

Coleção Ciências Contábeis

**MÉTODOS QUANTITATIVOS
APLICADOS ÀS CIÊNCIAS
CONTÁBEIS 2**

Lidia Maria Alves Rodella Amorim



ISBN 978-65-5962-127-9



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO



UNIVERSIDADE FEDERAL DE PERNAMBUCO (UFPE)
SECRETARIA DE PROGRAMAS DE EDUCAÇÃO ABERTA E DIGITAL (SPREAD)
COORDENAÇÃO GERAL UAB/UFPE

Reitor

Alfredo Macedo Gomes

Vice-Reitor

Moacyr Cunha de Araujo Filho

Secretário Geral SPREAD

José Alberto Miranda Poza

Coordenador Geral UAB/UFPE

Francisco Kennedy Silva dos Santos

Coordenador Adjunto UAB/UFPE

André Felipe Vieira da Cunha

Design Instrucional UAB/UFPE | Projeto gráfico

Gabriela Carvalho da Nóbrega

Diagramação

Hanna Kardenya da Silva

Revisão Textual

Daniel Carvalho Cisneiros Silva

Jade Maria Oliveira da Paz

Professora Conteudista Responsável

Lidia Maria Alves Rodella Amorim



Esta licença permite que outros remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito e que licenciem as novas criações sob termos idênticos.

Catálogo na fonte:
Bibliotecária Kalina Ligia França da Silva, CRB4-1408

A524m Amorim, Lidia Maria Alves Rodella.
Métodos quantitativos aplicados às ciências contábeis 2
[recurso eletrônico] / Lidia Maria Alves Rodella Amorim. – Recife :
Ed. UFPE, 2022.
(Coleção Ciências Contábeis).

Inclui referências.
ISBN 978-65-5962-127-9 (online)

1. Contabilidade. 2. Contabilidade – Métodos estatísticos. 3.
Pesquisa quantitativa. 4. Estatística. I. Título. II. Título da coleção.

657 CDD (23.ed.) UFPE (BC2022-055)

Como Coordenador do curso de Ciências Contábeis EAD, é uma satisfação e um privilégio apresentar mais um de nossos e-books, que fazem parte de uma coletânea cujo conteúdo foi inteiramente produzido por nossos professores, o que concede um selo de qualidade e uma identidade da UFPE aos materiais didáticos.

Não posso me furtar em ressaltar que o EAD se constituiu num divisor de águas no contexto educacional, notadamente num país com dimensões continentais como é o Brasil, tendo a sua essencialidade mais que evidenciada com a crise sanitária imposta pela pandemia da COVID-19, que obrigou a humanidade a copiar o modelo do EAD como saída para a continuidade das atividades acadêmicas nos mais diversos níveis da educação.

A interatividade, flexibilidade, praticidade e metodologias inovadoras permitem que o aluno do EAD assuma grande protagonismo na dinâmica de construção do seu conhecimento, sendo-lhe exigida, porém, muita disciplina e foco para o alcance dos seus objetivos. Neste momento, o professor figura como um imprescindível mediador desse processo de aprendizagem.

Mas, para além disso, o estudante precisa ter à sua disposição um material didático confiável e atualizado, que lhe ajudará a consolidar os seus conhecimentos, uma vez que nele o aluno vai encontrar o complemento da fala do professor e terá a chance de praticar o conteúdo aprendido através de exercícios e atividades.

Assim, tenho a convicção de que o discente de Ciências Contábeis EAD da UFPE, encontrará nesta obra todo o arcabouço de conhecimentos que lhe são necessários, para a sua formação nesta disciplina, pois a incomparável expertise de nossos docentes que, além de tecnicamente capacitados, têm a sensibilidade e o conhecimento do perfil dos seus alunos, nos trazem essa segurança.

Eu te desejo uma boa leitura!

