recém-nascidos em ordem crescente**

| Ordenamento | Peso (kg) |
|---|---|
| 1º valor | 2,97 |
| 2º valor | 2,98 |
| 3º valor | 3,04 |
| 4º valor | 3,05 |
| 5º valor | 3,09 |
| **6º valor** | **3,12** |
| **7º valor** | **3,13** |
| 8º valor | 3,13 |
| 9º valor | 3,18 |
| 10º valor | 3,21 |

| Ordenamento | Peso (kg) |
|---|---|
| 11º valor | 3,22 |
| 12º valor | 3,26 |

No caso da tabela 15, como temos 12 dados, a mediana é a média aritmética do 6º valor e do 7º valor (soma desses valores dividida por 2), o que resulta em MD = 3,125, conforme calculado a seguir.

$$MD = \frac{(6^{o}\,\text{valor}) + (7^{o}\,\text{valor})}{2} = \frac{3{,}12 + 3{,}13}{2} = 3{,}125$$

Para fixar o que vimos, vamos analisar o gráfico da figura 7, que mostra as quantidades de gols feitos por artilheiros das Copas do Mundo de 1930 a 2006.

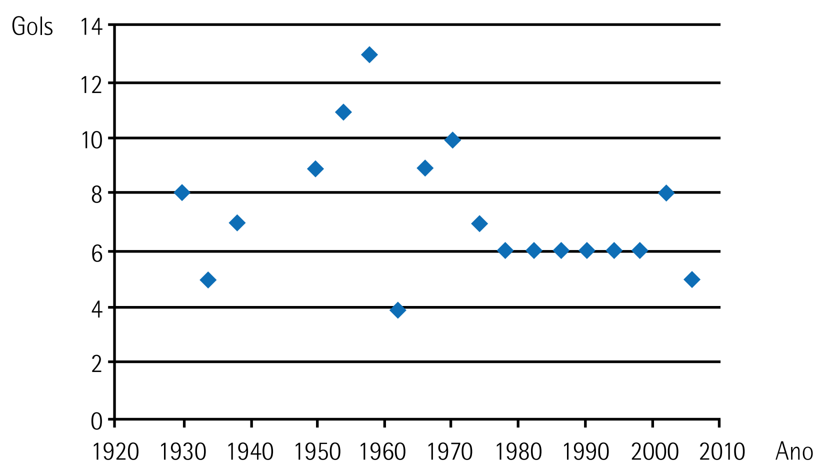



Figura 7 – Quantidades de gols dos artilheiros das Copas do Mundo de 1930 a 2006

Adaptada de: https://bit.ly/2Z1Vd9W. Acesso em: 17 nov. 2021.

Com base no gráfico apresentado, podemos fazer as contagens mostradas na tabela 16. Nessa tabela, para cada quantidade de gols, há o correspondente número de vezes que ela aparece no gráfico, ou seja, a sua frequência absoluta.

**Tabela 16 – Quantidades de gols e frequências absolutas**

| Quantidade de gols | Número de vezes que aparece no gráfico (frequência absoluta da medida) |
|---|---|
| $x_1 = 4$ | $FA_1 = 1$ |
| $x_2 = 5$ | $FA_2 = 2$ |
| $x_3 = 6$ | $FA_3 = 6$ |
| $x_4 = 7$ | $FA_4 = 2$ |
| $x_5 = 8$ | $FA_5 = 2$ |
| $x_6 = 9$ | $FA_6 = 2$ |

| Quantidade de gols | Número de vezes que aparece no gráfico (frequência absoluta da medida) |
|---|---|
| $x_7 = 10$ | $FA_7 = 1$ |
| $x_8 = 11$ | $FA_8 = 1$ |
| $x_9 = 13$ | $FA_9 = 1$ |

Na figura 8, temos a representação gráfica dos dados da tabela 16.

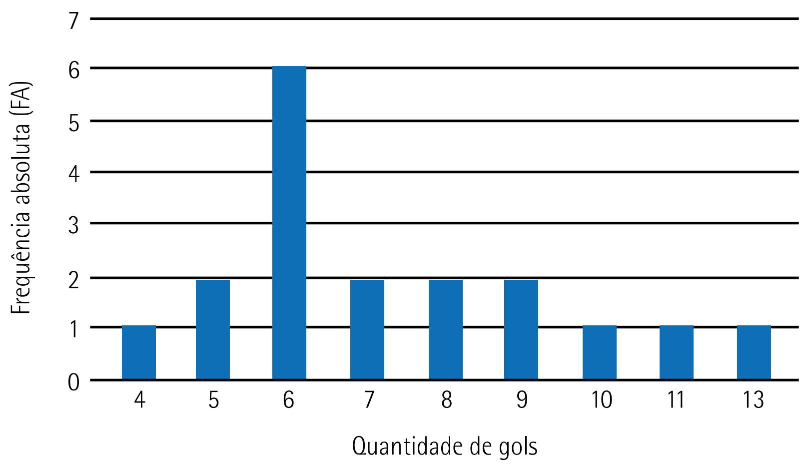



Figura 8 – Gráfico das frequências absolutas das quantidades de gols dos artilheiros das Copas do Mundo de 1930 a 2006

Podemos determinar a média $\overline{x}_Q$ da quantidade de gols, conforme calculado a seguir.

$$\overline{x}_Q = \frac{x_1 . FA_1 + x_2 . FA_2 + x_3 . FA_3 + x_4 . FA_4 + x_5 . FA_5 + x_6 . FA_6 + x_7 . FA_7 + x_8 . FA_8 + x_9 . FA_9}{18}$$

$$\overline{x}_Q = \frac{4 . 1 + 5 . 2 + 6 . 6 + 7 . 2 + 8 . 2 + 9 . 2 + 10 . 1 + 11 . 1 + 13 . 1}{18}$$

$$\overline{x}_Q = \frac{4 + 10 + 36 + 14 + 16 + 18 + 10 + 11 + 13}{18} = \frac{132}{18} = 7{,}33$$

A moda da quantidade de gols é 6, que é o valor mais frequente.

Se ordenarmos de forma crescente as 18 contagens de gols que aparecem na tabela, os dois valores centrais serão 6 e 7 gols.

Para determinarmos a mediana de um conjunto par de valores (18), precisamos calcular a média dos dois valores centrais, no caso a média entre 6 e 7. Vejamos:

$$\text{Mediana} = \frac{6+7}{2} = \frac{13}{2} = 6{,}5 \text{ gols}$$

Logo, a partir dos dados apresentados no gráfico, concluímos que a mediana das quantidades de gols marcados pelos artilheiros das Copas do Mundo de 1930 a 2006 é igual a 6,5 gols.

## 2.2 Medidas de dispersão de um conjunto de dados: amplitude, variância e desvio padrão

Podemos estar interessados em avaliar de modo mais amplo o comportamento de um conjunto de valores. Para isso, não é suficiente calcularmos apenas as medidas de tendência central.

Nesse âmbito, temos as medidas de dispersão, como a amplitude, a variância e o desvio padrão, que serão estudadas a seguir no contexto de situações que envolvem um conjunto de dados.

### 2.2.1 Amplitude de um conjunto de dados

A amplitude A de um conjunto de dados é a diferença entre o maior valor observado (MVO) e o menor valor observado (mvo).

Por exemplo, vamos reescrever a tabela 9 na tabela 17, indicando em negrito o MVO e o mvo dos pesos dos recém-nascidos.

**Tabela 17 – Pesos dos recém-nascidos em janeiro em determinada maternidade, com o MVO e o mvo destacados em negrito**

| Nome do recém-nascido | Peso do recém-nascido (kg) |
|---|---|
| Alice | 3,05 |
| Hugo | 3,12 |
| **Laura** | **2,97 (mvo)** |
| Silas | 3,21 |
| Tiago | 3,18 |
| Rodrigo | 3,09 |
| Carina | 2,98 |
| Gisele | 3,13 |
| Viviane | 3,04 |
| Ricardo | 3,13 |
| **Ulisses** | **3,26 (MVO)** |
| Marta | 3,22 |

A amplitude A do conjunto de dados da tabela 17 é igual a 0,29, pois:

A = MVO – mvo = 3,26 – 2,97 = 0,29

Podemos interpretar o seguinte: na situação em foco, a diferença de peso entre o "maior recém-nascido" e o "menor recém-nascido" é de 0,29 kg, ou seja, 290 gramas.

### 2.2.2 Variância e desvio padrão de um conjunto de dados

Vejamos o exemplo mostrado na tabela 18, em que temos as notas de 3 estudantes, Marcos, Vanessa e Leila, em 4 provas de ciências.

**Tabela 18 – Notas de 3 estudantes em 4 provas de ciências**

| Estudante | Notas em ciências | | | | Média das notas |
|---|---|---|---|---|---|
| Marcos | 4,75 | 5,25 | 5,5 | 4,5 | 5 |
| Vanessa | 5 | 5 | 5 | 5 | 5 |
| Leila | 0 | 0 | 10 | 10 | 5 |

Na tabela, observamos que os 3 estudantes (Marcos, Vanessa e Leila) obtiveram a mesma média em ciências, igual a 5, conforme calculado a seguir.

Marcos:

$$\overline{x}_{CD} = \frac{4,75 + 5,25 + 5,5 + 4,5}{4} = \frac{20}{4} = 5$$

Vanessa:

$$\overline{x}_{CD} = \frac{5 + 5 + 5 + 5}{4} = \frac{20}{4} = 5$$

Leila:

$$\overline{x}_{CD} = \frac{0 + 0 + 10 + 10}{4} = \frac{20}{4} = 5$$

Contudo, as notas de cada um dos estudantes nas provas foram bem diferentes.

- As notas do Marcos variaram de 4,5 a 5,5.

- As notas da Vanessa não variaram, pois ela tirou 5 em todas as provas.

- As notas da Leila variaram de 0 a 10.

Podemos fazer uma comparação entre as notas dos 3 estudantes e, assim, verificar que:

- a Vanessa teve a menor variação de notas (amplitude igual a 0);
- o Marcos teve variação intermediária de notas (amplitude igual a 1);
- a Leila teve a maior variação de notas (amplitude igual a 10).

O desvio padrão de um conjunto de dados, indicado por DP, é um valor que mostra se a variabilidade de um conjunto de valores é grande ou é pequena em relação à média dos valores. Assim, vejamos as situações a seguir.

- Se o desvio padrão de um conjunto de dados é zero (DP = 0), isso significa que esse conjunto é formado por valores idênticos (como no caso das 4 notas da Vanessa em ciências).
- Se o desvio padrão de um conjunto de dados é "pequeno" quando comparado à média, isso significa que esse conjunto é formado por valores "pouco dispersos", que variam pouco (como no caso das 4 notas do Marcos em ciências).
- Se o desvio padrão de um conjunto de dados é "grande" quando comparado à média, significa que esse conjunto é formado por valores "muito dispersos", que variam muito (como no caso das 4 notas da Leila em ciências).

Você deve estar se perguntando: como fazemos para calcular o desvio padrão de um conjunto de dados?

Primeiramente, vamos aprender a calcular uma medida de dispersão chamada de **variância** e indicada por VAR.

Para isso, suponha que você queira calcular a variância do conjunto formado pelos 3 dados mostrados na tabela 19.

**Tabela 19 – Conjunto de dados para cálculo de variância**

| Medida | Valor |
|---|---|
| $x_1$ | –2 |
| $x_2$ | 221 |
| $x_3$ | 56 |

Antes de calcularmos a variância dos dados da tabela 19, precisamos calcular a média desse conjunto. Conforme vimos:

$$\overline{x}_{CD} = \frac{x_1 + x_2 + x_3}{3}$$

$$\overline{x}_{CD} = \frac{(-2) + 221 + 56}{3} = \frac{275}{3} = 91,67$$

A variância dos valores apresentados na tabela anterior, indicada por VAR, é calculada pela expressão a seguir.

$$VAR = \frac{\left(x_1 - \overline{x}_{CD}\right)^2 + \left(x_2 - \overline{x}_{CD}\right)^2 + \left(x_3 - \overline{x}_{CD}\right)^2}{3}$$

$$VAR = \frac{\left((-2) - (91,67)\right)^2 + \left((221) - (91,67)\right)^2 + \left((56) - (91,67)\right)^2}{3}$$

$$VAR = \frac{\left(-93,67\right)^2 + \left(129,33\right)^2 + \left(-35,67\right)^2}{3}$$

$$VAR = \frac{8774,0689 + 16726,2489 + 1272,3489}{3} = \frac{26772,6667}{3} = 8924,2222$$

O desvio padrão é a raiz quadrada da variância. Logo, para o caso em estudo, o DP é:

$$DP = \sqrt{VAR} = \sqrt{8924,2222} = 94,47$$

De modo geral, a variância VAR das N medidas de um conjunto de dados, em que $x_i$ representa uma medida qualquer (lida como "i-ésima medida"), é dada por:

$$VAR = \frac{\left(x_1 - \overline{x}_{CD}\right)^2 + \left(x_2 - \overline{x}_{CD}\right)^2 + \ldots + \left(x_i - \overline{x}_{CD}\right)^2 + \ldots + \left(x_N - \overline{x}_{CD}\right)^2}{N}$$

A igualdade anterior também pode ser escrita como:

$$VAR = \frac{\sum_{i=1}^{N} \left(x_i - \overline{x}_{CD}\right)^2}{N}$$

Na expressão, $\sum_{i=1}^{N} \left(x_i - \overline{x}_{CD}\right)^2$ é lido como "somatória do quadrado de xis i menos xis barra, com i começando em 1 e indo até N".

O desvio padrão DP das N medidas de um conjunto de dados é dado por:

$DP = \sqrt{VAR}$

**Observação**

Veja que, independentemente dos sinais dos valores que compõem o conjunto de dados e do sinal da média (valor médio), a variância é sempre um número positivo, pois $(x_i - \bar{x}_{CD})$ está elevado ao quadrado.

Voltemos ao caso das notas do Marcos, da Vanessa e da Leila. Vamos calcular os desvios padrão das notas de cada um desses estudantes. Já sabemos que suas notas médias $\bar{x}_{CD}$ são iguais a 5.

**Tabela 20**

| Estudante | Notas em ciências | | | |
|---|---|---|---|---|
| Marcos | $x_1 = 4{,}75$ | $x_2 = 5{,}25$ | $x_3 = 5{,}5$ | $x_4 = 4{,}5$ |

$$VAR(Marcos) = \frac{(x_1 - \bar{x}_{CD})^2 + (x_2 - \bar{x}_{CD})^2 + (x_3 - \bar{x}_{CD})^2 + (x_4 - \bar{x}_{CD})^2}{4}$$

$$VAR(Marcos) = \frac{((4{,}75) - (5))^2 + ((5{,}25) - (5))^2 + ((5{,}5) - (5))^2 + ((4{,}5) - (5))^2}{4}$$

$$VAR(Marcos) = \frac{(-0{,}25)^2 + (0{,}25)^2 + (0{,}5)^2 + (-0{,}5)^2}{4}$$

$$VAR(Marcos) = \frac{0{,}0625 + 0{,}0625 + 0{,}25 + 0{,}25}{4} = \frac{0{,}625}{4}$$

$VAR(Marcos) = 0{,}15625$

$DP(Marcos) = \sqrt{0{,}15625} = 0{,}3953$

**Tabela 21**

| Estudante | Notas em ciências | | | |
|---|---|---|---|---|
| Vanessa | $x_1 = 5$ | $x_2 = 5$ | $x_3 = 5$ | $x_4 = 5$ |

$$VAR(Vanessa) = \frac{(x_1 - \overline{x}_{CD})^2 + (x_2 - \overline{x}_{CD})^2 + (x_3 - \overline{x}_{CD})^2 + (x_4 - \overline{x}_{CD})^2}{4}$$

$$VAR(Vanessa) = \frac{((5)-(5))^2 + ((5)-(5))^2 + ((5)-(5))^2 + ((5)-(5))^2}{4}$$

$$VAR(Vanessa) = \frac{(0)^2 + (0)^2 + (0)^2 + (0)^2}{4} = \frac{0}{4}$$

VAR(Vanessa) = 0

DP(Vanessa) = $\sqrt{0}$ = 0

**Tabela 22**

| Estudante | Notas em ciências | | | |
|---|---|---|---|---|
| Leila | $x_1 = 0$ | $x_2 = 0$ | $x_3 = 10$ | $x_4 = 10$ |

$$VAR(Leila) = \frac{(x_1 - \overline{x}_{CD})^2 + (x_2 - \overline{x}_{CD})^2 + (x_3 - \overline{x}_{CD})^2 + (x_4 - \overline{x}_{CD})^2}{4}$$

$$VAR(Leila) = \frac{((0)-(5))^2 + ((0)-(5))^2 + ((10)-(5))^2 + ((10)-(5))^2}{4}$$

$$VAR(Leila) = \frac{(-5)^2 + (-5)^2 + (5)^2 + (5)^2}{4} = \frac{25+25+25+25}{4} = \frac{100}{4}$$

VAR(Leila) = 25

DP(Leila) = $\sqrt{25}$ = 5

Assim, vemos que:

- a Vanessa tem média igual a 5 em ciências, com desvio padrão igual a 0 (variação nula de notas);
- o Marcos tem média igual a 5 em ciências, com desvio padrão igual a 0,3953 (variação intermediária de notas);
- a Leila tem média igual a 5 em ciências, com desvio padrão igual a 5 (variação alta de notas).

**Saiba mais**

Para saber mais sobre a notação de somatório, visite a página indicada.

NOTAÇÃO de somatório. *Khan Academy*, [s.d.].

Disponível em: https://bit.ly/3CuCzFd. Acesso em: 17 nov. 2021.

## 3 PROBABILIDADE: PARTE 1

### 3.1 Noções gerais sobre probabilidade

Você já deve ter ouvido frases como as seguintes.

- A probabilidade de chover amanhã é de 70%.
- Você tem chance de 1 em 50.063.860 de acertar na Mega-Sena se fizer uma aposta de 6 números.


Figura 9 – A chance de ganhar na Mega-Sena com uma aposta de 6 números é de 1 em 50.063.860.

Disponível em: https://bit.ly/3FuqKAK. Acesso em: 17 nov. 2021.

Veremos que o estudo das probabilidades é importante para a modelagem de fenômenos aleatórios, ou seja, de situações nas quais a incerteza está presente ou de acontecimentos cujos resultados não podem ser previstos com certeza (MAGALHÃES; LIMA, 2008).

**Observação**

O lançamento de uma moeda honesta é um exemplo típico de fenômeno aleatório. Antes de fazermos o lançamento, não podemos dizer, com certeza, se teremos cara ou coroa.

Uma maneira de calcularmos a probabilidade P de sucesso em um evento aleatório, como ganharmos um prêmio em um sorteio, é fazermos a divisão do "número de casos favoráveis para nosso sucesso" pelo "número total de casos possíveis". Ou seja:

$$P = P(\text{sucesso}) = \frac{\text{Número de casos favoráveis}}{\text{Número total de casos possíveis}}$$

Como o número de casos favoráveis é menor do que o número total de casos possíveis ou igual ao número total de casos possíveis, a probabilidade P é um número real que varia de 0 até 1 ($0 \leq P \leq 1$). Muitas vezes, a probabilidade P é multiplicada por 100% para expressarmos o resultado em percentual (%).

Vamos pensar em um caso bastante conhecido, o lançamento de um dado honesto, cujas faces estejam numeradas de 1 a 6, como mostrado na figura 10.

Figura 10 – Dado honesto com faces numeradas de 1 a 6

Disponível em: https://bit.ly/3ntevyf. Acesso em: 17 nov. 2021.

Qual é a probabilidade de que, em um único lançamento de um dado honesto, obtenhamos o resultado 3 (face 3 voltada para cima)?

O número total de casos possíveis é igual a 6, pois você pode obter os seguintes resultados: 1, 2, 3, 4, 5 ou 6.

Os resultados possíveis de um experimento também são conhecidos como **espaço amostral E**. No caso em estudo, temos que $E = \{1, 2, 3, 4, 5, 6\}$.

De todos os possíveis resultados, ou seja, de todo o espaço amostral E, somente se considera sucesso a obtenção do resultado 3. Ou seja, dos 6 resultados possíveis, apenas um deles é favorável.

Logo, a probabilidade P de que, em um único lançamento, obtenhamos o resultado 3 (face 3 voltada para cima) é de 1 em 6, ou seja, 1/6, conforme detalhado a seguir.

$$P = P(\text{sucesso}) = \frac{\text{Número de casos favoráveis}}{\text{Número total de casos possíveis}} = \frac{1}{6}$$

Podemos, também, expressar o resultado anterior em termos de percentuais:

$$P = P(\text{sucesso}) = \frac{\text{Número de casos favoráveis}}{\text{Número total de casos possíveis}} \,.\, 100\% = \frac{1}{6} \,.\, 100\% = 16{,}7\%$$

Continuaremos pensando no lançamento de um dado honesto, cujas faces estejam numeradas de 1 a 6. Qual é a probabilidade de que, em um único lançamento, obtenhamos como resultado um número ímpar?

O número total de casos possíveis permanece igual a 6, pois você pode obter os seguintes resultados: 1, 2, 3, 4, 5 ou 6. Ou seja, o espaço amostral E é $E = \{1, 2, 3, 4, 5, 6\}$.

Dos possíveis resultados, ou seja, de todo o espaço amostral E, somente se considera sucesso a obtenção dos resultados ímpares, ou seja, 1, 3 ou 5. Isto é, dos 6 resultados possíveis, apenas 3 são favoráveis.

Logo, a probabilidade P de que, em um único lançamento, obtenhamos como resultado um número ímpar é de 3 em 6, ou seja, ½ ou 0,5 ou 50%, conforme detalhado a seguir.

$$P = P(\text{sucesso}) = \frac{\text{Número de casos favoráveis}}{\text{Número total de casos possíveis}} = \frac{3}{6} = \frac{1}{2} = 0{,}5$$

$$P = P(\text{sucesso}) = \frac{\text{Número de casos favoráveis}}{\text{Número total de casos possíveis}} \,.\, 100\% = \frac{3}{6} \,.\, 100\% = 50\%$$

Vamos continuar lançando um dado honesto. Qual é a probabilidade de que, em dois lançamentos independentes, obtenhamos como resultado favorável a soma dos números das faces voltadas para cima em cada um dos lançamentos ser maior do que 5?

A tabela 23 mostra, para o caso em estudo, todos os resultados possíveis e os casos em que há resultado favorável.

**Tabela 23 – Resultados possíveis e indicação dos favoráveis**

| 1º lançamento | 2º lançamento | Soma das faces | Sucesso? |
|---|---|---|---|
| 1 | 1 | 2 | Não |
| 1 | 2 | 3 | Não |
| 1 | 3 | 4 | Não |

| 1º lançamento | 2º lançamento | Soma das faces | Sucesso? |
|---|---|---|---|
| 1 | 4 | 5 | Não |
| 1 | 5 | 6 | Sim |
| 1 | 6 | 7 | Sim |
| 2 | 1 | 3 | Não |
| 2 | 2 | 4 | Não |
| 2 | 3 | 5 | Não |
| 2 | 4 | 6 | Sim |
| 2 | 5 | 7 | Sim |
| 2 | 6 | 8 | Sim |
| 3 | 1 | 4 | Não |
| 3 | 2 | 5 | Não |
| 3 | 3 | 6 | Sim |
| 3 | 4 | 7 | Sim |
| 3 | 5 | 8 | Sim |
| 3 | 6 | 9 | Sim |
| 4 | 1 | 5 | Não |
| 4 | 2 | 6 | Sim |
| 4 | 3 | 7 | Sim |
| 4 | 4 | 8 | Sim |
| 4 | 5 | 9 | Sim |
| 4 | 6 | 10 | Sim |
| 5 | 1 | 6 | Sim |
| 5 | 2 | 7 | Sim |
| 5 | 3 | 8 | Sim |
| 5 | 4 | 9 | Sim |
| 5 | 5 | 10 | Sim |
| 5 | 6 | 11 | Sim |
| 6 | 1 | 7 | Sim |
| 6 | 2 | 8 | Sim |
| 6 | 3 | 9 | Sim |
| 6 | 4 | 10 | Sim |
| 6 | 5 | 11 | Sim |
| 6 | 6 | 12 | Sim |

O número total de casos possíveis é igual a 36, dos quais 26 representam sucesso para a condição que escolhemos (em 2 lançamentos independentes de um dado honesto, obtemos a soma dos números das faces voltadas para cima em cada um dos lançamentos maior do que 5). Logo, a probabilidade P de sucesso é de 26 casos em 36 casos, ou seja, 13/18 ou 72,2%, pois:

$$P = P(\text{sucesso}) = \frac{\text{Número de casos favoráveis}}{\text{Número total de casos possíveis}} = \frac{26}{36} = \frac{13}{18} = 0{,}722$$

$$P = P(\text{sucesso}) = \frac{\text{Número de casos favoráveis}}{\text{Número total de casos possíveis}} \,.\, 100\% = \frac{26}{36} \,.\, 100\% = 72{,}2\%$$

Veja que, para a situação em estudo, há 10 casos desfavoráveis (em 2 lançamentos independentes de um dado honesto, obtemos a soma dos números das faces voltadas para cima em cada um dos lançamentos igual a 5 ou menor do que 5). Logo, a probabilidade P de fracasso é de 10 casos em 36, ou seja, 5/18 ou 27,8%, pois:

$$P = P(\text{fracasso}) = \frac{\text{Número de casos desfavoráveis}}{\text{Número total de casos possíveis}} = \frac{10}{36} = \frac{5}{18} = 0{,}278$$

$$P = P(\text{sucesso}) = \frac{\text{Número de casos favoráveis}}{\text{Número total de casos possíveis}} \,.\, 100\% = \frac{10}{36} \,.\, 100\% = 27{,}8\%$$

De modo geral, se P(X) representar a função de probabilidade do evento X e E representar o espaço amostral, teremos:

- $0 \leq P(X) \leq 1$
- $P(X = E) = P(E) = 1$

Em termos de percentuais, podemos escrever:

- $0 \leq P(X) \leq 100\%$
- $P(X = E) = P(E) = 100\%$

Observe que a probabilidade de ocorrência de um evento que inclua todo o espaço amostral é igual a 1 (ou 100%). Por exemplo, no caso de um lançamento de um dado honesto, a probabilidade de obtermos as faces 1, 2, 3, 4, 5 ou 6 voltadas para cima, o que constitui todo o espaço amostral, é igual a 1.

Vamos estudar um exemplo de cálculo de probabilidades.

Imagine que na festa de formatura de uma turma que concluiu o curso de Ciência da Computação cada um dos 45 formandos tenha levado 1, 2, 3 ou 4 acompanhantes, conforme mostrado no gráfico da figura 11.

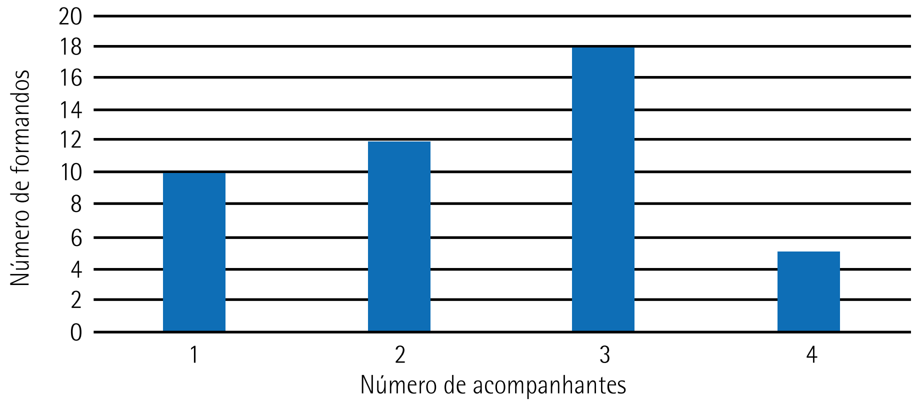



Figura 11 – Número de formandos e número de acompanhantes na festa de formatura da turma de Ciência da Computação

Qual é a quantidade total de acompanhantes que foram levados à festa?

Pela leitura do gráfico, vemos o que segue.

- 10 formandos levaram apenas 1 acompanhante cada, ou seja, juntos, eles levaram 10 . 1 = 10 acompanhantes.

- 12 formandos levaram 2 acompanhantes cada, ou seja, juntos, eles levaram 12 . 2 = 24 acompanhantes.

- 18 formandos levaram 3 acompanhantes cada, ou seja, juntos, eles levaram 18 . 3 = 54 acompanhantes.

- 5 formandos levaram 4 acompanhantes cada, ou seja, juntos, eles levaram 5 . 4 = 20 acompanhantes.

Logo, a quantidade total de acompanhantes é 108, pois:

10 + 24 + 54 + 20 = 108

Na festa, um dos convidados foi sorteado para fazer um brinde à turma. Qual é a probabilidade de que o acompanhante sorteado para fazer o brinde tenha sido levado por um formando que levou 4 acompanhantes?

Do gráfico dado, verificamos que a quantidade total de acompanhantes que os formandos levaram à festa foi igual a 108. Desse total de 108 acompanhantes, 20 foram levados por formandos que levaram 4 acompanhantes.

Logo, a probabilidade P de que o acompanhante sorteado para fazer o brinde à turma tenha sido levado por um formando que levou 4 acompanhantes é de 20 em 108, ou seja, 5/27 ou 18,5%, conforme calculado a seguir.

$$P = \frac{20}{108} = \frac{5}{27} \text{ ou } P = \frac{20}{108} . 100\% = 18{,}5\%$$

Vamos pensar em mais uma situação. Imagine que, nas faces de um cubo, estejam impressos os números mostrados na figura a seguir.

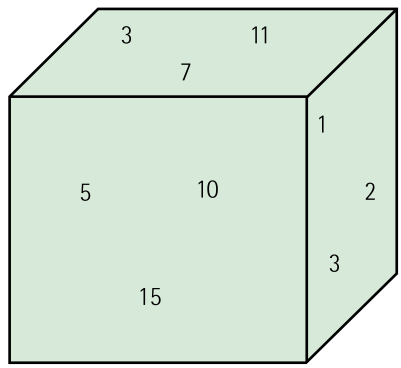



Figura 12 – Números em faces de um cubo

Nas faces opostas às mostradas na figura, que não podem ser nela visualizadas, temos a repetição dos números indicados.

Se jogarmos esse cubo de modo aleatório, qual é a probabilidade de obtermos a soma dos números da face voltada para cima maior do que 20?

Na tabela 24, estão mostrados os números gravados em cada face do cubo, as suas correspondentes somas de valores e as indicações de obtermos ou não soma dos números da face voltada para cima maior do que 20 no lançamento do cubo. Veja que as faces do cubo foram nomeadas como faces 1, 2, 3, 4, 5 e 6.

**Tabela 24 – Números gravados em cada face do cubo**

| Face | Números | Soma | Soma maior do que 20? |
|---|---|---|---|
| Face 1 | 5, 10 e 15 | 5 + 10 + 15 = 30 | Sim |
| Face 2 | 3, 7 e 11 | 3 + 7 + 11 = 21 | Sim |
| Face 3 | 1, 2 e 3 | 1 + 2 + 3 = 6 | Não |
| Face 4 | 5, 10 e 15 | 5 + 10 + 15 = 30 | Sim |
| Face 5 | 3, 7 e 11 | 3 + 7 + 11 = 21 | Sim |
| Face 6 | 1, 2 e 3 | 1 + 2 + 3 = 6 | Não |

Pelas informações contidas na tabela, verificamos que a probabilidade de que, ao jogarmos aleatoriamente o cubo em foco, a face voltada para cima tenha a soma de números maior do que 20 é de 2/3, conforme calculado a seguir.

$$P(\text{soma} > 20) = \frac{\text{Número de faces com soma maior do que 20}}{\text{Número total de faces}} = \frac{4}{6} = \frac{2}{3}$$

## 3.2 Probabilidade da união e da intersecção de eventos

Vamos indicar por A e por B dois eventos do espaço amostral E, cujas probabilidades de ocorrência são, respectivamente, P(A) e P(B).

A probabilidade de ocorrência de A ou B, ou seja, a probabilidade da ocorrência da união dos eventos A e B, indicada por $P(A \cup B)$, é dada por:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Na igualdade acima, chamada de regra de adição de probabilidades, $P(A \cap B)$ indica a probabilidade de ocorrência de A e B, ou seja, a probabilidade de ocorrência da intersecção dos eventos A e B.

Isso ficará mais claro se pensarmos na situação fictícia expressa na tabela 25, em que temos os números de estudantes que cursam, por turno, Ciência da Computação, Matemática e Administração em determinada localidade.

**Tabela 25 – Números de estudantes que cursam, por turno, Ciência da Computação, Matemática e Administração em determinada localidade**

| Curso | Diurno | Noturno | Total |
|---|---|---|---|
| Ciência da Computação | 1253 | 2356 | 3609 |
| Matemática | 238 | 327 | 565 |
| Administração | 1126 | 983 | 2109 |
| Total | 2617 | 3666 | 6283 |

Da tabela 25, vemos que nosso universo é de 6283 estudantes.

Imagine que, desse total de 6283 estudantes, façamos o sorteio de 1 aluno.

Nesse caso, vamos definir os eventos a seguir.

- **C**: o estudante sorteado cursa Ciência da Computação.
- **M**: o estudante sorteado cursa Matemática.
- **A**: o estudante sorteado cursa Administração.

- **D**: o estudante sorteado cursa o turno diurno.
- **N**: o estudante sorteado cursa o turno noturno.

Vejamos as probabilidades de cada um dos eventos definidos anteriormente.

- $P(C) = \dfrac{3609}{6283} = 0{,}5744$
- $P(M) = \dfrac{565}{6283} = 0{,}0899$
- $P(A) = \dfrac{2109}{6283} = 0{,}3357$
- $P(D) = \dfrac{2617}{6283} = 0{,}4165$
- $P(N) = \dfrac{3666}{6283} = 0{,}5835$

Observamos que a probabilidade de o estudante sorteado cursar Ciência da Computação, Matemática ou Administração é igual a 1, o que pode ser confirmado pelo cálculo a seguir.

$P(C) + P(M) + P(A) = 0{,}5744 + 0{,}0899 + 0{,}3357 = 1$

Verificamos que a probabilidade de o estudante sorteado cursar turno diurno ou turno noturno é igual a 1, o que pode ser confirmado pelo cálculo a seguir.

$P(D) + P(N) = 0{,}4165 + 0{,}5835 = 1$

Qual é a probabilidade de o estudante sorteado cursar Ciência da Computação no turno diurno?

O que estamos querendo encontrar é a probabilidade da intersecção entre os eventos "o estudante sorteado cursa Ciência da Computação" e "o estudante sorteado estuda no turno diurno", indicada por $P(C \cap D)$.

Para calcular tal probabilidade, observamos a tabela 25. Nela, vemos que, do total de 6283 estudantes, 1253 cursam Ciência da Computação no turno diurno. Logo:

$$P(C \cap D) = \frac{1253}{6283} = 0{,}1994$$

Assim, para a situação em análise, a probabilidade de o estudante sorteado cursar Ciência da Computação no turno diurno é de quase 20%.

Qual é a probabilidade de o estudante sorteado cursar Ciência da Computação ou estudar no turno diurno?

O que estamos querendo encontrar é a probabilidade da união entre os eventos "o estudante sorteado cursa Ciência da Computação" e "o estudante sorteado estuda no turno diurno", indicada por $P(C \cup D)$.

Para calcular tal probabilidade, podemos usar a regra de adição de probabilidades:

$P(C \cup D) = P(C) + P(D) - P(C \cap D)$

$P(C \cup D) = 0{,}5744 + 0{,}4165 - 0{,}1994 = 0{,}7915$

Assim, para a situação em análise, a probabilidade de o estudante sorteado cursar Ciência da Computação ou estudar no turno diurno é pouco superior a 79%.

### 3.3 Probabilidade condicional

Vamos indicar por A e por B dois eventos do espaço amostral E, cujas probabilidades de ocorrência são, respectivamente, P(A) e P(B).

A probabilidade de ocorrência de A dado que B ocorreu, ou seja, a probabilidade condicional de A dada a ocorrência de B, indicada por P(A/B) e lida como "pê de A dado B", é dada por:

$$P(A / B) = \frac{P(A \cap B)}{P(B)}$$

Vemos que, nessa igualdade, precisamos atender à condição $P(B) \neq 0$.

Podemos pensar que calcular uma probabilidade condicional equivale a termos uma redução no espaço amostral.

Tomemos novamente a tabela 25. Qual é a probabilidade de sortearmos um estudante do curso de Ciência da Computação sabendo que esse estudante estuda no turno diurno?

Estamos em uma situação na qual devemos usar o conceito de probabilidade condicional, visto que queremos determinar a probabilidade de o estudante sorteado ser do curso de Ciência da Computação dado que ele estuda no turno diurno, indicada por P(C/D).

Assim:

$$P(C / D) = \frac{P(C \cap D)}{P(D)} = \frac{0{,}1994}{0{,}4165} = 0{,}4787$$

Também poderíamos ter calculado essa probabilidade usando diretamente os dados da tabela 25:

$$P(C/D) = \frac{\text{Quantidade de estudantes de Ciência da Computação no diurno}}{\text{Quantidade total de estudantes no diurno}}$$

$$P(C/D) = \frac{1253}{2617} = 0{,}4787$$

Vejamos agora outra situação. Imagine que, em uma urna, tenhamos 2 bolas cinza e 3 bolas verdes, conforme esquematizado na figura 13.

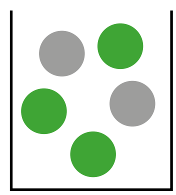


Figura 13 – Urna com 2 bolas cinza e 3 bolas verdes

Dessa urna, que contém o total de 5 bolas, retiramos ao acaso e sucessivamente 2 bolas, sem fazer a reposição da bola inicialmente retirada. Qual é a probabilidade de retirarmos duas bolas verdes?

Primeiramente, vamos fazer o que chamamos de árvore de probabilidades, em que:

- P(C) é a probabilidade de retirarmos uma bola cinza;
- P(V) é a probabilidade de retirarmos uma bola verde;
- P(C/C) é a probabilidade de retirarmos uma bola cinza dado que já retiramos uma bola cinza;
- P(C/V) é a probabilidade de retirarmos uma bola cinza dado que já retiramos uma bola verde;
- P(V/V) é a probabilidade de retirarmos uma bola verde dado que já retiramos uma bola verde;
- P(V/C) é a probabilidade de retirarmos uma bola verde dado que já retiramos uma bola cinza;
- $P(C \cap C)$ é a probabilidade de retirarmos uma bola cinza e, depois, outra bola cinza;
- $P(C \cap V)$ é a probabilidade de retirarmos uma bola cinza e, depois, uma bola verde;
- $P(V \cap V)$ é a probabilidade de retirarmos uma bola verde e, depois, outra bola verde;
- $P(V \cap C)$ é a probabilidade de retirarmos uma bola verde e, depois, uma bola cinza.

No caso em estudo, temos o seguinte.

- $P(C) = \dfrac{\text{Número de bolas cinza}}{\text{Número total de bolas}} = \dfrac{2}{5}$

- $P(V) = \dfrac{\text{Número de bolas verdes}}{\text{Número total de bolas}} = \dfrac{3}{5}$

Depois da primeira retirada, como não há reposição, sobram 4 das 5 bolas e ficamos com:

- $P(C / C) = \dfrac{\text{Número de bolas cinza depois de retirarmos uma bola cinza}}{\text{Número total de bolas restantes}} = \dfrac{1}{4}$

- $P(C / V) = \dfrac{\text{Número de bolas cinza depois de retirarmos uma bola verde}}{\text{Número total de bolas restantes}} = \dfrac{2}{4} = \dfrac{1}{2}$

- $P(V / V) = \dfrac{\text{Número de bolas verdes depois de retirarmos uma bola verde}}{\text{Número total de bolas restantes}} = \dfrac{2}{4} = \dfrac{1}{2}$

- $P(V / C) = \dfrac{\text{Número de bolas verdes depois de retirarmos uma bola cinza}}{\text{Número total de bolas restantes}} = \dfrac{3}{4}$

Logo, chegamos a:

- $P(C / C) = \dfrac{P(C \cap C)}{P(C)} \Rightarrow P(C \cap C) = P(C) \cdot P(C / C) = \dfrac{2}{5} . \dfrac{1}{4} = \dfrac{1}{10}$

- $P(C / V) = \dfrac{P(C \cap V)}{P(V)} \Rightarrow P(C \cap V) = P(V) \cdot P(C / V) = \dfrac{3}{5} . \dfrac{1}{2} = \dfrac{3}{10}$

- $P(V / C) = \dfrac{P(V \cap C)}{P(C)} \Rightarrow P(V \cap C) = P(C) . P(V / C) = \dfrac{2}{5} . \dfrac{3}{4} = \dfrac{3}{10}$

- $P(V / V) = \dfrac{P(V \cap V)}{P(V)} \Rightarrow P(V \cap V) = P(V) . P(V / V) = \dfrac{3}{5} . \dfrac{1}{2} = \dfrac{3}{10}$

Vamos à árvore de probabilidades, que sintetiza os cálculos que fizemos e está representada na figura 14.