Daniel José Cardoso da Silva, Dr.
Coordenador do Curso de Ciências Contábeis EAD da UFPE

Sumário



1	Cálculo integral	7
	<i>Objetivos de aprendizagem</i>	8
	<i>Introdução</i>	8
1.1	Integrais indefinidas	9
1.2	Regras de integração	12
1.3	Condições iniciais	14
1.4	Técnicas de integração	16
1.4.1	Integração por substituição	16
1.4.2	Integração por partes	20
	<i>Revisando</i>	23
	<i>Saiba Mais</i>	24
	<i>Referências</i>	25
2	Integral definida, integrais impróprias e suas aplicações	26
	<i>Objetivos de aprendizagem</i>	27
	<i>Introdução</i>	27
2.1	A área como limite de uma soma	28
2.2	Integral definida (ou integral de Riemann)	32
2.3	Teorema fundamental do cálculo	33
2.4	Propriedades de integrais definidas	38
2.5	Área entre duas curvas	39
2.6	Integrais impróprias	41
2.6.1	Tipo 1: Limites de integração infinitos	41
2.6.2	Tipo 2: Integrandos infinitos	43
2.7	Teorema do valor médio para integrais	45
2.8	Aplicações de integrais definidas	46
	<i>Revisando</i>	50
	<i>Saiba Mais</i>	52

	<i>Referências</i>	53
3	Estatística descritiva	54
	<i>Objetivos de aprendizagem</i>	55
	<i>Introdução</i>	55
3.1	Conceitos básicos de estatística	56
3.2	Distribuições de frequências	59
3.3	Gráficos	62
3.3.1	Gráficos para variáveis qualitativas	63
3.3.2	Gráficos para variáveis quantitativas	64
3.4	Medidas descritivas	67
3.4.1	Medidas de tendência central	67
3.4.2	Medidas de dispersão	69
3.4.3	Forma	71
3.4.4	Medidas de posição	72
3.4.5	Medidas de associação	72
	<i>Revisando</i>	75
	<i>Saiba Mais</i>	76
	<i>Referências</i>	76
4	Probabilidade e variáveis aleatórias	77
	<i>Objetivos de aprendizagem</i>	78
	<i>Introdução</i>	78
4.1	Probabilidade	79
4.2	Variáveis aleatórias	89
4.2.1	Valor médio ou valor esperado (esperança matemática) de uma variável aleatória	94
4.2.2	Variância de uma variável aleatória	95
	<i>Revisando</i>	98
	<i>Saiba Mais</i>	99
	<i>Referências</i>	100

5	Modelos de distribuição de probabilidade e inferência estatística	101
	<i>Objetivos de aprendizagem</i>	102
	<i>Introdução</i>	102
5.1	Modelos de distribuição de probabilidade para variáveis aleatórias discretas	103
5.1.1	Distribuição de Bernoulli	103
5.1.2	Distribuição Binomial	104
5.1.3	Distribuição Hipergeométrica	106
5.1.4	Distribuição de Poisson	107
5.2	Modelos de distribuição de probabilidade para variáveis aleatórias contínuas	109
5.2.1	Distribuição Normal	109
5.2.2	Distribuição t de Student	115
5.3	Inferência Estatística	117
5.3.1	Estimação por ponto	118
5.3.2	Amostras e distribuições amostrais	120
5.3.2.1	Distribuição amostral de \bar{x}	120
5.3.2.2	Distribuição amostral de \bar{p}	122
5.3.3	Estimação por intervalo	124
5.3.3.1	Média da população: σ conhecido	125
5.3.3.2	Média da população: σ desconhecido	126
5.3.3.3	Tamanho da amostra para intervalo de confiança de uma média populacional	128
5.3.3.4	Proporção da população	128
5.3.4	Testes de hipóteses	129
5.3.4.1	Teste de hipótese sobre a média da população com σ conhecido	131
5.3.4.2	Teste de hipótese sobre a média da população com σ desconhecido	135
	<i>Revisando</i>	138
	<i>Saiba Mais</i>	139
	<i>Referências</i>	140
	Anexo A - Distribuição Normal Padrão	141
	Anexo B - Distribuição t de Student	142

1

Cálculo integral

Métodos Quantitativos Aplicados
às Ciências Contábeis 2

Lidia Maria Alves Rodella Amorim

Objetivos da aprendizagem

- 1 Entender a relação entre o processo de derivação e o processo de integração;
- 2 Aprender as regras e as técnicas de integração;
- 3 Ser capaz de, a partir de uma função marginal de determinada grandeza, encontrar sua função acumulada.