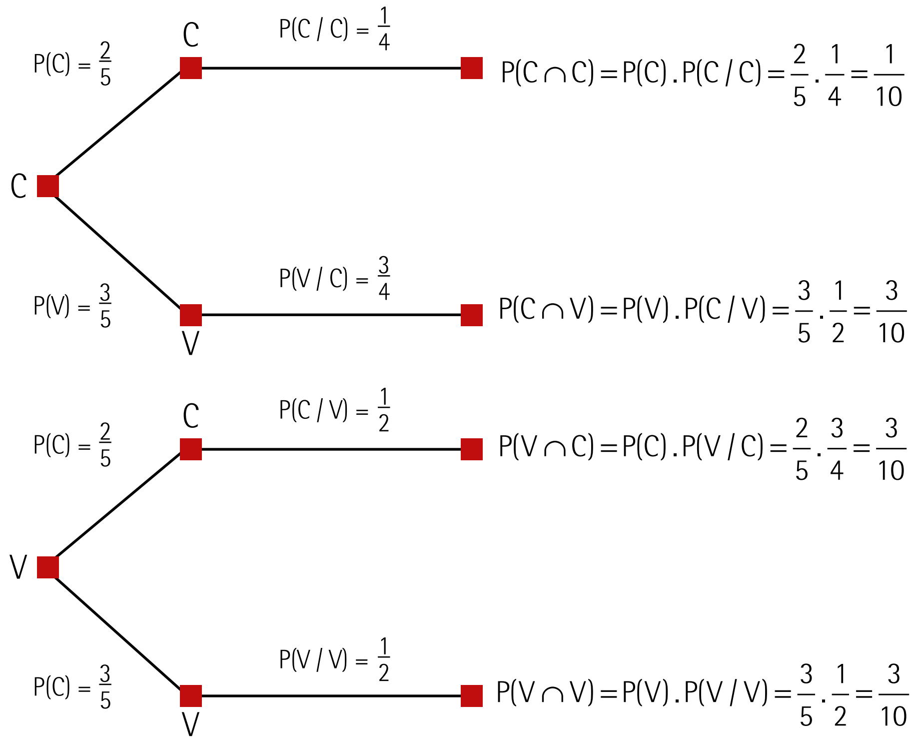



Figura 14 – Árvore de probabilidades para as retiradas sucessivas e sem reposição de 2 bolas de uma urna que contém 2 bolas cinza (C) e 3 bolas verdes (V)

Na tabela 26, temos um resumo dos resultados e de suas probabilidades para o exemplo em estudo.

**Tabela 26 – Resultados e probabilidades de ocorrência para retiradas de 2 bolas de uma urna com 5 bolas (2 bolas cinza e 3 bolas verdes) sem reposição**

| Resultado obtido | Probabilidade de ocorrer o resultado |
|---|---|
| Retiro uma bola cinza e, depois, outra bola cinza | $P(C \cap C) = P(C) . P(C / C) = \frac{2}{5} . \frac{1}{4} = \frac{1}{10}$ |
| Retiro uma bola cinza e, depois, uma bola verde | $P(C \cap V) = P(V) . P(C / V) = \frac{3}{5} . \frac{1}{2} = \frac{3}{10}$ |
| Retiro uma bola verde e, depois, uma bola cinza | $P(V \cap C) = P(C) . P(V / C) = \frac{2}{5} . \frac{3}{4} = \frac{3}{10}$ |
| Retiro uma bola verde e, depois, outra bola verde | $P(V \cap V) = P(V) . P(V / V) = \frac{3}{5} . \frac{1}{2} = \frac{3}{10}$ |
| Todo tipo possível de combinação de retiradas | $P(C \cap C) + P(C \cap V) + P(V \cap C) + P(V \cap V) = \frac{1}{10} + \frac{3}{10} + \frac{3}{10} + \frac{3}{10} = 1$ |

Vemos, na tabela 26, que a probabilidade de retirarmos duas bolas verdes no caso em que não há reposição da primeira bola retirada é igual a 3/10, ou seja, 30%.

Lembrete

Como vimos, a probabilidade de ocorrência de todo o espaço amostral é igual a 1 (100%). Por isso, obtivemos 1 na última linha e última coluna da tabela 26.

## 3.4 Eventos independentes

Vamos indicar por A e por B dois eventos do espaço amostral E, cujas probabilidades de ocorrência são, respectivamente, P(A) e P(B).

Os eventos A e B são independentes se informações a respeito da ocorrência ou da não ocorrência de A não interferem na probabilidade de ocorrência de B. Nesse caso, temos:

$P(A/B) = P(A)$ e $P(A \cap B) = P(A) \,.\, P(B)$

Vamos aplicar o conceito de eventos independentes para resolver a situação a seguir.

Pense novamente em uma urna com 2 bolas cinza e 3 bolas verdes, conforme esquematizado na figura 13.

Dessa urna, que contém o total de 5 bolas, retiramos ao acaso e sucessivamente 2 bolas, mas, agora, fazendo a reposição da bola inicialmente retirada (colocando-a novamente na urna).

Nesse caso, qual é a probabilidade de retirarmos duas bolas verdes?

Esse caso é bem mais simples do que o caso em que não havia a reposição da bola inicialmente retirada, pois trata-se de uma situação em que a primeira retirada e a segunda retirada são eventos independentes. Logo, ficamos com:

- $P(C \cap C) = P(C) \,.\, P(C) = \frac{2}{5} \,.\, \frac{2}{5} = \frac{4}{25}$
- $P(C \cap V) = P(C) \,.\, P(V) = \frac{2}{5} \,.\, \frac{3}{5} = \frac{6}{25}$
- $P(V \cap C) = P(V) \,.\, P(C) = \frac{3}{5} \,.\, \frac{2}{5} = \frac{6}{25}$
- $P(V \cap V) = P(V) \,.\, P(V) = \frac{3}{5} \,.\, \frac{3}{5} = \frac{9}{25}$

Na tabela 27, temos um resumo dos resultados e de suas probabilidades para o exemplo em estudo.

**Tabela 27 – Resultados e probabilidades de ocorrência para retiradas de 2 bolas de uma urna com 5 bolas (2 bolas cinza e 3 bolas verdes) com reposição**

| Resultado obtido | Probabilidade de ocorrer o resultado |
|---|---|
| Retiro uma bola cinza e, depois, outra bola cinza | $P(C \cap C) = P(C) \,.\, P(C) = \frac{2}{5} \,.\, \frac{2}{5} = \frac{4}{25}$ |
| Retiro uma bola cinza e, depois, uma bola verde | $P(C \cap V) = P(C) \,.\, P(V) = \frac{2}{5} \,.\, \frac{3}{5} = \frac{6}{25}$ |
| Retiro uma bola verde e, depois, uma bola cinza | $P(V \cap C) = P(V) \,.\, P(C) = \frac{3}{5} \,.\, \frac{2}{5} = \frac{6}{25}$ |
| Retiro uma bola verde e, depois, outra bola verde | $P(V \cap V) = P(V) \,.\, P(V) = \frac{3}{5} \,.\, \frac{3}{5} = \frac{9}{25}$ |
| Todo tipo possível de combinação de retiradas | $P(C \cap C) + P(C \cap V) + P(V \cap C) + P(V \cap V) = \frac{4}{25} + \frac{6}{25} + \frac{6}{25} + \frac{9}{25} = 1$ |

Vemos, na tabela 27, que a probabilidade de retirarmos duas bolas verdes no caso em que há reposição da primeira bola retirada é igual a 9/25, ou seja, 36%.

Saiba mais

Para saber mais sobre como calcular probabilidades, assista ao vídeo indicado.

PROBABILIDADE: como calcular? 2019. 1 vídeo (6 min). Publicado pelo canal Ferreto Matemática. Disponível em: https://bit.ly/3Duyouq. Acesso em: 17 nov. 2021.

## 4 PROBABILIDADE: PARTE 2

### 4.1 Distribuição discreta de probabilidade (função discreta de probabilidade)

Vamos pensar novamente em um dado cujas faces estejam numeradas de 1 a 6, como o dado mostrado na figura a seguir.

Figura 15 – Dado com faces numeradas de 1 a 6

Disponível em: https://bit.ly/3qRqgkc. Acesso em: 17 nov. 2021.

Imagine o seguinte experimento: lançamos o dado uma vez e observamos o valor indicado na face que ficará voltada para cima.

Esse experimento gera os eventos apresentados na tabela a seguir, em que também adicionamos suas probabilidades de ocorrência.

**Tabela 28 – Eventos gerados pela observação da face voltada para cima no lançamento de um dado e suas probabilidades**

| Evento | Número da face voltada para cima | Probabilidade |
|---|---|---|
| Evento 1 | 1 | 1/6 |
| Evento 2 | 2 | 1/6 |
| Evento 3 | 3 | 1/6 |
| Evento 4 | 4 | 1/6 |
| Evento 5 | 5 | 1/6 |
| Evento 6 | 6 | 1/6 |

Se chamarmos de X a variável aleatória que representa o número da face voltada para cima no lançamento de um dado, os valores assumidos por X são:

- $x_1 = 1$
- $x_2 = 2$
- $x_3 = 3$
- $x_4 = 4$

- $x_5 = 5$
- $x_6 = 6$

As probabilidades de termos cada um dos valores anteriores são:

- $p(x_1 = 1) = 1/6$
- $p(x_2 = 2) = 1/6$
- $p(x_3 = 3) = 1/6$
- $p(x_4 = 4) = 1/6$
- $p(x_5 = 5) = 1/6$
- $p(x_6 = 6) = 1/6$

De modo geral, segundo Magalhães e Lima (2008), a quantidade X associada a cada possível resultado de um evento amostral é chamada de **variável aleatória discreta** se assumir valores em um conjunto enumerável e com certa probabilidade.

A função discreta de probabilidade de uma variável aleatória X que assume valores $x_1, x_2, ..., x_i, ..., x_N$ é dada conforme segue, em que $p(x_i)$ representa a probabilidade de ocorrência do valor $x_i$ da variável X.

**Tabela 29**

| X | $x_1$ | $x_2$ | ... | $x_i$ | ... | $x_N$ |
|---|---|---|---|---|---|---|
| p | $p(x_1)$ | $p(x_2)$ | ... | $p(x_i)$ | ... | $p(x_N)$ |

Na função discreta de probabilidade, temos:

- $0 \leq p(x_i) \leq 1$
- $p(x_1) + p(x_2) + ... + p(x_i) + ... + p(x_N) = \sum_{i=1}^{N} p(x_i) = 1$

Observação

A indicação **aleatória** usada na nomeação de uma variável aleatória indica que não sabemos, com certeza, o seu resultado. O que temos é a probabilidade, ou a chance, de sua ocorrência.

Assim, no caso que estudamos, ficamos com a seguinte função discreta de probabilidade da variável aleatória X, que representa o número da face voltada para cima no lançamento de um dado.

**Tabela 30**

| X | $x_1 = 1$ | $x_2 = 2$ | $x_3 = 3$ | $x_4 = 4$ | $x_5 = 5$ | $x_6 = 6$ |
|---|---|---|---|---|---|---|
| p | $p(x_1) = 1/6$ | $p(x_2) = 1/6$ | $p(x_3) = 1/6$ | $p(x_4) = 1/6$ | $p(x_5) = 1/6$ | $p(x_6) = 1/6$ |

Podemos simplificar nossa representação:

**Tabela 31**

| X | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| p | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 |

Veja que a soma de todas as probabilidades dá 1.

Agora, imagine que, em determinada cidade ABC, tenhamos o seguinte.

- A probabilidade de uma família não ter carro é 0,23 (ou 23%).
- A probabilidade de uma família ter 1 carro é 0,32 (ou 32%).
- A probabilidade de uma família ter 2 carros é 0,21 (ou 21%).

Sabe-se que, na cidade ABC, nenhuma família tem mais de 3 carros. Para essa cidade, determine a probabilidade de uma família ter 3 carros e indique a função de probabilidade da variável aleatória que representa o número de carros que uma família tem.

Vamos chamar de X a variável aleatória discreta que representa o número de carros que uma família da cidade ABC tem. Logo, podemos ter:

- $x_1 = 0$
- $x_2 = 1$
- $x_3 = 2$
- $x_4 = 3$

As probabilidades de termos cada um dos valores anteriores são:

- $p(x_1 = 0) = 0{,}23$
- $p(x_2 = 1) = 0{,}32$
- $p(x_3 = 2) = 0{,}21$
- $p(x_4 = 3) = ?$

Sabemos que a soma de todas as probabilidades dá 1. Se aplicarmos isso à situação em estudo, podemos encontrar a probabilidade $p(x_4 = 3)$ de uma família da cidade ABC ter 3 carros. Vejamos:

$p(x_1 = 0) + p(x_2 = 1) + p(x_3 = 2) + p(x_4 = 3) = 1$

$0{,}23 + 0{,}32 + 0{,}21 + p(x_4 = 3) = 1$

$0{,}76 + p(x_4 = 3) = 1$

$p(x_4 = 3) = 1 - 0{,}76$

$p(x_4 = 3) = 0{,}24$

Logo, ficamos com a seguinte função discreta de probabilidade da variável aleatória X, que representa as quantidades de carros que as famílias da cidade ABC têm.

**Tabela 32**

| X | $x_1 = 0$ | $x_2 = 1$ | $x_3 = 2$ | $x_4 = 3$ |
|---|---|---|---|---|
| p | $p(x_1) = 0{,}23$ | $p(x_2) = 0{,}32$ | $p(x_3) = 0{,}21$ | $p(x_4) = 0{,}24$ |

Podemos simplificar nossa representação:

**Tabela 33**

| X | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| p | 0,23 | 0,32 | 0,21 | 0,24 |

A esperança E(X) de uma variável aleatória discreta é calculada por:

$$E(X) = x_1 \,.\, p(x_1) + x_2 \,.\, p(x_2) + \ldots + x_i \,.\, p(x_i) + \ldots + x_N \,.\, p(x_N) = \sum_{i=1}^{N} x_i \,.\, p(x_i)$$

No caso do exemplo da cidade ABC, a esperança é igual a 1,46, pois:

$E(X) = x_1 . p(x_1) + x_2 . p(x_2) + x_3 . p(x_3) + x_4 . p(x_4)$

$E(X) = 0 . 0{,}23 + 1 . 0{,}32 + 2 . 0{,}21 + 3 . 0{,}24 = 0 + 0{,}32 + 0{,}42 + 0{,}72$

$E(X) = 1{,}46$

**Observação**

A esperança E(X) de uma variável aleatória discreta também pode ser representada por $\mu$.

A variância VAR(X) de uma variável aleatória discreta é calculada por:

$VAR(X) = E(X^2) - [E(X)]^2$

Na igualdade, temos:

$$E(X^2) = x_1^2 . p(x_1) + x_2^2 . p(x_2) + ... + x_i^2 . p(x_i) + ... + x_N^2 . p(x_N) = \sum_{i=1}^{N} x_i^2 . p(x_i)$$

No caso do exemplo da cidade ABC, a variância é igual a 1,1884, como segue.

$$E(X^2) = x_1^2 . p(x_1) + x_2^2 . p(x_2) + x_3^2 . p(x_3) + x_4^2 . p(x_4)$$

$E(X^2) = 0^2 . 0{,}23 + 1^2 . 0{,}32 + 2^2 . 0{,}21 . 3^2 . 0{,}24 = 0 + 0{,}32 + 0{,}84 + 2{,}16$

$E(X^2) = 3{,}32$

$VAR(X) = E(X^2) - [E(X)]^2 = 3{,}32 - (1{,}46)^2 = 3{,}32 - 2{,}1316$

$VAR(X) = 1{,}1884$

**Observação**

A variância VAR(X) de uma variável aleatória discreta também pode ser representada por $\sigma^2$.

O desvio padrão DP(X) de uma variável aleatória discreta é calculado por:

$$DP(X) = \sqrt{VAR(X)}$$

No caso do exemplo da cidade ABC, o desvio padrão é igual a 1,09, conforme calculado a seguir.

$$DP(X) = \sqrt{1{,}1884} = 1{,}09$$

Observação

O desvio padrão DP(X) de uma variável aleatória discreta também pode ser representado por $\sigma$.

## 4.2 Principais modelos de distribuição discreta de probabilidade

A seguir, veremos os principais modelos de distribuição discreta de probabilidade, a saber:

- modelo uniforme discreto;
- modelo de Bernoulli;
- modelo binomial.

Observação

Há outros modelos de função discreta de probabilidade, como o geométrico, o hipergeométrico e o de Poisson.

### 4.2.1 Modelo uniforme discreto

No modelo uniforme discreto, todos os possíveis valores da variável aleatória discreta têm a mesma probabilidade P de ocorrência.

Voltemos ao caso de lançamento de um dado, em que a função discreta de probabilidade da variável aleatória X que representa o número da face voltada para cima no lançamento de um dado é dada, como vimos, nas representações a seguir.

**Tabela 34**

| X | $x_1 = 1$ | $x_2 = 2$ | $x_3 = 3$ | $x_4 = 4$ | $x_5 = 5$ | $x_6 = 6$ |
|---|---|---|---|---|---|---|
| p | $p(x_1) = 1/6$ | $p(x_2) = 1/6$ | $p(x_3) = 1/6$ | $p(x_4) = 1/6$ | $p(x_5) = 1/6$ | $p(x_6) = 1/6$ |

**Tabela 35**

| X | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| p | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 | 1/6 |

Veja que cada um dos valores $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$, $x_5 = 5$ e $x_6 = 6$ tem a mesma probabilidade $p = 1/6$ de ocorrer. Logo, trata-se de um modelo uniforme discreto de probabilidade. Na figura a seguir, temos um gráfico que mostra essa distribuição.

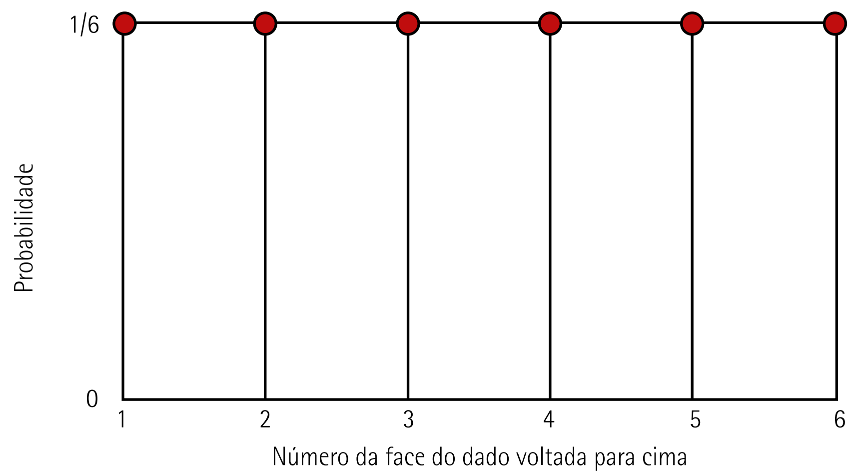



Figura 16 – Gráfico de um modelo uniforme discreto de probabilidade

De modo geral, dizemos que, no modelo uniforme discreto, todos os N possíveis valores da variável aleatória discreta X, representados por $x_1$, $x_2$, ..., $x_i$, ..., $x_N$, têm a mesma probabilidade P de ocorrência, ou seja:

$$p(x_1) = p(x_2) = \ldots = p(x_i) = \ldots = p(x_N) = p$$

Isso pode ser representado por:

$$p(X = x_i) = \frac{1}{N}$$

A esperança E(X) (ou $\mu$), a variância VAR(X) (ou $\sigma^2$) e o desvio padrão DP(X) (ou $\sigma$) de uma variável aleatória discreta X que segue o modelo uniforme discreto de probabilidades são calculados por:

$$\mu = E(X) = \frac{1}{N}(x_1 + x_2 + \ldots + x_i + \ldots + x_N) = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\sigma^2 = VAR(X) = \frac{1}{N}\left(\sum_{i=1}^{N} x_i^2 - \frac{\left(\sum_{i=1}^{N} x_i\right)^2}{N}\right)$$

$$\sigma = DP(X) = \sqrt{VAR(X)}$$

Logo, no caso de lançamento de um dado em estudo, temos o que segue.

$$\mu = E(X) = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{1}{6}\sum_{i=1}^{6} x_i = \frac{1}{6}.(x_1 + x_2 + x_3 + x_4 + x_5 + x_6)$$

$$\mu = E(X) = \frac{1}{6}.(1+2+3+4+5+6) = \frac{1}{6}.21 = 3{,}5$$

$$\sigma^2 = VAR(X) = \frac{1}{N}\left(\sum_{i=1}^{N} x_i^2 - \frac{\left(\sum_{i=1}^{N} x_i\right)^2}{N}\right) = \frac{1}{6}\left(\sum_{i=1}^{6} x_i^2 - \frac{\left(\sum_{i=1}^{6} x_i\right)^2}{6}\right)$$

$$\sigma^2 = VAR(X) = \frac{1}{6}\left(\left(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2\right) - \frac{(x_1 + x_2 + x_3 + x_4 + x_5 + x_6)^2}{6}\right)$$

$$\sigma^2 = VAR(X) = \frac{1}{6}\left(\left(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2\right) - \frac{(1+2+3+4+5+6)^2}{6}\right)$$

$$\sigma^2 = VAR(X) = \frac{1}{6}\left((1+4+9+16+25+36) - \frac{(21)^2}{6}\right) = \frac{1}{6}\left(91 - \frac{441}{6}\right) = 2{,}917$$

$$\sigma = DP(X) = \sqrt{VAR(X)} = \sqrt{2{,}917} = 1{,}708$$

### 4.2.2 Modelo de Bernoulli

Usamos o modelo de Bernoulli em um experimento aleatório que admite apenas um dos dois resultados a seguir.

- Sucesso, em que a variável aleatória discreta assume o valor 1.
- Fracasso, em que a variável aleatória discreta assume o valor 0.

Se a probabilidade de sucesso for p, a probabilidade de fracasso será 1 – p, visto que há apenas duas situações (sucesso ou fracasso) e a soma das probabilidades de todos os resultados possíveis dá 1.

Imagine que X seja uma variável aleatória discreta que segue o modelo de Bernoulli de probabilidade P de sucesso (ou parâmetro p). Nesse caso, indicamos "X ~ Bernoulli(p)" e escrevemos:

$$X = \begin{cases} 1, \text{se ocorrer sucesso} \\ 0, \text{se ocorrer fracasso} \end{cases}$$

Logo, a função de probabilidade de uma distribuição "X ~ Bernoulli(p)" é a que segue.

**Tabela 36**

| X | 1 | 0 |
|---|---|---|
| P(X = x) | p | 1 – p |

A esperança E(X) (ou $\mu$), a variância VAR(X) (ou $\sigma^2$) e o desvio padrão DP(X) (ou $\sigma$) de uma variável aleatória discreta X que segue o modelo de Bernoulli de parâmetro p são calculados por:

$\mu = E(X) = p$

$\sigma^2 = VAR(X) = p \,.\, (1 - p)$

$\sigma = DP(X) = \sqrt{VAR(X)} = \sqrt{p \,.\, (1 - p)}$

Vejamos uma aplicação do modelo de Bernoulli. Imagine que você esteja resolvendo uma questão objetiva (tipo teste) com 5 alternativas e que você chute sua resposta. Nesse caso, ocorre uma das duas situações a seguir.

- Você acerta a questão, o que representa sucesso.
- Você erra a questão, o que representa fracasso.

Caso haja apenas 1 alternativa correta, a probabilidade P de você acertar a resposta nesse chute é de 1 em 5, ou 1/5 = 0,2 (pois há apenas 1 alternativa correta entre as 5 alternativas possíveis).

Assim, se X é a variável aleatória discreta associada ao caso em estudo, em que X = 1 representa sucesso e X = 0 representa fracasso, temos um modelo de Bernoulli de parâmetro p = 0,2 e com a função de probabilidade a seguir.

**Tabela 37**

| X | 1 | 0 |
|---|---|---|
| P(X = x) | 0,2 | 0,8 |

Veja que, se p = 0,2, então 1 – p = 1 – 0,2 = 0,8.

Nessa aplicação do modelo de Bernoulli, temos:

E(X) = p = 0,2 e VAR(X) = 0,2 . (1 - 0,2) = 0,2 . 0,8 = 0,16

O fato de a esperança ser 0,2 no exemplo em estudo indica que se espera que, ao chutar a resposta da questão com 5 alternativas, você tenha 20% de probabilidade de acertar a questão.

### 4.2.3 Modelo binomial

A variável aleatória discreta X que corresponde ao número de sucessos obtidos na realização de n ensaios de Bernoulli independentes, todos com a mesma probabilidade P de sucesso, segue modelo binomial de parâmetros n e p e é indicada por "X ~ b(n;p)".

Imagine que, nesses n ensaios, consideremos a probabilidade de ocorrência de k sucessos (e, evidentemente, n – k fracassos). Essa probabilidade, indicada por P(X = k), é calculada por:

$$P(X=k)=\binom{n}{k}p^{k}(1-p)^{n-k}$$

Na expressão, temos:

$$\binom{n}{k}=\frac{n!}{k!(n-k)!}$$

Lembrete

O símbolo n! é lido como "ene fatorial" e é dado por:

$n! = n \, . \, (n - 1) \, . \, (n - 2) \ldots 1$, com n inteiro e $n > 0$

Veja os exemplos a seguir.

- $1! = 1$
- $2! = 2 \, . \, 1 = 2$
- $3! = 3 \, . \, 2 \, .1 = 6$
- $4! = 4 \, . \, 3 \, . \, 2 \, . \, 1 = 24$

Observe que podemos reescrever as igualdades dos exemplos.

- $2! = 2 \, . \, 1!$
- $3! = 3 \, . \, 2!$
- $4! = 4 \, . \, 3!$

A esperança E(X) (ou $\mu$), a variância VAR(X) (ou $\sigma^2$) e o desvio padrão DP(X) (ou $\sigma$) de uma variável aleatória discreta X que segue o modelo binomial de parâmetros n e p são calculados por:

$$\mu = E(X) = n \, . \, p$$

$$\sigma^2 = VAR(X) = n \, . \, p \, . \, (1 - p)$$

$$\sigma = DP(X) = \sqrt{VAR(X)} = \sqrt{n \, . \, p \, . \, (1 - p)}$$

Vejamos uma aplicação do modelo binomial. Imagine que você esteja resolvendo 10 testes com 5 alternativas cada e que você chute as alternativas corretas. Qual é a probabilidade de você não acertar questão alguma? Quais são as probabilidades de você acertar 1 questão, 2 questões, 3 questões, 4 questões, 5 questões, 6 questões, 7 questões, 8 questões, 9 questões e 10 questões?

Nesse caso, em cada chute, ocorre uma das duas situações a seguir.

- Você acerta a questão, o que representa sucesso.
- Você erra a questão, o que representa fracasso.

A probabilidade p de você acertar a resposta em cada chute é de 1 em 5, ou 1/5 = 0,2, como já vimos. Aqui, o número n de tentativas é o número total de questões, ou seja, n = 10.

Vamos chamar de X a variável aleatória discreta associada ao caso em estudo. Logo, temos um modelo binomial de parâmetros n = 10 e p = 0,2, cuja função probabilidade é dada pela seguinte expressão:

$$P(X=k)=\binom{n}{k}p^{k}(1-p)^{n-k}=\binom{10}{k}0,2^{k}(1-0,2)^{10-k}$$

$$P(X=k)=\binom{10}{k}.0,2^{k}.0,8^{10-k}$$

Na expressão, k representa o número de acertos. Assim:

- P(X = 0) é a probabilidade de você não acertar questão alguma, em que usamos k = 0;
- P(X = 1) é a probabilidade de você acertar 1 questão, em que usamos k = 1;
- P(X = 2) é a probabilidade de você acertar 2 questões, em que usamos k = 2;
- P(X = 3) é a probabilidade de você acertar 3 questões, em que usamos k = 3;
- P(X = 4) é a probabilidade de você acertar 4 questões, em que usamos k = 4;
- P(X = 5) é a probabilidade de você acertar 5 questões, em que usamos k = 5;
- P(X = 6) é a probabilidade de você acertar 6 questões, em que usamos k = 6;
- P(X = 7) é a probabilidade de você acertar 7 questões, em que usamos k = 7;
- P(X = 8) é a probabilidade de você acertar 8 questões, em que usamos k = 8;
- P(X = 9) é a probabilidade de você acertar 9 questões, em que usamos k = 9;
- P(X = 10) é a probabilidade de você acertar 10 questões, em que usamos k = 10.

Vamos calcular cada uma dessas probabilidades.

$$P(X=0)=\binom{10}{0}.0,2^{0}.0,8^{10-0}=\frac{10!}{0!10!}.0,2^{0}.0,8^{10}=\frac{1}{1}.0,2^{0}.0,8^{10}=0,1073742$$

$$P(X=1)=\binom{10}{1}.\,0{,}2^{1}\,.\,0{,}8^{10-1}=\frac{10!}{1!9!}.\,0{,}2^{1}\,.\,0{,}8^{9}=\frac{10\cdot 9!}{1!9!}.\,0{,}2^{1}\,.\,0{,}8^{9}=10\,.\,0{,}2^{1}\,.\,0{,}8^{9}=0{,}2684355$$

$$P(X=2)=\binom{10}{2}.\,0{,}2^{2}\,.\,0{,}8^{10-2}=\frac{10!}{2!8!}.\,0{,}2^{2}\,.\,0{,}8^{8}=\frac{10\,.\,9\,.\,8!}{2!8!}.\,0{,}2^{2}\,.\,0{,}8^{8}=45\,.\,0{,}2^{2}\,.\,0{,}8^{8}=0{,}3019899$$

$$P(X=3)=\binom{10}{3}.\,0{,}2^{3}\,.\,0{,}8^{10-3}=\frac{10!}{3!7!}.\,0{,}2^{3}\,.\,0{,}8^{7}=\frac{10\,.\,9\,.\,8\,.\,7!}{6\cdot 7!}.\,0{,}2^{3}\,.\,0{,}8^{7}=120\,.\,0{,}2^{3}\,.\,0{,}8^{7}=0{,}2013266$$

$$P(X=4)=\binom{10}{4}.\,0{,}2^{4}\,.\,0{,}8^{10-4}=\frac{10!}{4!6!}.\,0{,}2^{4}\,.\,0{,}8^{6}=\frac{10\,.\,9\,.\,8\,.\,7\,.\,6!}{24\cdot 6!}.\,0{,}2^{4}\,.\,0{,}8^{6}=210\,.\,0{,}2^{4}\,.\,0{,}8^{6}=0{,}0880804$$

$$P(X=5)=\binom{10}{5}.\,0{,}2^{5}\,.\,0{,}8^{10-5}=\frac{10!}{5!5!}.\,0{,}2^{5}\,.\,0{,}8^{5}=\frac{10\,.\,9\,.\,8\,.\,7\,.\,6\,.\,5!}{120\,.\,5!}.\,0{,}2^{5}\,.\,0{,}8^{5}=252\,.\,0{,}2^{5}\,.\,0{,}8^{5}=0{,}0264241$$

$$P(X=6)=\binom{10}{6}.\,0{,}2^{6}\,.\,0{,}8^{10-6}=\frac{10!}{6!4!}.\,0{,}2^{6}\,.\,0{,}8^{4}=\frac{10\,.\,9\,.\,8\,.\,7\,.\,6!}{24\cdot 6!}.\,0{,}2^{6}\,.\,0{,}8^{4}=210\,.\,0{,}2^{6}\,.\,0{,}8^{4}=0{,}0055050$$

$$P(X=7)=\binom{10}{7}.\,0{,}2^{7}\,.\,0{,}8^{10-7}=\frac{10!}{7!3!}.\,0{,}2^{7}\,.\,0{,}8^{3}=\frac{10\,.\,9\,.\,8\,.\,7!}{6\cdot 7!}.\,0{,}2^{7}\,.\,0{,}8^{3}=120\,.\,0{,}2^{7}\,.\,0{,}8^{3}=0{,}0007864$$

$$P(X=8)=\binom{10}{8}.\,0{,}2^{8}\,.\,0{,}8^{10-8}=\frac{10!}{8!2!}.\,0{,}2^{8}\,.\,0{,}8^{2}=\frac{10\,.\,9\,.\,8!}{2\cdot 8!}.\,0{,}2^{8}\,.\,0{,}8^{2}=45\,.\,0{,}2^{8}\,.\,0{,}8^{2}=0{,}0000737$$

$$P(X=9)=\binom{10}{9}.\,0{,}2^{9}\,.\,0{,}8^{10-1}=\frac{10!}{9!1!}.\,0{,}2^{9}\,.\,0{,}8^{1}=\frac{10\,.\,9!}{1.\,9!}.\,0{,}2^{9}\,.\,0{,}8^{1}=10\,.\,0{,}2^{9}\,.\,0{,}8^{1}=0{,}0000041$$

$$P(X=10)=\binom{10}{10}.\,0{,}2^{10}\,.\,0{,}8^{10-10}=\frac{10!}{10!0!}.\,0{,}2^{10}\,.\,0{,}8^{0}=1.\,0{,}2^{10}\,.\,0{,}8^{0}=0{,}0000001$$

**Lembrete**

Por definição, 0! é igual a 1.

Na tabela e no gráfico da figura a seguir, temos um resumo do que calculamos.

**Tabela 38 – Probabilidades de acertos (no chute)**

| Quantidade de testes corretos (no chute) | Probabilidade | Probabilidade em % |
|---|---|---|
| 0 | 0,1073742 | 10,73742% |
| 1 | 0,2684355 | 26,84355% |
| 2 | 0,3019899 | 30,19899% |
| 3 | 0,2013266 | 20,13266% |
| 4 | 0,0880804 | 8,80804% |
| 5 | 0,0264241 | 2,64241% |
| 6 | 0,0055050 | 0,55050% |
| 7 | 0,0007864 | 0,07864% |
| 8 | 0,0000737 | 0,00737% |
| 9 | 0,0000041 | 0,00041% |
| 10 | 0,0000001 | 0,00001% |
| Soma | 1,0000000 | 100% |

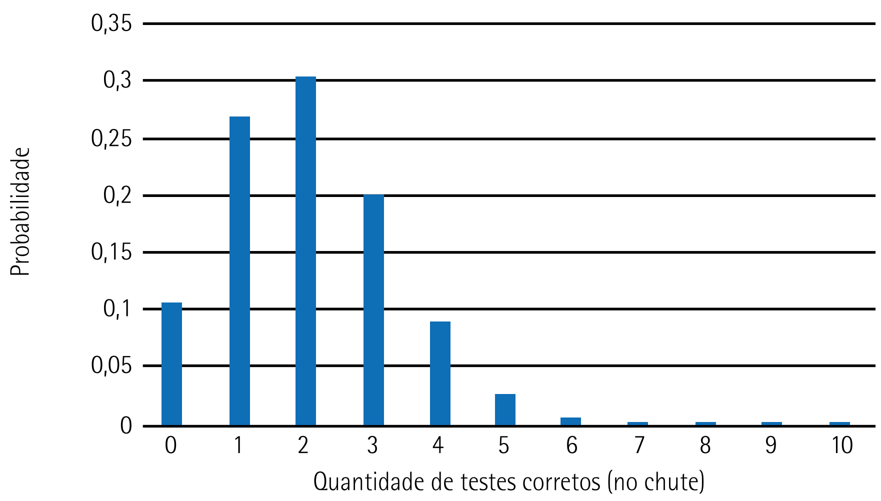



Figura 17 – Gráfico de um modelo binomial de probabilidade

Veja, no gráfico, que, no chute, a probabilidade de você acertar 9 dos 10 testes é de 0,00041%.

Para o exemplo que estamos estudando, qual é probabilidade de você acertar no máximo 3 questões no chute?

Essa probabilidade, indicada por $P(X \leq 3)$, é a soma das probabilidades de você não acertar questão alguma, acertar 1 questão, acertar 2 questões e acertar 3 questões. Logo:

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$$

$$P(X \leq 3) = 0{,}1073742 + 0{,}2684355 + 0{,}3019899 + 0{,}2013266 = 0{,}879126$$

Logo, em uma prova composta por 10 testes (cada um deles com 5 alternativas), você tem quase 88% de chance de, apenas chutando, acertar no máximo 3 testes. Isso equivale a dizer que, nessa prova, você tem cerca de 12% de chance de acertar mais de 3 testes exclusivamente com chutes.

A esperança E(X) (ou $\mu$), a variância VAR(X) (ou $\sigma^2$) e o desvio padrão DP(X) (ou $\sigma$) para o exemplo em estudo estão calculados a seguir.

$$\mu = E(X) = n \,.\, p = 10 \,.\, 0{,}2 = 2$$

$$\sigma^2 = VAR(X) = n \,.\, p \,.\, (1 - p) = 10 \,.\, 0{,}2 \,.\, (1 - 0{,}2) = 2 \,.\, 0{,}8 = 1{,}6$$

$$\sigma = DP(X) = \sqrt{VAR(X)} = \sqrt{1{,}6} = 1{,}265$$

O fato de a esperança ser 2 no exemplo em estudo indica que se espera que, ao resolver as 10 questões com 5 alternativas cada utilizando exclusivamente chutes, você tenha 2 acertos. Obviamente, você pode acertar mais questões ou menos questões. Uma forma de interpretar essa esperança é que, se um número muito grande de alunos realizasse a mesma prova de 10 questões com 5 alternativas cada e se todos eles respondessem às questões com base em chutes, a média de acertos dos alunos seria de 2 acertos.

### 4.3 Distribuição contínua de probabilidade (função densidade de probabilidade)

As distribuições contínuas de probabilidade operam com variáveis aleatórias contínuas, ou seja, com variáveis cujos valores ocorrem de modo aleatório e cujo domínio está em intervalo de números reais.

Em outras palavras, diferentemente do que acontece com as variáveis aleatórias discretas, as variáveis aleatórias contínuas podem assumir "infinitos" valores, ou seja, não conseguimos listar individualmente todos os seus possíveis valores. Logo, no caso de uma distribuição contínua de probabilidade, falamos de probabilidades associadas a intervalos de valores da variável X, e não da probabilidade de um valor "específico" x da variável.

Nessa situação, representamos graficamente o que chamamos de função densidade de probabilidade, indicada por f.d.p. e representada por f(x), como mostrado de modo genérico na figura a seguir.

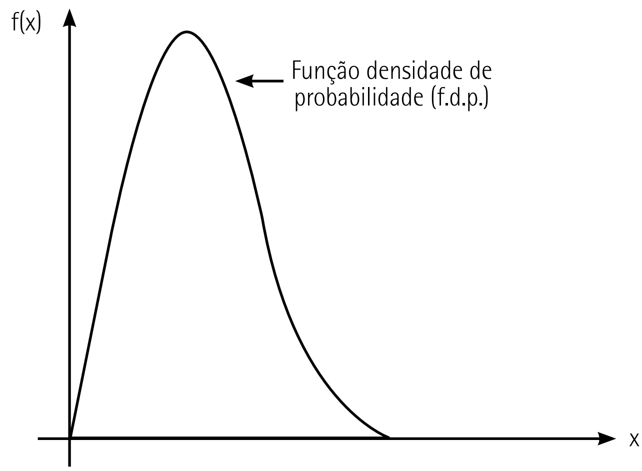



Figura 18 – Gráfico genérico de função densidade de probabilidade, indicada por f.d.p. e representada por f(x)

A função densidade de probabilidade f(x) que caracteriza uma variável aleatória contínua X apresenta as propriedades I e II a seguir.

I – Qualquer que seja o valor de x, $f(x) \geq 0$. Ou seja, o gráfico de f(x) "nunca" ocupa região abaixo do eixo x, pois não existem probabilidades negativas.

II – Numericamente, a área delimitada pelo gráfico de f(x) e pelo eixo x é igual a 1, como ilustrado em laranja na figura a seguir. Essa área representa a probabilidade de ocorrência de todos os possíveis valores de x, o que resulta em 100%.

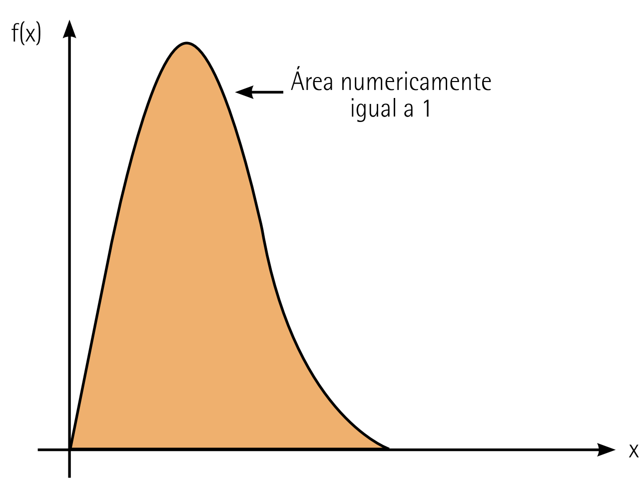



Figura 19 – Numericamente, a área em laranja delimitada pelo gráfico de f(x) e pelo eixo x é igual a 1, em que f(x) é uma f.d.p.

A probabilidade de termos $a \leq x < b$, representada por $P(a \leq x < b)$, é numericamente igual à área delimitada pelo gráfico de f(x), pelo eixo x e pelas retas verticais $x = a$ e $x = b$, como ilustrado na figura a seguir.

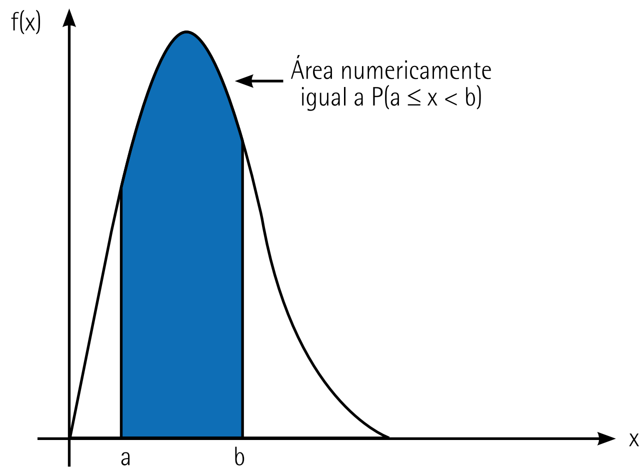



Figura 20 – A probabilidade de termos $a \leq x < b$, representada por $P(a \leq x < b)$, é a área em azul delimitada pelo gráfico de f(x), pelo eixo x e pelas retas verticais $x = a$ e $x = b$

Conforme dissemos, em uma f.d.p., a probabilidade associada a um valor "específico" (pontual) da variável é zero. Assim, temos o que segue.