Introdução

Um empresário ciente do custo marginal de sua produção pode querer encontrar a variação total no custo em um determinado intervalo da produção. Um cientista que conhece a taxa segundo a qual o número de curados da covid-19 está crescendo pode querer saber qual é o tamanho da população curada em determinado momento futuro. Ainda, como exemplifica Stewart (2017): um físico que conhece a velocidade de uma partícula pode desejar saber sua posição em um dado instante; um engenheiro que sabe a taxa de variação segundo a qual a água escoada de um tanque pode querer saber a quantidade escoada durante certo período.

Em todas estas situações, a derivada (taxa de variação) de uma determinada grandeza é conhecida e estamos interessados em encontrar o valor da própria grandeza. Este processo é chamado de integração (ou antiderivação).

Neste capítulo vamos aprender as técnicas usadas para determinar as integrais (ou antiderivadas). Depois de obtida uma integral, para verificar se ela está correta, basta derivá-la e confirmar se o resultado é igual à função original (HOFFMANN; BRADLEY, 2008).

1.1 Integrais indefinidas

Em seu livro, Hoffmann e Bradley (2008) fazem as seguintes indagações: Como podemos usar a taxa de inflação para prever os futuros preços? Como podemos usar o conhecimento da taxa de crescimento de uma população para estimar o número futuro de habitantes?

Em cada uma dessas situações o problema é encontrar uma função F cuja derivada é uma função conhecida f . Se a função F existir, ela é denominada integral indefinida (ou primitiva ou antiderivada) de f .

Assim, temos que na derivação $y = F(x) \rightarrow \frac{dy}{dx}$ ou $f(x)$ (taxa de variação de y). Enquanto na integração $\frac{dy}{dx}$ ou $f(x) \rightarrow y = F(x)$.

Portanto, a integração é o processo de obter uma função a partir de sua derivada, ou seja, a integração é o inverso da derivação.

Vejam o seguinte exemplo (STEWART, 2017). Seja $f(x) = x^2$. Pela regra da potência (que veremos detalhadamente mais adiante) não é difícil descobrir uma integral indefinida para tal função: $F(x) = \frac{1}{3}x^3$, logo $F'(x) = x^2 = f(x)$. Mas a função $G(x) = \frac{1}{3}x^3 + 100$ também satisfaz $G'(x) = x^2$. Portanto, F e G são primitivas ou antiderivadas de f . De fato, qualquer função na forma $H(x) = \frac{1}{3}x^3 + C$, onde C é uma constante, é uma integral indefinida (primitiva/antiderivada) de f .

Como vimos acima, no processo de integração, a derivada (taxa de variação) de uma função já é conhecida e o objetivo é encontrar a própria função.

Vejam o exemplo de Chiang e Wainwright (2006): considere que o tamanho de uma população P varia ao longo do tempo à taxa $\frac{dp}{dt} = t^{-\frac{1}{2}}$. Vamos encontrar qual é trajetória temporal da população $P = P(t)$ que pode resultar na referida taxa de variação?

Encontramos $P(t)$ através do método conhecido como integração. Por enquanto, vamos nos satisfazer com a observação de que a função $P(t) = 2t^{1/2}$ tem uma derivada da forma apresentada acima e, portanto, aparentemente pode ser considerada como uma solução para o nosso problema.

A questão é que existem outras funções similares com a mesma derivada: $P(t) = 2t^{1/2} + 10$, $P(t) = 2t^{1/2} + 50$ ou, de modo mais geral, $P(t) = 2t^{1/2} + C$, onde C é a constante de integração (arbitrária).