- $P(x = a) = 0$ e $P(x = b) = 0$
- $P(a \leq x < b) = P(a < x < b) = P(a \leq x \leq b) = P(a < x \leq b)$

## 4.4 Principais modelos de distribuição contínua de probabilidade

A seguir, veremos os principais modelos de distribuição contínua de probabilidade, a saber:

- modelo uniforme contínuo;
- modelo normal.

**Observação**

Há outros modelos de distribuição contínua de probabilidade, como o exponencial.

### 4.4.1 Modelo uniforme contínuo

No modelo uniforme contínuo, a função densidade de probabilidade f(x) da variável aleatória contínua X no intervalo [a;b], ou seja, para $a \leq x \leq b$, é dada por:

$$f(x) = \frac{1}{b - a}$$

Para esse tipo de distribuição, se $x < a$ ou se $x > b$, $f(x) = 0$.

A esperança E(X) (ou $\mu$), a variância VAR(X) (ou $\sigma^2$) e o desvio padrão DP(X) (ou $\sigma$) de uma variável aleatória contínua X que segue o modelo uniforme contínuo são calculados por:

$$\mu = E(X) = \frac{a+b}{2}$$

$$\sigma^2 = VAR(X) = \frac{(a-b)^2}{12}$$

$$\sigma = DP(X) = \sqrt{VAR(X)}$$

Observação

A letra grega $\mu$, que indica a média populacional, é pronunciada como "mi" ou "mu". A letra grega $\sigma$, que indica o desvio padrão populacional, é pronunciada como "sigma". Assim, é comum pronunciarmos o símbolo da variância populacional ($\sigma^2$) como "sigma ao quadrado".

Vamos estudar um exemplo em que se usa o modelo uniforme contínuo de probabilidades. Pense na reta real (reta "contínua" em que temos números reais). Imagine que um ponto seja escolhido aleatoriamente no segmento [3;8] dessa reta. Qual é a probabilidade de que o ponto escolhido se encontre entre 5,2 e 7,1?

Primeiramente, precisamos conhecer a função densidade de probabilidade dessa variável. Chamamos de X a variável aleatória contínua que representa o ponto escolhido. A função densidade de probabilidade f(x) da variável X segue modelo uniforme contínuo no intervalo $[a;b] = [3;8]$, com $a = 3$ e $b = 8$, e é dada por:

$$f(x) = \frac{1}{b-a} = \frac{1}{8-3} = \frac{1}{5}$$

$F(x) = 0{,}2$

Para esse tipo de distribuição, se $x < 3$ ou se $x > 8$, $f(x) = 0$.

Graficamente, temos o que se mostra na figura a seguir.

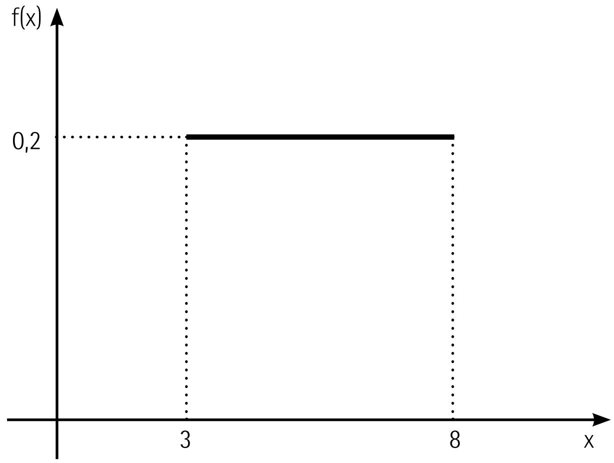



Figura 21 – Função densidade de probabilidade dada por f(x) = 0,2 para $3 \leq x \leq 8$ e f(x) = 0 para x < 3 ou para x > 8

Queremos determinar a probabilidade de que o ponto escolhido se encontre entre 5,2 e 7,1, indicada por $P(5{,}2 \leq X \leq 7{,}1)$. Numericamente, essa probabilidade é igual à área do retângulo mostrado na figura a seguir.

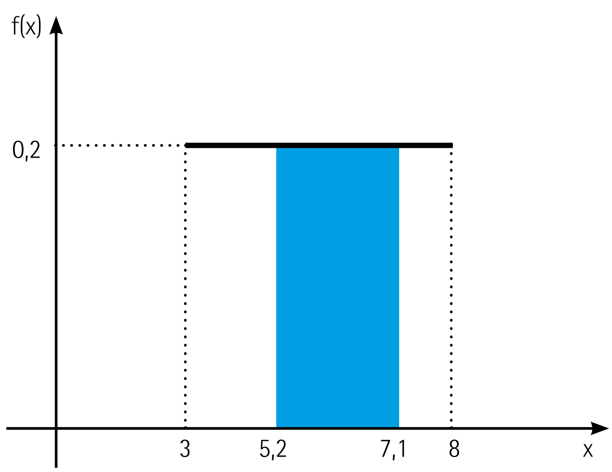



Figura 22 – A área do retângulo é numericamente igual a $P(5{,}2 \leq X \leq 7{,}1)$

A base B do retângulo mede 7,1 – 5,2 = 1,9 e a sua altura H mede 0,2. Logo, a área A do retângulo é 0,38, visto que A = B . H = 1,9 . 0,2 = 0,38.

Concluímos que a probabilidade de que o ponto escolhido se encontre entre 5,2 e 7,1 é igual a 0,38 (ou 38%).

A esperança E(X) (ou $\mu$), a variância VAR(X) (ou $\sigma^2$) e o desvio padrão DP(X) (ou $\sigma$) para o exemplo em estudo são calculados a seguir.

$$\mu = E(X) = \frac{a+b}{2} = \frac{3+8}{2} = \frac{11}{2} = 5{,}5$$

$$\sigma^2 = VAR(X) = \frac{(a-b)^2}{12} = \frac{(3-8)^2}{12} = \frac{(-5)^2}{12} = \frac{25}{12} = 2{,}083$$

$$\sigma = DP(X) = \sqrt{VAR(X)} = \sqrt{2{,}083} = 1{,}443$$

### 4.4.2 Modelo normal

No modelo normal, um dos tipos mais importantes de distribuição contínua de probabilidade, a função densidade de probabilidade f(x) da variável aleatória contínua X é dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Na igualdade:

- $\mu$ é a esperança ou média E(X) da variável aleatória contínua X;
- $\sigma$ é o desvio padrão DP(X) da variável aleatória contínua X.

Observação

A letra **e** da função em estudo representa o número de Neper, que é um número irracional. Ele é definido como o valor que a função $\left(1+\frac{1}{x}\right)^x$ tende a assumir quando a variável x recebe valores cada vez mais altos.

O valor de e é 2,71828182845904523536028 74...

Dizemos que a distribuição normal tem parâmetros $\mu$ e $\sigma^2$, em que $\sigma^2$ é a variância VAR(X) da variável aleatória contínua X, e fazemos sua indicação por $X \sim N(\mu;\sigma^2)$.

A distribuição normal de probabilidade também é chamada de gaussiana, em homenagem a Carl Friedrich Gauss, conhecido como O Príncipe da Matemática.

## Saiba mais

Para saber mais sobre Carl Friedrich Gauss, leia o seguinte texto.

KARL Friedrich Gauss. *Biblioteca Matemática*, [s.d.]. Disponível em: https://bit.ly/30FHFRK. Acesso em: 17 nov. 2021.

Na figura 23, temos um gráfico que mostra uma distribuição normal $X \sim N(\mu;\sigma^2)$. Veja que se trata de uma curva em forma de sino, simétrica em relação à média $\mu$, cujo ponto de máximo ocorre em $x = \mu$.

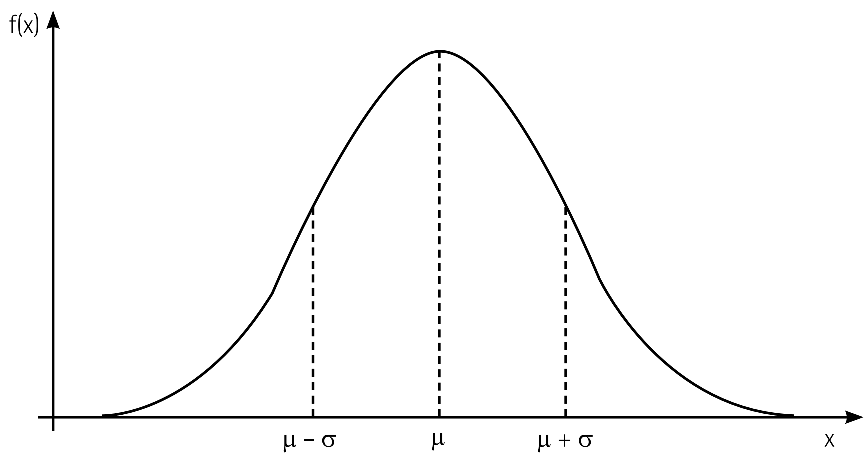



Figura 23 – Distribuição normal $X \sim N(\mu;\sigma^2)$

Na próxima figura, temos dois gráficos que mostram duas distribuições normais, $N_1(\mu_1;\sigma^2)$ e $N_2(\mu_2;\sigma^2)$, que apresentam o mesmo valor de desvio padrão ($\sigma$) e diferentes valores de médias ($\mu_1$ e $\mu_2$, com $\mu_1 < \mu_2$). Tais gráficos têm o mesmo formato, mas posicionam-se de modos distintos segundo o eixo horizontal.

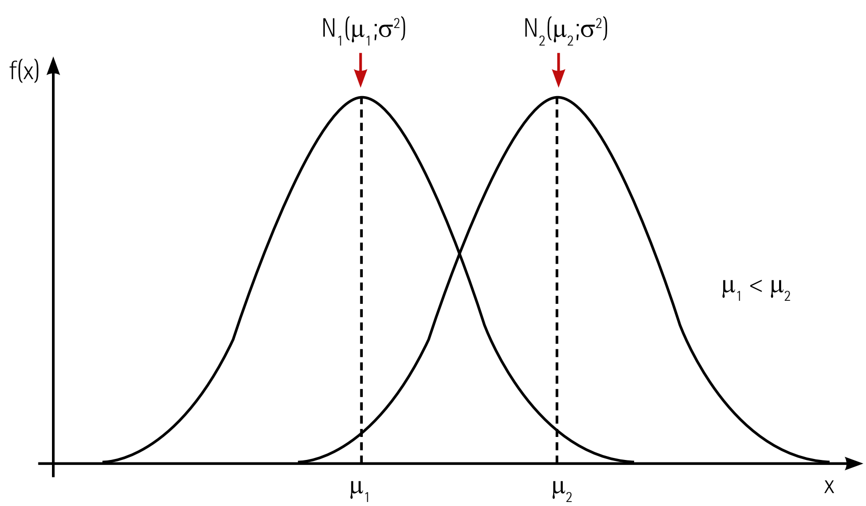



Figura 24 – Duas distribuições normais, $N_1(\mu_1;\sigma^2)$ e $N_2(\mu_2;\sigma^2)$, que apresentam o mesmo valor de desvio padrão ($\sigma$) e diferentes valores de médias ($\mu_1$ e $\mu_2$)

Na figura a seguir, temos dois gráficos que mostram duas distribuições normais, $N_3(\mu;\sigma_3^2)$ e $N_4(\mu;\sigma_4^2)$, que apresentam o mesmo valor de média ($\mu$) e diferentes valores de desvios padrão $\sigma_3$ e $\sigma_4$, com $\sigma_3 < \sigma_4$. Tais gráficos têm formatos diferentes. Quanto maior é a variância de uma distribuição normal, mais "achatado" é o seu gráfico.

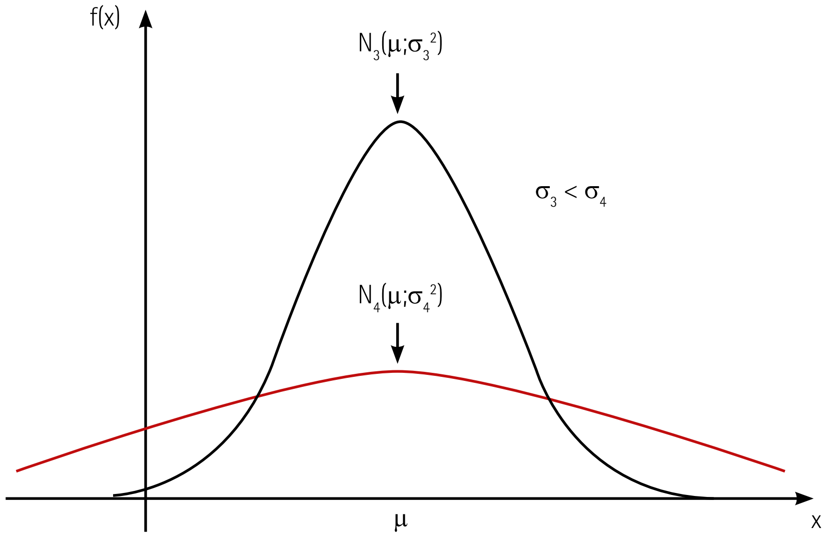



Figura 25 – Duas distribuições normais, $N_3(\mu;\sigma_3^2)$ e $N_4(\mu;\sigma_4^2)$, que apresentam o mesmo valor de média ($\mu$) e diferentes valores de desvios padrão ($\sigma_3$ e $\sigma_4$)

A probabilidade de termos $a \leq x < b$ no caso de uma variável aleatória contínua X que segue modelo normal, representada por $P(a \leq x < b)$, é numericamente igual à área delimitada pelo gráfico de f(x), pelo eixo x e pelas retas verticais $x = a$ e $x = b$, como ilustrado na figura a seguir.

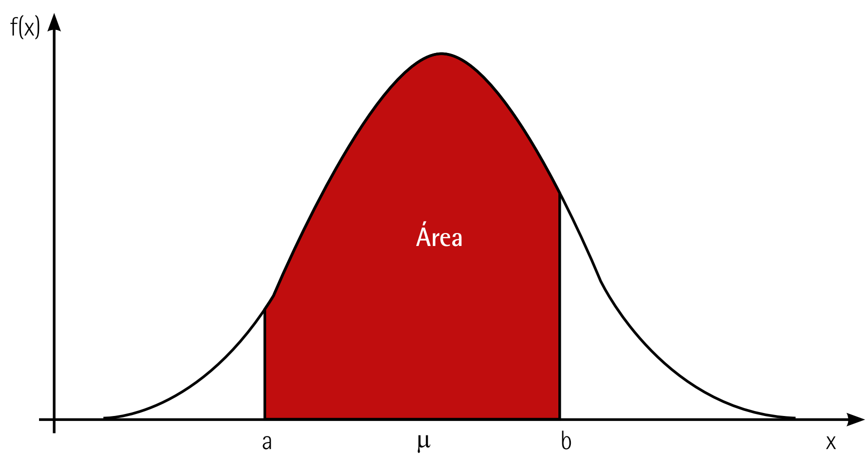



Figura 26 – A área indicada em vermelho na gaussiana fornece $P(a \leq x < b)$

E como podemos calcular essa área? Com o auxílio de uma tabela, que trabalha com o que chamamos de **normal reduzida** (ou normal padrão). Para isso, definimos a variável aleatória contínua Z a seguir.

$$Z = \frac{X - \mu}{\sigma}$$

Veja que essa expressão faz uma "conversão" da variável X para a variável Z. Por exemplo:

- se X = a, então $Z = \frac{a - \mu}{\sigma}$
- se X = b, então $Z = \frac{b - \mu}{\sigma}$

Para a variável Z, temos o que segue.

- Média E(Z) = 0
- Variância VAR(Z) = 1
- Desvio padrão DP(Z) = 1

Em resumo, se a variável aleatória contínua X segue uma distribuição normal de probabilidades de parâmetros E(X) = $\mu$ e VAR(X) = $\sigma^2$, ou seja, X ~ N($\mu$;$\sigma^2$), então a variável aleatória contínua Z dada por $Z = \frac{X - \mu}{\sigma}$ segue uma distribuição normal reduzida de probabilidades de parâmetros E(Z) = 0 e VAR(Z) = $1^2$ = 1, ou seja, Z ~ N(0;1). Além disso, temos o que segue.

- A média E(X) = $\mu$ corresponde à média E(Z) = 0.
- A variância VAR(X) = $\sigma^2$ corresponde à média VAR(Z) = 1.
- O desvio padrão DP(X) = $\sigma$ corresponde ao desvio padrão DP(Z) = 1.
- A abscissa X = a corresponde à abscissa $Z = \frac{a - \mu}{\sigma}$.
- A abscissa X = b corresponde à abscissa $Z = \frac{b - \mu}{\sigma}$.

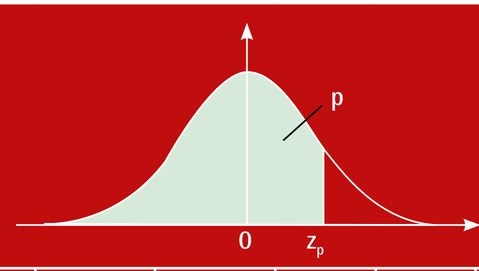
Logo, ficamos com:

$$P(a < X < b) = P\left(\frac{a - \mu}{\sigma} < Z < \frac{b - \mu}{\sigma}\right)$$

Vamos começar aprendendo a usar a tabela relativa à normal reduzida Z ~ N(0;1). Para isso, tomaremos como base a tabela a seguir, que fornece a probabilidade de termos $Z < z_p$, indicada por $P(Z < z_p)$.

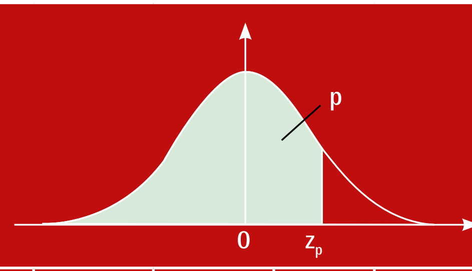
## Tabela 39 – Tabela normal reduzida (normal padrão)

Tabela normal reduzida
(normal padrão)
Z ~ N(0;1)




| z | 0,00 | 0,01 | 0,02 | 0,03 | 0,04 | 0,05 | 0,06 | 0,07 | 0,08 | 0,09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0,0 | 0,5000 | 0,5040 | 0,5080 | 0,5120 | 0,5160 | 0,5199 | 0,5239 | 0,5279 | 0,5319 | 0,5359 |
| 0,1 | 0,5398 | 0,5438 | 0,5478 | 0,5517 | 0,5557 | 0,5596 | 0,5636 | 0,5675 | 0,5714 | 0,5753 |
| 0,2 | 0,5793 | 0,5832 | 0,5871 | 0,5910 | 0,5948 | 0,5987 | 0,6026 | 0,6064 | 0,6103 | 0,6141 |
| 0,3 | 0,6179 | 0,6217 | 0,6255 | 0,6293 | 0,6331 | 0,6368 | 0,6406 | 0,6443 | 0,6480 | 0,6517 |
| 0,4 | 0,6554 | 0,6591 | 0,6628 | 0,6664 | 0,6700 | 0,6736 | 0,6772 | 0,6808 | 0,6844 | 0,6879 |
| 0,5 | 0,6915 | 0,6950 | 0,6985 | 0,7019 | 0,7054 | 0,7088 | 0,7123 | 0,7157 | 0,7190 | 0,7224 |
| 0,6 | 0,7257 | 0,7291 | 0,7324 | 0,7357 | 0,7389 | 0,7422 | 0,7454 | 0,7486 | 0,7517 | 0,7549 |
| 0,7 | 0,7580 | 0,7611 | 0,7642 | 0,7673 | 0,7704 | 0,7734 | 0,7764 | 0,7794 | 0,7823 | 0,7852 |
| 0,8 | 0,7881 | 0,7910 | 0,7939 | 0,7967 | 0,7995 | 0,8023 | 0,8051 | 0,8078 | 0,8106 | 0,8133 |
| 0,9 | 0,8159 | 0,8186 | 0,8212 | 0,8238 | 0,8264 | 0,8289 | 0,8315 | 0,8340 | 0,8365 | 0,8389 |
| 1,0 | 0,8413 | 0,8438 | 0,8461 | 0,8485 | 0,8508 | 0,8531 | 0,8554 | 0,8577 | 0,8599 | 0,8621 |
| 1,1 | 0,8643 | 0,8665 | 0,8686 | 0,8708 | 0,8729 | 0,8749 | 0,8770 | 0,8790 | 0,8810 | 0,8830 |
| 1,2 | 0,8849 | 0,8869 | 0,8888 | 0,8907 | 0,8925 | 0,8944 | 0,8962 | 0,8980 | 0,8997 | 0,9015 |
| 1,3 | 0,9032 | 0,9049 | 0,9066 | 0,9082 | 0,9099 | 0,9115 | 0,9131 | 0,9147 | 0,9162 | 0,9177 |
| 1,4 | 0,9192 | 0,9207 | 0,9222 | 0,9236 | 0,9251 | 0,9265 | 0,9279 | 0,9292 | 0,9306 | 0,9319 |
| 1,5 | 0,9332 | 0,9345 | 0,9357 | 0,9370 | 0,9382 | 0,9394 | 0,9406 | 0,9418 | 0,9429 | 0,9441 |
| 1,6 | 0,9452 | 0,9463 | 0,9474 | 0,9484 | 0,9495 | 0,9505 | 0,9515 | 0,9525 | 0,9535 | 0,9545 |
| 1,7 | 0,9554 | 0,9564 | 0,9573 | 0,9582 | 0,9591 | 0,9599 | 0,9608 | 0,9616 | 0,9625 | 0,9633 |
| 1,8 | 0,9641 | 0,9649 | 0,9656 | 0,9664 | 0,9671 | 0,9678 | 0,9686 | 0,9693 | 0,9699 | 0,9706 |
| 1,9 | 0,9713 | 0,9719 | 0,9726 | 0,9732 | 0,9738 | 0,9744 | 0,9750 | 0,9756 | 0,9761 | 0,9767 |
| 2,0 | 0,9772 | 0,9778 | 0,9783 | 0,9788 | 0,9793 | 0,9798 | 0,9803 | 0,9808 | 0,9812 | 0,9817 |
| 2,1 | 0,9821 | 0,9826 | 0,9830 | 0,9834 | 0,9838 | 0,9842 | 0,9846 | 0,9850 | 0,9854 | 0,9857 |
| 2,2 | 0,9861 | 0,9864 | 0,9868 | 0,9871 | 0,9875 | 0,9878 | 0,9881 | 0,9884 | 0,9887 | 0,9890 |
| 2,3 | 0,9893 | 0,9896 | 0,9898 | 0,9901 | 0,9904 | 0,9906 | 0,9909 | 0,9911 | 0,9913 | 0,9916 |
| 2,4 | 0,9918 | 0,9920 | 0,9922 | 0,9925 | 0,9927 | 0,9929 | 0,9931 | 0,9932 | 0,9934 | 0,9936 |
| 2,5 | 0,9938 | 0,9940 | 0,9941 | 0,9943 | 0,9945 | 0,9946 | 0,9948 | 0,9949 | 0,9951 | 0,9952 |
| 2,6 | 0,9953 | 0,9955 | 0,9956 | 0,9957 | 0,9959 | 0,9960 | 0,9961 | 0,9962 | 0,9963 | 0,9964 |
| 2,7 | 0,9965 | 0,9966 | 0,9967 | 0,9968 | 0,9969 | 0,9970 | 0,9971 | 0,9972 | 0,9973 | 0,9974 |
| 2,8 | 0,9974 | 0,9975 | 0,9976 | 0,9977 | 0,9977 | 0,9978 | 0,9979 | 0,9979 | 0,9980 | 0,9981 |
| 2,9 | 0,9981 | 0,9982 | 0,9982 | 0,9983 | 0,9984 | 0,9984 | 0,9985 | 0,9985 | 0,9986 | 0,9986 |
| 3,0 | 0,9987 | 0,9987 | 0,9987 | 0,9988 | 0,9988 | 0,9989 | 0,9989 | 0,9989 | 0,9990 | 0,9990 |
| 3,1 | 0,9990 | 0,9991 | 0,9991 | 0,9991 | 0,9992 | 0,9992 | 0,9992 | 0,9992 | 0,9993 | 0,9993 |
| 3,2 | 0,9993 | 0,9993 | 0,9994 | 0,9994 | 0,9994 | 0,9994 | 0,9994 | 0,9995 | 0,9995 | 0,9995 |

Tabela normal reduzida
(normal padrão)
Z ~ N(0;1)




| z | 0,00 | 0,01 | 0,02 | 0,03 | 0,04 | 0,05 | 0,06 | 0,07 | 0,08 | 0,09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 3,3 | 0,9995 | 0,9995 | 0,9995 | 0,9996 | 0,9996 | 0,9996 | 0,9996 | 0,9996 | 0,9996 | 0,9997 |
| 3,4 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9997 | 0,9998 |
| 3,5 | 0,9998 | 0,9998 | 0,9998 | 0,9998 | 0,9998 | 0,9998 | 0,9998 | 0,9998 | 0,9998 | 0,9998 |

Disponível em: https://bit.ly/3x4b04s. Acesso em: 17 nov. 2021.

## Exemplo de aplicação

Com base na tabela normal reduzida apresentada, vamos calcular as probabilidades dos exemplos 1 a 6.

**Exemplo 1. P(Z < 1,73)**

A probabilidade procurada está ilustrada a seguir.

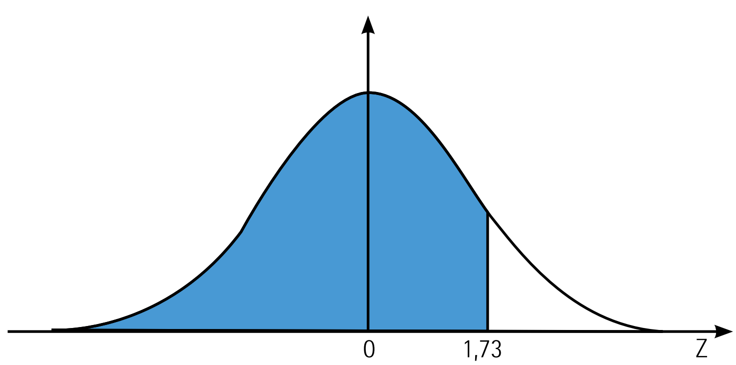



Figura 27

Como 1,73 é o resultado da soma de 1,7 e 0,03, na tabela normal reduzida, entramos com 1,7 na horizontal e com 0,03 na vertical, como destacado a seguir, e chegamos a P(Z < 1,73) = 0,9968.

**Tabela 40**

| Z | 0,03 |
|---|---|
| 1,7 | 0,9582 |

Vimos que, em uma f.d.p., a probabilidade associada a um valor "específico" (pontual da variável) é zero. Logo, temos o que segue.

- $P(x = a) = 0$ e $P(x = b) = 0$
- $P(a \leq x < b) = P(a < x < b) = P(a \leq x \leq b) = P(a < x \leq b)$

Desse modo, para o exemplo 1 que acabamos de resolver, concluímos:

$P(Z < 1{,}73) = P(Z \leq 1{,}73) = 0{,}9582$, pois $P(Z = 1{,}73) = 0$

**Exemplo 2. P(0 < Z < 1,68)**

A probabilidade procurada está ilustrada a seguir.

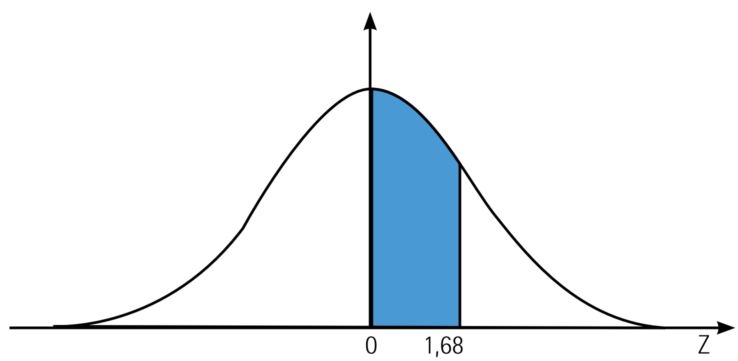



Figura 28

Inicialmente, vamos determinar $P(Z < 1{,}68)$. Como 1,68 é o resultado da soma de 1,6 e 0,08, na tabela normal reduzida, entramos com 1,6 na horizontal e com 0,08 na vertical, como destacado a seguir, e chegamos a $P(Z < 1{,}68) = 0{,}9535$.

**Tabela 41**

| Z | 0,08 |
|---|---|
| 1,6 | 0,9535 |

A probabilidade $P(Z < 1{,}68)$ está ilustrada a seguir.

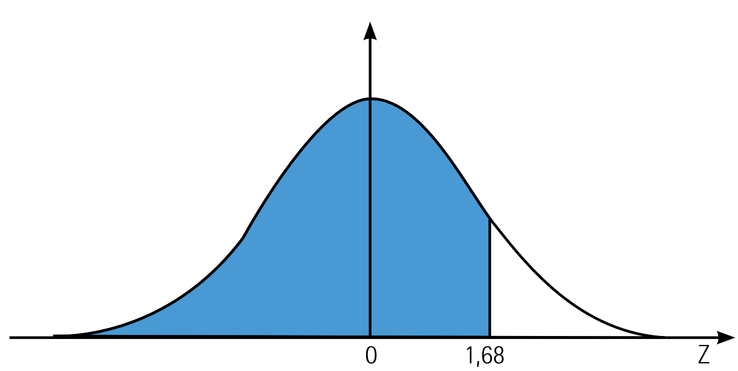



Figura 29

Sabemos que a área total delimitada pela curva normal (gaussiana) e pelo eixo horizontal vale 1. Como a curva normal reduzida tem média 0 (ela é simétrica em relação a essa média), a área da região à esquerda da média vale 0,5, que é metade da área total sob a curva, conforme ilustrado a seguir. Essa área equivale à probabilidade $P(Z < 0)$.

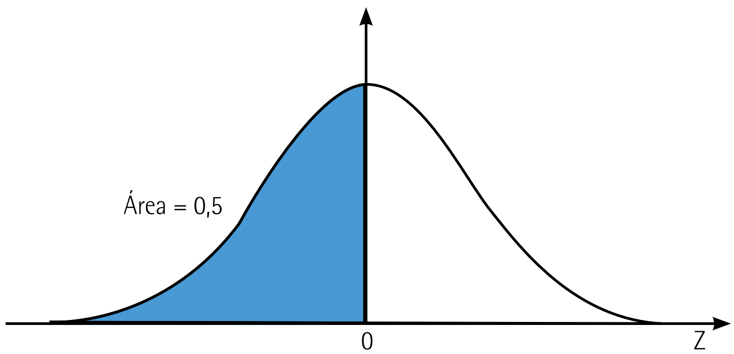



Figura 30

Como queremos $P(0 < Z < 1{,}68)$, devemos fazer:

$P(0 < Z < 1{,}68) = P(Z < 1{,}68) - P(Z < 0)$

$P(0 < Z < 1{,}68) = 0{,}9535 - 0{,}5$

$P(0 < Z < 1{,}68) = 0{,}4535$

**Exemplo 3. $P(-0{,}96 < Z \leq 0)$**

A probabilidade procurada está ilustrada a seguir.

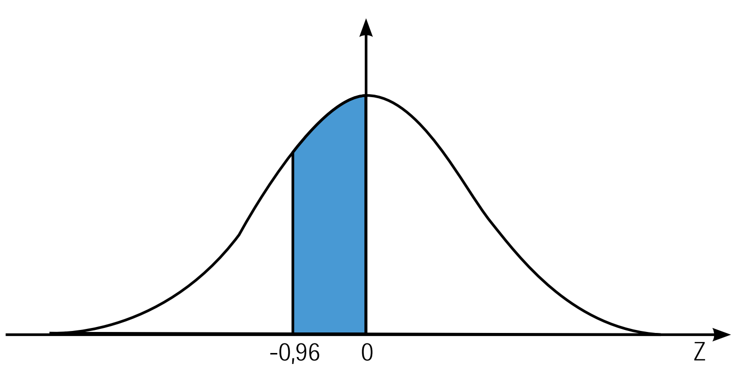



Figura 31

Sabemos que o gráfico da distribuição normal é simétrico em relação à média. Vemos, conforme ilustrado na figura a seguir, que $P(-0{,}96 < Z \leq 0)$ é igual a $P(0 < Z < 0{,}96)$.

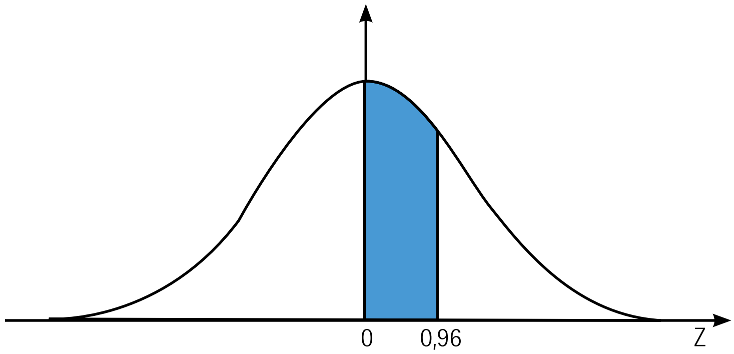



Figura 32

Inicialmente, vamos determinar P(Z < 0,96). Como 0,96 é o resultado da soma de 0,9 e 0,06, na tabela normal reduzida, entramos com 0,9 na horizontal e com 0,06 na vertical, como destacado a seguir, e chegamos a P(Z < 0,96) = 0,8315.

**Tabela 42**

| Z | 0,06 |
|---|---|
| 0,9 | 0,8315 |

A probabilidade P(Z < 0,96) está ilustrada a seguir.

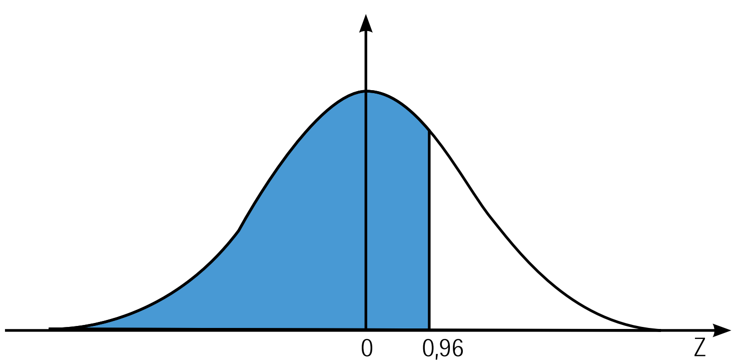



Figura 33

Sabemos que a área total delimitada pela curva normal (gaussiana) e pelo eixo horizontal vale 1. Como a curva normal reduzida tem média 0 (ela é simétrica em relação a essa média), a área da região à esquerda da média vale 0,5, que é metade da área total sob a curva, conforme ilustrado a seguir. Essa área equivale à probabilidade P(Z < 0).

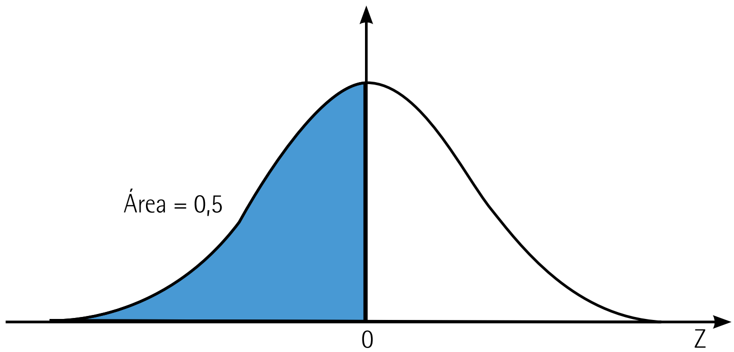



Figura 34

Como queremos $P(0 < Z < 0{,}96) = P(-0{,}96 < Z \leq 0)$, devemos fazer:

$P(0 < Z < 0{,}96) = P(-0{,}96 < Z \leq 0) = P(Z < 0{,}96) - P(Z < 0) = 0{,}8315 - 0{,}5$

$P(0 < Z < 1{,}68) = 0{,}3315$

**Exemplo 4. $P(1{,}23 < Z < 2{,}17)$**

A probabilidade procurada está ilustrada a seguir.

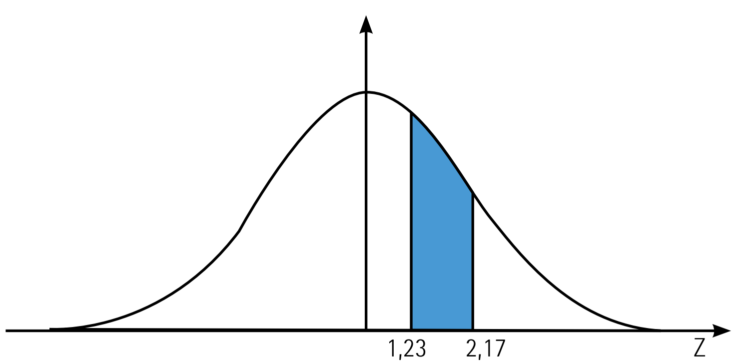



Figura 35

Inicialmente, vamos determinar $P(Z < 1{,}23)$. Como 1,23 é o resultado da soma de 1,2 e 0,03, na tabela normal reduzida, entramos com 1,2 na horizontal e com 0,03 na vertical, como destacado a seguir, e chegamos a $P(Z < 1{,}23) = 0{,}8907$.

**Tabela 43**

| Z | 0,03 |
|---|---|
| 1,2 | 0,8907 |

Agora, vamos determinar $P(Z < 2,17)$. Como 2,17 é o resultado da soma de 2,1 e 0,07, na tabela normal reduzida, entramos com 2,1 na horizontal e com 0,07 na vertical, como destacado a seguir, e chegamos a $P(Z < 2,17) = 0,9850$.

**Tabela 44**

| Z | 0,07 |
|---|---|
| 2,1 | 0,9850 |

Como queremos $P(1,23 < Z < 2,17)$, devemos fazer:

$$P(1,23 < Z < 2,17) = P(Z < 2,17) - P(Z < 1,23) = 0,9850 - 0,8907$$

$$P(1,23 < Z < 2,17) = 0,0943$$

**Exemplo 5. $P(-2,29 < Z < -1,31)$**

A probabilidade procurada está ilustrada a seguir.

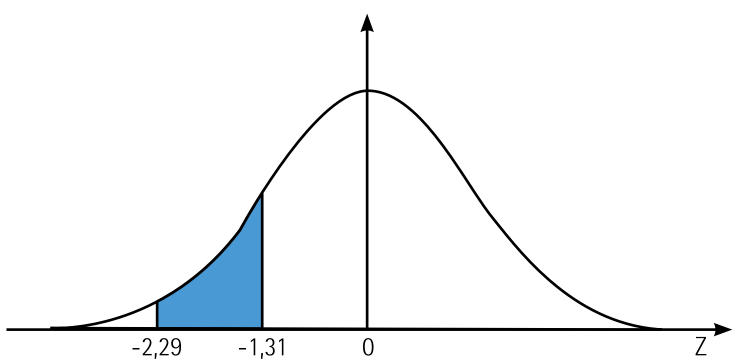



Figura 36

Uma vez que a distribuição normal é simétrica em relação à média, vemos, conforme ilustrado na figura a seguir, que $P(-2,29 < Z < -1,31)$ é igual a $P(1,31 < Z < 2,29)$.

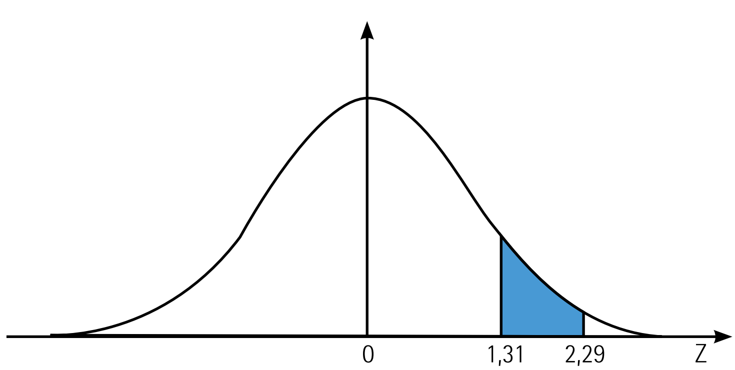



Figura 37

Vamos determinar $P(Z < 2{,}29)$. Como 2,29 é o resultado da soma de 2,2 e 0,09, na tabela normal reduzida, entramos com 2,2 na horizontal e com 0,09 na vertical, como destacado a seguir, e chegamos a $P(Z < 2{,}29) = 0{,}9890$.

**Tabela 45**

| Z | 0,09 |
|---|---|
| 2,2 | 0,9890 |

Vamos determinar $P(Z < 1{,}31)$. Como 1,31 é o resultado da soma de 1,3 e 0,01, na tabela normal reduzida, entramos com 1,3 na horizontal e com 0,01 na vertical, como destacado a seguir, e chegamos a $P(Z < 1{,}31) = 0{,}9049$.