Dessa forma, não é possível encontrarmos uma função específica, a menos que tenhamos alguma informação adicional que defina o valor da constante C . Essa informação adicional é chamada de condição inicial ou condição de fronteira. Nesse exemplo, se soubermos que a população inicial $P(0) = 100$, então poderemos determinar o valor da constante C : $P(0) = 2 \cdot (0)^2 + C = 100 \therefore C = 100$. Assim, nesse exemplo, o tamanho da população em qualquer ponto do tempo pode ser obtido pela seguinte função: $P(t) = 2t^{1/2} + 100$.

Uma função $F(x)$ será chamada de integral (ou primitiva ou antiderivada) de uma função $f(x)$ se $F'(x) = f(x)$ para todo x no domínio de $f(x)$.

No exemplo destacado por Hoffmann e Bradley (2008), verifique que $F(x) = \frac{x^3}{3} + 5x + 2$ é uma integral de $f(x) = x^2 + 5$. Temos que $F'(x) = x^2 + 5 = f(x)$. Logo, $F(x)$ é “uma” integral de $f(x)$. É “uma” porque existem várias. Veja: se $F(x) = \frac{x^3}{3} + 5x - 8$, então $F'(x) = x^2 + 5 = f(x) \therefore F(x)$ também é uma integral de $f(x)$. De um modo geral, qualquer função G na forma $G(x) = F(x) + C$ é uma integral (antiderivada) de $f(x)$.

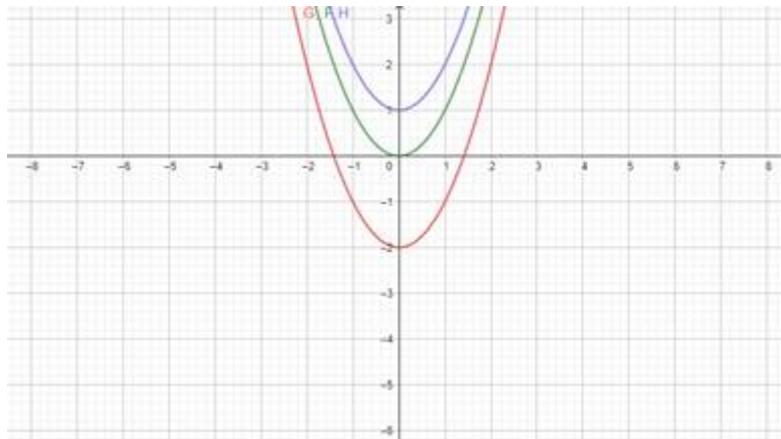
Diante do exposto, os autores destacam a seguinte propriedade: se $F(x)$ é uma integral de uma função contínua $f(x)$, qualquer outra integral de $f(x)$ tem a forma $G(x) = F(x) + C$, onde C é uma constante.

Para entendermos geometricamente, digamos que temos a seguinte função: $f(x) = 2x$. Vamos encontrar algumas integrais dessa função, por exemplo:

- $F(x) = x^2$, em que $F'(x) = 2x = f(x)$;
- $G(x) = x^2 - 2$, tal que $G'(x) = 2x = f(x)$;
- $H(x) = x^2 + 1$, onde $H'(x) = 2x = f(x)$.

Graficamente, teremos a Figura 1, a seguir.

Figura 1 – Gráfico das funções $F(x) = x^2$, $G(x) = x^2 - 2$ e $H(x) = x^2 + 1$



Fonte: Elaborado pela autora.

Observem na Figura 1 que, para um dado x , a inclinação da curva (derivada) é igual para as três funções. Em outras palavras, as curvas $F(x)$, $G(x)$ e $H(x)$ diferem apenas por uma translação (movimento) vertical.

Como Chiang e Wainwright (2006) destacam, precisamos de uma notação especial para representar a integração solicitada de $f(x)$ em relação a x . A notação padrão é:

$$\int f(x)dx = F(x) + C$$

Devemos lê-la da seguinte forma: “integral indefinida (ou integral) de $f(x)$ ”, onde:

- o símbolo \int é chamado de sinal de integração;
- a função $f(x)$ é o integrando;
- dx indica que a integral deve ser calculada em relação a x ;
- C é uma constante de integração e serve para indicar que o integrando tem várias integrais.