**Tabela 46**

| Z | 0,01 |
|---|---|
| 1,3 | 0,9049 |

Como queremos $P(-2{,}29 < Z < -1{,}31)$, que é igual a $P(1{,}31 < Z < 2{,}29)$, devemos fazer:

$$P(-2{,}29 < Z < -1{,}31) = P(1{,}31 < Z < 2{,}29) = P(Z < 2{,}29) - P(Z < 1{,}31) = 0{,}9890 - 0{,}9049$$

$$P(1{,}23 < Z < 2{,}17) = 0{,}0841$$

**Exemplo 6. $P(Z > 1{,}06)$**

A probabilidade procurada está ilustrada a seguir.

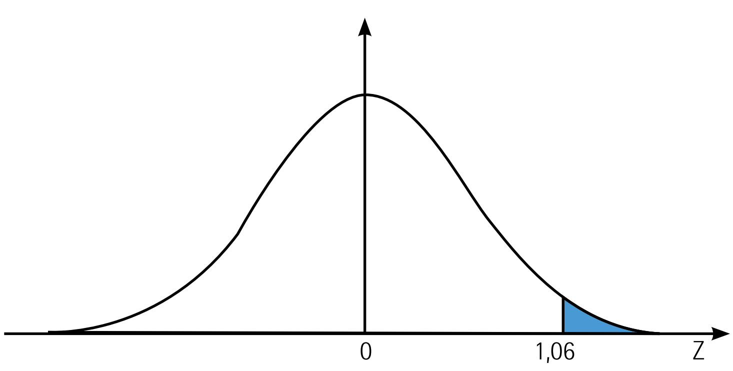



Figura 38

Inicialmente, vamos determinar $P(Z < 1{,}06)$. Como 1,06 é o resultado da soma de 1,0 e 0,06, na tabela normal reduzida, entramos com 1,0 na horizontal e com 0,06 na vertical, como destacado a seguir, e chegamos a $P(Z < 1{,}06) = 0{,}8554$.

**Tabela 47**

| Z | 0,06 |
|---|---|
| 1,0 | 0,8554 |

A probabilidade $P(Z < 1{,}06)$ está ilustrada a seguir.

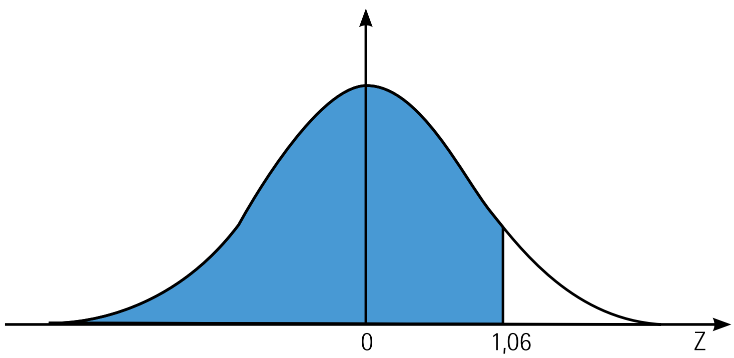



Figura 39

Sabemos que a área total delimitada pela curva normal (gaussiana) e pelo eixo horizontal vale 1, conforme ilustrado a seguir.

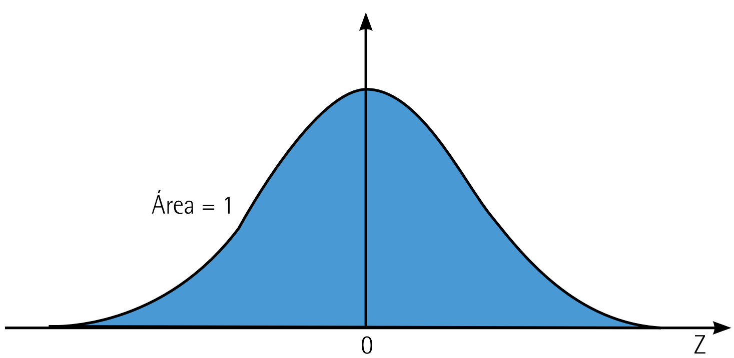



Figura 40

Como queremos $P(Z > 1{,}06)$, devemos fazer:

$P(Z > 1{,}06) = 1 - P(Z < 1{,}06) = 1 - 0{,}8554$

$P(Z > 1{,}06) = 0{,}1446$

Agora, considere que a variável aleatória contínua X siga modelo normal com parâmetros $\mu = 23$ (média 23) e $\sigma^2 = 81$ (variância 81), ou seja, $X \sim N(23;81)$. Para essa variável, vamos calcular as probabilidades solicitadas nos exemplos 7 a 9. Destacamos que, como o desvio padrão $\sigma$ é a raiz quadrada da variância, no caso temos $\sigma = 9$.

**Exemplo 7. P(X < 28)**

A probabilidade procurada está ilustrada a seguir.

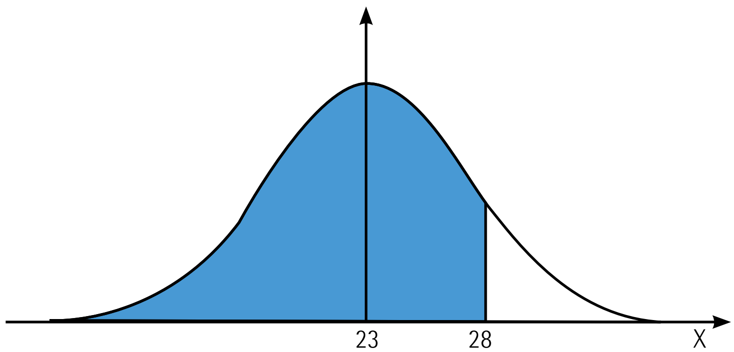



Figura 41

Vamos fazer a “conversão” da variável X para a variável Z, em que Z ~ N(0;1). Logo, ficamos com:

$$P(X<28)=P\left(Z<\frac{28-\mu}{\sigma}\right)=P\left(Z<\frac{28-23}{9}\right)=P\left(Z<\frac{5}{9}\right)=P\left(Z<0,55\right)$$

A probabilidade procurada, em termos da variável Z, está ilustrada a seguir.

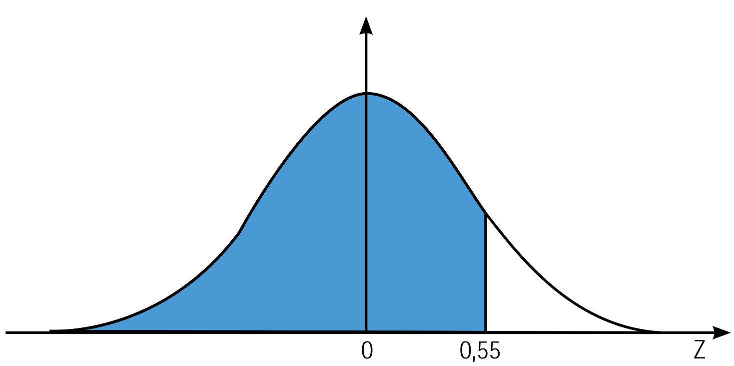



Figura 42

Como 0,55 é o resultado da soma de 0,5 e 0,05, na tabela normal reduzida, entramos com 0,5 na horizontal e com 0,05 na vertical, como destacado a seguir, e chegamos a P(Z < 0,55) = 0,7088.

**Tabela 48**

| Z | 0,05 |
|---|---|
| 0,5 | 0,7088 |

Logo, P(X < 28) = P(Z < 0,55) = 0,7088.

**Exemplo 8. P(20,5 < X < 24)**

A probabilidade procurada está ilustrada a seguir.

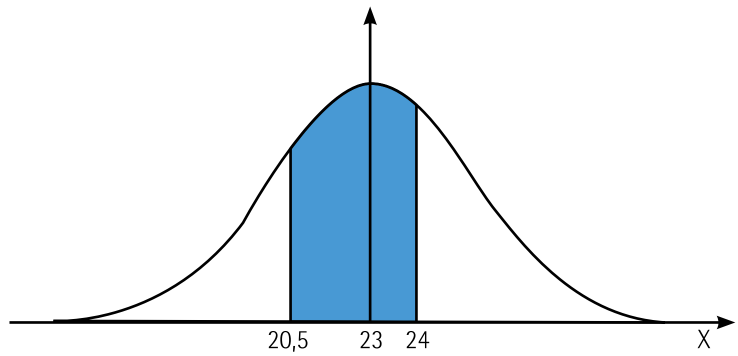



Figura 43

Vamos fazer a "conversão" da variável X para a variável Z, em que Z ~ N(0;1). Logo, ficamos com:

$$P(20{,}5 < X < 24) = P\left(\frac{20{,}5-\mu}{\sigma} < Z < \frac{24-\mu}{\sigma}\right) = P\left(\frac{20{,}5-23}{9} < Z < \frac{24-23}{9}\right)$$

$$P(20{,}5 < X < 24) = P\left(\frac{-2{,}5}{9} < Z < \frac{1}{9}\right) = P(-0{,}28 < Z < 0{,}11)$$

A probabilidade procurada, em termos da variável Z, está ilustrada a seguir.

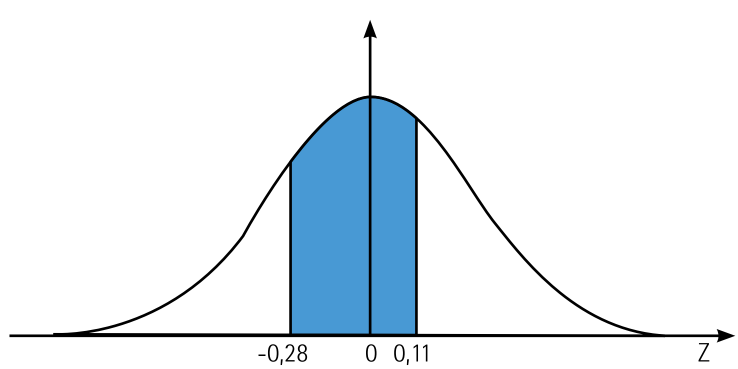



Figura 44

Logo, queremos saber o resultado de:

$P(20{,}5 < X < 24) = P(-0{,}28 < Z < 0{,}11) = P(Z < 0{,}11) - P(Z < -0{,}28)$

Por simetria, temos que $P(Z < -0{,}28) = P(Z > 0{,}28)$, como pode ser visualizado a seguir.

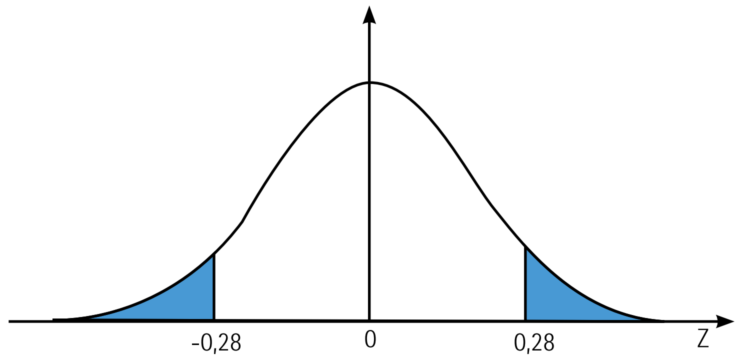



Figura 45

Desse modo, ficamos com:

$P(20{,}5 < X < 24) = P(-0{,}28 < Z < 0{,}11) = P(Z < 0{,}11) - P(Z < -0{,}28)$

Temos que $P(Z > 0{,}28) = 1 - P(Z < 0{,}28)$. Assim, chegamos a:

$P(20{,}5 < X < 24) = P(-0{,}28 < Z < 0{,}11) = P(Z < 0{,}11) - (1 - P(Z < -0{,}28))$

$P(20{,}5 < X < 24) = P(-0{,}28 < Z < 0{,}11) = P(Z < 0{,}11) - 1 + P(Z < -0{,}28)$

$P(20{,}5 < X < 24) = P(-0{,}28 < Z < 0{,}11) = P(Z < 0{,}11) + P(Z < -0{,}28) - 1$

Como 0,11 é o resultado da soma de 0,1 e 0,01, na tabela normal reduzida, entramos com 0,1 na horizontal e com 0,01 na vertical, como destacado a seguir, e chegamos a $P(Z < 0{,}11) = 0{,}5438$.

**Tabela 49**

| Z | 0,01 |
|---|---|
| 0,1 | 0,5438 |

Como 0,28 é o resultado da soma de 0,2 e 0,08, na tabela normal reduzida, entramos com 0,2 na horizontal e com 0,08 na vertical, como destacado a seguir, e chegamos a $P(Z < 0{,}28) = 0{,}6103$.

**Tabela 50**

| Z | 0,08 |
|---|---|
| 0,2 | 0,6103 |

Logo:

$P(20{,}5 < X < 24) = P(-0{,}28 < Z < 0{,}11) = P(Z < 0{,}11) + P(Z < -0{,}28) - 1$

$P(20{,}5 < X < 24) = P(-0{,}28 < Z < 0{,}11) = 0{,}5438 + 0{,}6103 - 1$

$P(20{,}5 < X < 24) = P(-0{,}28 < Z < 0{,}11) = 0{,}1568$

**Exemplo 9. P(X < 21,7 ou X > 26)**

A probabilidade procurada está ilustrada a seguir.

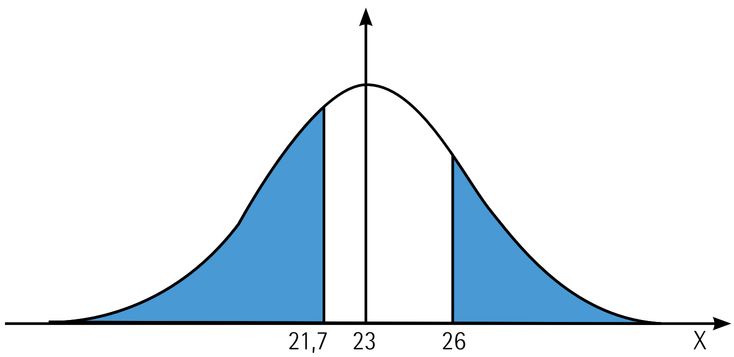



Figura 46

Vamos fazer a "conversão" da variável X para a variável Z, em que Z ~ N(0;1). Logo, ficamos com:

$$P(X < 21{,}7 \text{ ou } X > 26) = P\left(Z < \frac{21{,}7 - \mu}{\sigma} \text{ ou } Z > \frac{26 - \mu}{\sigma}\right)$$

$$P(X < 21{,}7 \text{ ou } X > 26) = P\left(Z < \frac{21{,}7 - 23}{9} \text{ ou } Z > \frac{26 - 23}{9}\right)$$

$P(X < 21{,}7 \text{ ou } X > 26) = P(Z < -0{,}14 \text{ ou } Z > 0{,}33)$

A probabilidade procurada, em termos da variável Z, está ilustrada a seguir.

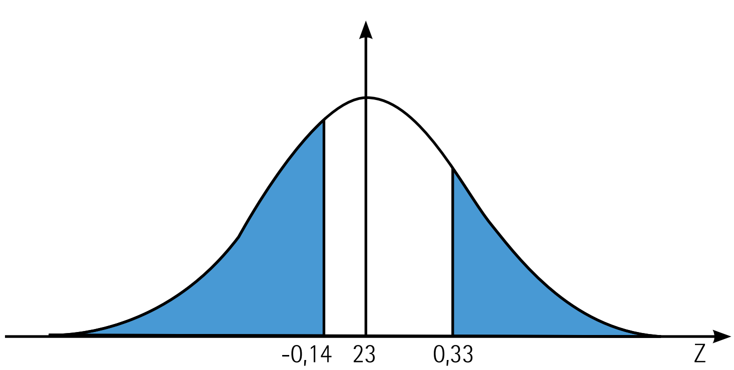



Figura 47

Logo, queremos saber o resultado de:

$P(X < 21{,}7 \text{ ou } X > 26) = P(Z < -0{,}14 \text{ ou } Z > 0{,}33) = P(Z < -0{,}14) + P(Z > 0{,}33)$

Por simetria, temos que $P(Z < -0{,}14) = P(Z > 0{,}14)$, como pode ser visualizado a seguir.

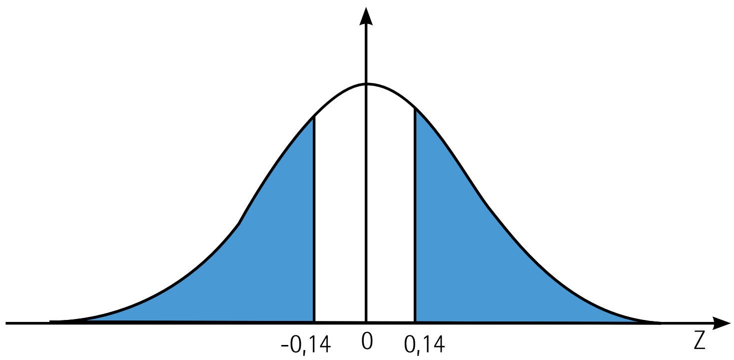



Figura 48

Desse modo, ficamos com:

$P(X < 21{,}7 \text{ ou } X > 26) = P(Z < -0{,}14 \text{ ou } Z > 0{,}33) = P(Z > 0{,}14) + P(Z > 0{,}33)$

Temos que $P(Z > 0{,}14) = 1 - P(Z < 0{,}14)$ e $P(Z > 0{,}33) = 1 - P(Z < 0{,}33)$.

Assim, chegamos a:

$P(X < 21{,}7 \text{ ou } X > 26) = P(Z < -0{,}14 \text{ ou } Z > 0{,}33) = 1 - P(Z < 0{,}14) + 1 - P(Z < 0{,}33)$

$P(X < 21{,}7 \text{ ou } X > 26) = P(Z < -0{,}14 \text{ ou } Z > 0{,}33) = 2 - P(Z < 0{,}14) - P(Z < 0{,}33)$

Como 0,14 é o resultado da soma de 0,1 e 0,04, na tabela normal reduzida, entramos com 0,1 na horizontal e com 0,04 na vertical, como destacado a seguir, e chegamos a $P(Z < 0{,}14) = 0{,}5557$.

**Tabela 51**

| Z | 0,04 |
|---|---|
| 0,1 | 0,5557 |

Como 0,33 é o resultado da soma de 0,3 e 0,03, na tabela normal reduzida, entramos com 0,3 na horizontal e com 0,03 na vertical, como destacado a seguir, e chegamos a $P(Z < 0{,}33) = 0{,}6293$.

**Tabela 52**

| Z | 0,03 |
|---|---|
| 0,3 | 0,6293 |

Logo:

$P(X < 21{,}7 \text{ ou } X > 26) = P(Z < -0{,}14 \text{ ou } Z > 0{,}33) = 2 - P(Z < 0{,}14) - P(Z < 0{,}33)$

$P(X < 21{,}7 \text{ ou } X > 26) = P(Z < -0{,}14 \text{ ou } Z > 0{,}33) = 2 - 0{,}5557 - 0{,}6293$

$P(X < 21{,}7 \text{ ou } X > 26) = P(Z < -0{,}14 \text{ ou } Z > 0{,}33) = 0{,}815$

**Exemplo 10**

Vamos estudar uma situação em que aplicaremos o que vimos neste item.

Imagine que um programa de computador tenha de processar um conjunto de imagens cujo tamanho siga uma distribuição normal de probabilidades com média igual a 2,1 Mb e desvio padrão igual a 0,6 Mb, em que Mb representa megabit.

Chamamos de T a variável aleatória contínua que representa o tamanho da imagem, em Mb. Sabemos que T segue uma distribuição normal de probabilidades.

Para esse caso, qual é a probabilidade de uma imagem ter tamanho inferior a 2,5 Mb?

A média $\mu$ da variável T é 2,1 Mb e seu desvio padrão $\sigma$ é 0,6 Mb. Logo, a variância $\sigma^2$ da variável T é 0,36 $Mb^2$, pois $0{,}6^2 = 0{,}36$.

Assim, temos $T \sim N(2{,}1;0{,}36)$.

Queremos determinar a probabilidade de uma imagem ter tamanho inferior a 2,5 Mb, ou seja, queremos determinar $P(T < 2{,}5)$.

A probabilidade procurada está ilustrada a seguir.

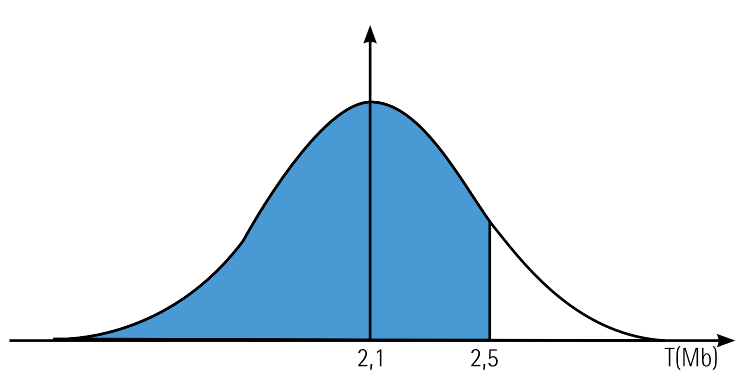



Figura 49

Vamos fazer a "conversão" da variável T para a variável Z, em que Z ~ N(0;1). Logo, ficamos com:

$$P(T<2{,}5)=P\left(Z<\frac{2{,}5-\mu}{\sigma}\right)=P\left(Z<\frac{2{,}5-2{,}1}{0{,}6}\right)=P\left(Z<0{,}67\right)$$

Como 0,67 é o resultado da soma de 0,6 e 0,07, na tabela normal reduzida, entramos com 0,6 na horizontal e com 0,07 na vertical, como destacado a seguir, e chegamos a P(Z < 0,67) = 0,7486.

**Tabela 53**

| Z | 0,07 |
|---|---|
| 0,6 | 0,7486 |

Logo, P(T < 2,5) = P(Z < 0,67) = 0,7486.

Ainda com base na situação em estudo, qual deve ser o espaço mínimo de memória que deve ser reservado para que 96% das imagens sejam processadas, com uma imagem de cada vez?

Nesse caso, queremos achar o tamanho t*, para o qual P(T < t*) = 0,96.

Vamos fazer a "conversão" da variável T para a variável Z, em que Z ~ N(0;1). Logo, ficamos com:

$$P(T<t^*)=P\left(Z<\frac{t^*-\mu}{\sigma}\right)=P\left(Z<\frac{t^*-2{,}1}{0{,}6}\right)=0{,}96$$

Vamos chamar $\frac{t^*-2{,}1}{0{,}6}$ de z*. Logo, queremos achar z*, de modo que P(Z < z*) = 0,96, o que está ilustrado a seguir.

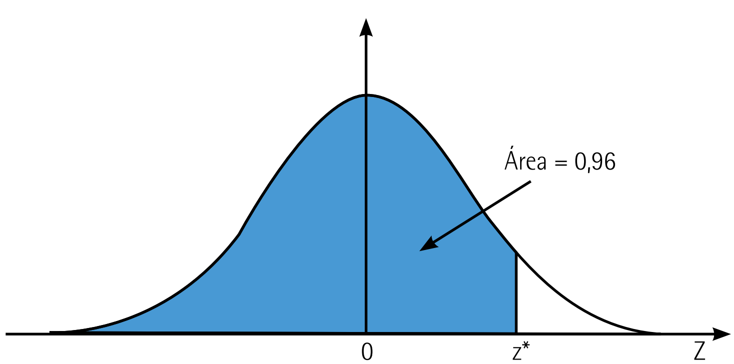



Figura 50

Precisamos encontrar, "dentro" da tabela normal reduzida, o valor 0,96. Vemos que, na tabela 39, o valor mais próximo de 0,96 é 0,9599, e ele corresponde a $z^* = 1,75$ (1,7 na horizontal e 0,05 na vertical).

**Tabela 54**

| Z | 0,05 |
|---|---|
| 1,7 | $0,9599 \approx 0,96$ |

Desse modo, temos:

$$z^* = 1,75 = \frac{t^* - 2,1}{0,6} \Rightarrow t^* - 2,1 = 1,75 . 0,6 \Rightarrow t^* = 3,15$$

Concluímos que o espaço mínimo de memória reservado para que 96% das imagens possam ser processadas é de 3,15 Mb.

## Resumo

Começamos a unidade I estudando a estatística descritiva, com destaque para a noção de conjunto de dados.

Vimos que a estatística descritiva visa a organizar, a descrever, a explorar, a expressar e a resumir informações brutas. Nesse sentido, estamos lidando com algo "certo", que já ocorreu, que "pertence ao passado", ou seja, estamos lidando com um conjunto de dados vindos, por exemplo, da observação de algum evento já acontecido, do relato de um questionário já aplicado ou de uma contagem de casos já vistos. Em síntese, os resultados observados formam o conjunto de dados que, usualmente, mostra os possíveis valores da característica estudada e suas respectivas contagens.

Podemos organizar as informações contidas em uma tabela de dados brutos de modo a obtermos tabelas de frequência absoluta, de frequência relativa e de percentuais. As frequências relativas correspondem às frequências absolutas divididas pelo número total de medições, e os percentuais são as frequências relativas multiplicadas por 100.

Com base nas tabelas de dados organizados, podemos, por exemplo, expressar as informações por meio de diferentes tipos de gráfico e fazer cálculos das medidas de tendência central (média, moda e mediana) e das medidas de dispersão (variância e desvio padrão).

Prosseguimos nossos estudos com a parte das probabilidades, em que lidamos com a chance de algo acontecer, ou seja, trata-se do campo da incerteza, dos eventos aleatórios, dos acontecimentos que não podem ser previstos com total exatidão.

Observamos que o lançamento de uma moeda honesta é um exemplo típico de fenômeno aleatório. Antes de fazermos o lançamento, não podemos dizer, com certeza, se teremos cara ou coroa. Sabemos que há 50% de probabilidade de obtermos cara quando lançamos essa moeda, mas não sabemos se isso de fato acontecerá.

Uma maneira de calcularmos a probabilidade P de sucesso em um evento aleatório, como ganharmos um prêmio em um sorteio, é fazermos a divisão do "número de casos favoráveis para nosso sucesso" pelo "número total de casos possíveis".

Como o número de casos favoráveis é menor do que número total de casos possíveis ou igual ao número total de casos possíveis, a probabilidade P é um número real que varia de 0 até 1 ($0 \leq P \leq 1$). Muitas vezes, a probabilidade P é multiplicada por 100% para expressarmos o resultado em percentual (%).

Definimos probabilidade da união de eventos, probabilidade da intersecção de eventos, probabilidade condicional e eventos independentes, fizemos diversos exemplos de aplicação com base em situações do dia a dia e também exploramos contextos específicos da área da computação.

Vimos que a quantidade X associada a cada possível resultado de um evento amostral é chamada de variável aleatória discreta se ela assumir valores em um conjunto enumerável e com certa probabilidade de ocorrência. Com base nesse conceito, estudamos os principais modelos de distribuição discreta de probabilidade:

- o modelo uniforme discreto, em que todos os possíveis valores da variável aleatória discreta têm a mesma probabilidade de acontecer;

- o modelo de Bernoulli, em que um experimento aleatório admite como resultado apenas sucesso (no qual a variável aleatória discreta assume o valor 1) ou fracasso (no qual a variável aleatória discreta assume o valor 0);

- o modelo binomial, em que a variável aleatória discreta corresponde ao número de sucessos obtidos na realização de n ensaios de Bernoulli independentes, todos com a mesma probabilidade P de sucesso.

Apontamos que as distribuições contínuas de probabilidade operam com variáveis aleatórias contínuas, ou seja, com variáveis cujos valores ocorrem de modo aleatório e cujo domínio está em intervalo de números reais.

Ressaltamos que, diferentemente do que acontece com as variáveis aleatórias discretas, as variáveis aleatórias contínuas podem assumir "infinitos" valores, ou seja, não conseguimos listar individualmente todos os seus possíveis valores. Concluímos que, no caso de uma distribuição contínua de probabilidade, falamos de probabilidades associadas a intervalos de valores da variável X, e não da probabilidade de um valor "específico" dessa variável.

Nessa situação, representamos graficamente o que chamamos de função densidade de probabilidade, indicada por f.d.p. e representada por f(x). Mostramos que essa função apresenta as propriedades a seguir.

- Qualquer que seja o valor de x, $f(x) \geq 0$. Ou seja, o gráfico de f(x) "nunca" ocupa região abaixo do eixo x.

- Numericamente, a área delimitada pelo gráfico de f(x) e pelo eixo x é igual a 1. Essa área representa a probabilidade de ocorrência de todos os possíveis valores de x, o que resulta em 100%.

Explicamos que a probabilidade de termos $a \leq x < b$, representada por $P(a \leq x < b)$, é numericamente igual à área delimitada pelo gráfico de f(x), pelo eixo x e pelas retas verticais $x = a$ e $x = b$.

Abordamos os principais modelos de distribuição contínua de probabilidade, a saber:

- o modelo uniforme contínuo, em que f.d.p. da variável aleatória contínua X no intervalo [a;b] é dada por $f(x) = 1/(b - a)$, com $f(x) = 0$ se $x < a$ ou se $x > b$;

- o modelo normal, em que f.d.p. da variável aleatória contínua X é a famosa curva em forma de sino, simétrica em relação à média.

Vimos que, se a variável aleatória contínua X segue uma distribuição normal de probabilidades de parâmetros média E(X) igual a $\mu$ e variância VAR(X) igual a $\sigma^2$, ou seja, $X \sim N(\mu;\sigma^2)$, então a variável aleatória contínua Z dada por $Z = \frac{X-\mu}{\sigma}$ segue uma distribuição normal reduzida (padronizada) de parâmetros média E(Z) igual a 0 e variância VAR(Z) igual a 1.

Fizemos diversos exemplos de determinação de probabilidades usando valores da tabela normal reduzida (normal padrão) e estudamos um caso de distribuição normal aplicada ao tamanho de um conjunto de imagens que um computador terá de processar.

## Exercícios

**Questão 1**. A empresa Software Para Você fornece soluções computacionais para empresas que atuam em várias áreas do comércio. No gráfico a seguir, temos a distribuição do tempo, em horas, que os desenvolvedores dessa empresa levaram para responder às demandas dos 200 clientes atendidos no último mês.

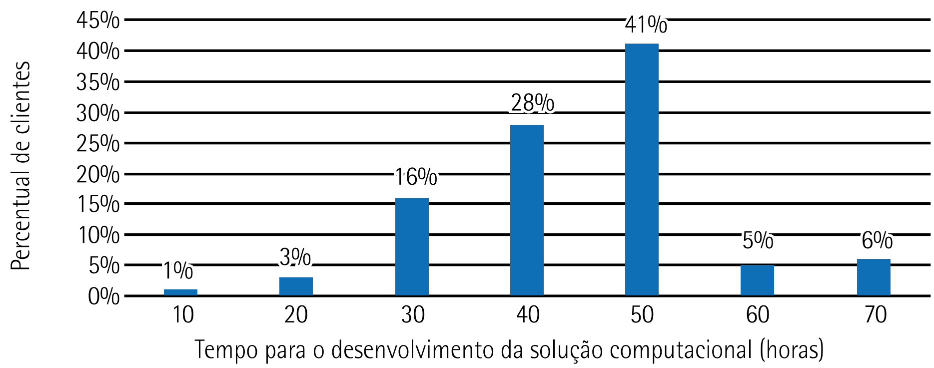



Figura 51

O tempo médio que os desenvolvedores da Software Para Você levaram para responder às demandas dos 200 clientes atendidos no último mês é igual a:

A) 44,4 horas.

B) 35,0 horas.

C) 50,5 horas.

D) 37,8 horas.

E) 50,5 horas.

Resposta correta: alternativa A.

**Análise da questão**

A quantidade total de clientes atendidos no último mês foi igual a 200.

Pela leitura do gráfico do enunciado, podemos concluir que, dos 200 clientes:

- 2 deles (1% de 200) demandaram 10 horas para a resposta às demandas;

- 6 deles (3% de 200) demandaram 20 horas para a resposta às demandas;
- 32 deles (16% de 200) demandaram 30 horas para a resposta às demandas;
- 56 deles (28% de 200) demandaram 40 horas para a resposta às demandas;
- 82 deles (41% de 200) demandaram 50 horas para a resposta às demandas;
- 10 deles (5% de 200) demandaram 60 horas para a resposta às demandas;
- 12 deles (6% de 200) demandaram 70 horas para a resposta às demandas.

Os cálculos feitos podem ser resumidos na tabela a seguir.

**Tabela 55**

| Tempo (horas) | Frequência de clientes |
|---|---|
| 10 | 2 |
| 20 | 6 |
| 30 | 32 |
| 40 | 56 |
| 50 | 82 |
| 60 | 10 |
| 70 | 12 |
| Total | 200 clientes |

Na tabela, temos o total de 200 clientes. Para calcularmos o tempo médio, precisamos somar todos os 200 valores da tabela e dividir essa soma por 100. Observe que, com base na tabela:

- 10 (horas) é um valor que precisa ser somado 2 vezes;
- 20 (horas) é um valor que precisa ser somado 6 vezes;
- 30 (horas) é um valor que precisa ser somado 32 vezes;
- 40 (horas) é um valor que precisa ser somado 56 vezes;
- 50 (horas) é um valor que precisa ser somado 82 vezes;
- 60 (horas) é um valor que precisa ser somado 10 vezes;
- 70 (horas) é um valor que precisa ser somado 12 vezes.

Logo, o tempo médio é igual a 44,4 horas, conforme calculado a seguir.

$$\text{Tempo médio} = \frac{10\,.\,2 + 20\,.\,6 + 30\,.\,32 + 40\,.\,56 + 50\,.\,82 + 60\,.\,10 + 70\,.\,12}{200}$$

$$\text{Tempo médio} = \frac{20 + 120 + 960 + 2240 + 4100 + 600 + 840}{200} = \frac{8880}{200} = 44,4$$

Tempo médio = 44,4

**Questão 2**. (Enade 2009) Uma empresa produz um tipo especial de motor. A quantidade em estoque desse motor segue uma distribuição normal com média de 200 unidades e desvio padrão de 20. O gráfico a seguir representa a distribuição normal padrão (média igual a 0 e desvio padrão igual a 1), em que as porcentagens representam as probabilidades entre os valores de desvio padrão.

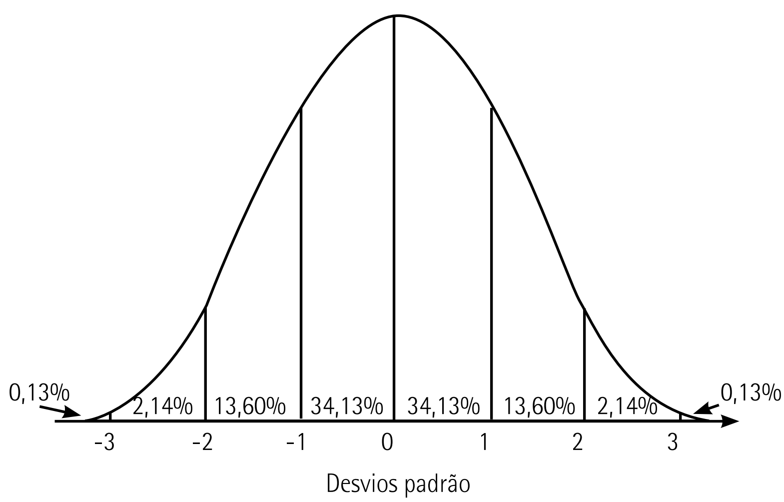



Figura 52

Qual é a probabilidade de, em dado momento, o estoque da empresa apresentar mais de 220 unidades?

A) 84,13%.

B) 68,26%.

C) 34,13%.

D) 15,87%.

E) 13,60%.

Resposta correta: alternativa D.

**Análise da questão**

A questão envolve o uso da distribuição normal para o cálculo de probabilidades. A normal é a mais frequente de todas as distribuições de probabilidade, sendo definida pela média $\mu$ e pelo desvio padrão $\sigma$.

Para facilitar os cálculos, é utilizada uma curva normal padrão (também chamada de normal reduzida), que assume, por definição, os seguintes valores.

- 0 para a média $\mu$.
- 1 para o desvio padrão $\sigma$.

Os valores relativos à distribuição normal padronizada são obtidos via tabela.

Ou seja, os cálculos são feitos pela utilização da tabela da curva normal padronizada, cujo dado de entrada é o valor da variável normal reduzida (Z), calculado pela seguinte fórmula:

$$Z = \frac{X - \mu}{\sigma}$$

A curva normal padrão varia no eixo horizontal desde –4 até +4, correspondendo ao número de desvios padrão distantes da média, cujo valor é 0.

Pode-se afirmar que a probabilidade de um valor estar entre 4 desvios padrão para menos e 4 desvios padrão para mais é de aproximadamente 100%. Assim, por exemplo, entre a média e um desvio padrão para mais, está a probabilidade de 34,13% (informação tabelada). Qualquer outro cálculo de probabilidade similar pode ser feito pela análise das áreas envolvidas por meio da utilização da tabela da distribuição normal reduzida ou padronizada.

Para ilustrar o exposto, considere o seguinte exemplo. Uma turma de alunos fez uma prova de estatística, e as notas seguem uma distribuição normal cuja média foi 5,2 e cujo desvio padrão foi 0,6. Qual é a probabilidade de que um aluno qualquer dessa turma tenha nota entre 5,8 e 6,4? Utilizando a fórmula da variável normal reduzida, obtém-se o número de desvios padrão correspondente a cada nota, conforme mostrado a seguir.

$$Z = \frac{X - \mu}{\sigma} \rightarrow Z = \frac{5{,}8 - 5{,}2}{0{,}6} \rightarrow Z = 1$$

$$Z = \frac{X - \mu}{\sigma} \rightarrow Z = \frac{6{,}4 - 5{,}2}{0{,}6} \rightarrow Z = 2$$

Deseja-se, portanto, o cálculo da probabilidade entre 1 e 2 desvios padrão acima da média (observe o fato de os valores de Z serem positivos; caso fossem negativos, estariam abaixo da média). Isso é feito pela determinação da área correspondente. O gráfico da figura a seguir mostra essa situação.

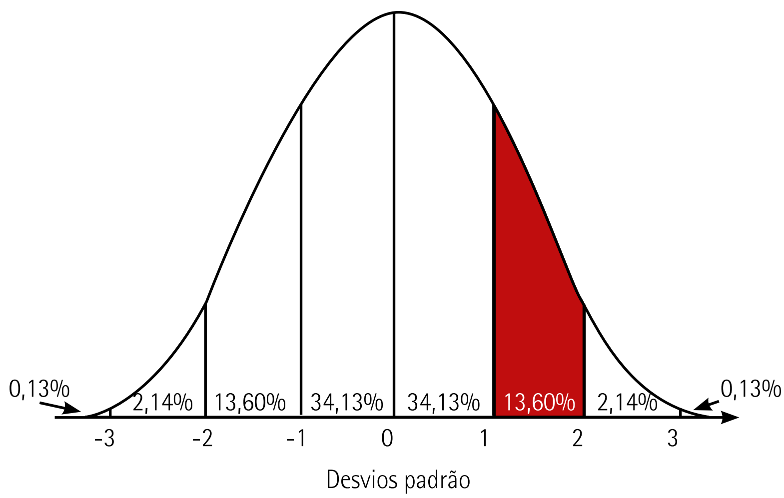



Figura 53

A área preenchida corresponde ao desejado, ou seja, alunos que tiraram notas entre 5,8 e 6,4 (em termos de variáveis reduzidas, entre 1 e 2). Nota-se também que essa área vale 13,60%. Assim, é possível afirmar que a probabilidade de um aluno qualquer dessa turma ter tirado nota entre 5,8 e 6,4 é de 13,60%.

No caso da questão, devemos ter em mente os parâmetros estatísticos fornecidos (média igual a 200 e desvio padrão igual a 20). É pedida a probabilidade de ocorrência da faixa acima de 220 unidades, ou seja, um desvio padrão ou mais acima da média. Vejamos:

$$Z = \frac{X - \mu}{\sigma} \rightarrow Z = \frac{220 - 200}{20} \rightarrow Z = 1$$

O enunciado não fornece a tabela da normal reduzida, mas mostra a curva normal com as áreas e, portanto, as probabilidades compreendidas entre desvios padrão inteiros. O gráfico da figura a seguir mostra a situação proposta.

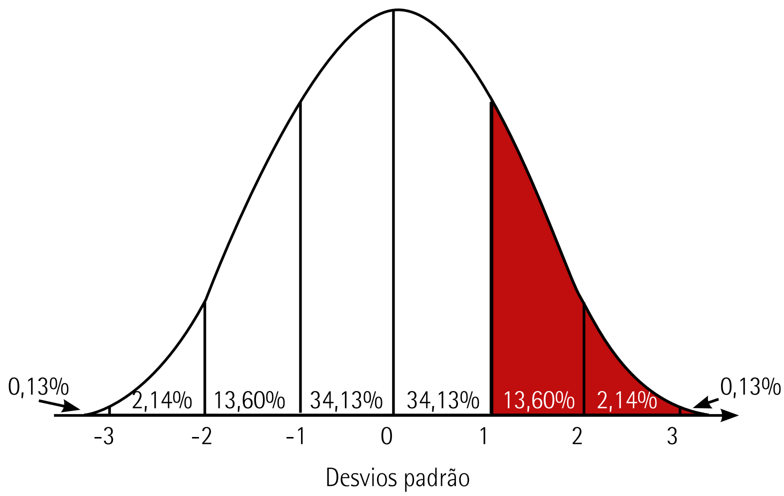



Figura 54

A área preenchida corresponde ao solicitado, ou seja:

13,60% + 2,14% + 0,13% = 15,87%