Atenção!

A integral indefinida é chamada assim em contraponto à integral definida, porque ela não tem nenhum valor numérico definido. Ela é uma função da variável x .

1.2 Regras de integração¹

Como observa Leithold (1994), a integração é a operação inversa da diferenciação e as regras de integração podem ser obtidas das regras de derivação. Assim sendo, os teoremas a seguir podem ser provados a partir dos teoremas correspondentes da derivação.

a) Regra da constante: $\int k dx = kx + C$, onde k é uma constante.

Prova: $F(x) = kx + C \rightarrow F'(x) = k$. Verifica-se que a derivada da integral é igual ao integrando, como sempre deve ser.

Exemplo: $\int 2 dx = 2x + C$

b) Regra da potência: $\int x^n dx = \frac{x^{n+1}}{n+1} + C$, para $n \neq -1$.

Prova: $F(x) = \frac{x^{n+1}}{n+1} + C \rightarrow F'(x) = \frac{(n+1).x^n}{(n+1)} = x^n$

Exemplo 1: $\int x^3 dx = \frac{x^{3+1}}{3+1} + C = \frac{x^4}{4} + C$

Exemplo 2: $\int \frac{1}{x^2} dx = \int x^{-2} dx = \frac{x^{-2+1}}{-2+1} + C = \frac{x^{-1}}{-1} + C = -\frac{1}{x} + C$

Exemplo 3: $\int \sqrt[3]{x} dx = \int x^{\frac{1}{3}} dx = \frac{x^{\frac{1}{3}+1}}{\frac{1}{3}+1} + C = \frac{x^{\frac{4}{3}}}{\frac{4}{3}} + C = \frac{3}{4} x^{\frac{4}{3}} + C$

c) Regra do logaritmo: $\int \frac{1}{x} dx = \ln|x| + C$, para qualquer $x \neq 0$.

Prova:

- Se $x > 0 \therefore |x| = x: \frac{d \ln x}{dx} = \frac{1}{x}$

¹ As regras a, b, c e d e seus exemplos foram retirados de Leithold (1994).

- Se $x < 0 \therefore |x| = -x: \frac{d \ln -x}{dx} = \frac{1}{-x} \cdot -1 = \frac{1}{x}$

d) Regra da exponencial: $\int e^{kx} dx = \frac{1}{k} e^{kx} + C$, onde k é uma constante $\neq 0$.

Prova: $F(x) = \frac{1}{k} e^{kx} + C \rightarrow F'(x) = \frac{1}{k} e^{kx} \cdot k = e^{kx}$

Exemplo: $\int e^{-3x} dx = \frac{1}{-3} e^{-3x} + C = -\frac{1}{3} e^{-3x} + C$

As próximas regras são para os casos em que temos combinações de funções, também chamadas de regras de operação. Veremos a seguir como elas são apresentadas por Chiang e Wainwright (2006):

e) Integral de uma soma/subtração:

$$\int [f(x) \pm g(x)] dx = \int f(x) dx \pm \int g(x) dx$$

Prova: $\int f(x) dx = F(x)$ e $\int g(x) dx = G(x)$

$$\frac{d}{dx} [F(x) \pm G(x)] = \frac{d}{dx} F(x) \pm \frac{d}{dx} G(x) = f(x) \pm g(x)$$

$$\int [f(x) \pm g(x)] dx = F(x) \pm G(x) = \int f(x) dx \pm \int g(x) dx$$

Exemplo: $\int (x^3 + x + 1) dx = \int x^3 dx + \int x dx + \int dx$

$$\begin{aligned} &= \frac{x^4}{4} + C_1 + \frac{x^2}{2} + C_2 + x + C_3 = \frac{x^4}{4} + \frac{x^2}{2} + x + C, \text{ onde } C \\ &= C_1 + C_2 + C_3 \end{aligned}$$

Refleta

Por que todas as constantes de integração que surgem durante o processo sempre podem ser combinadas em uma única constante arbitrária na resposta final?

f) Integral de um múltiplo:

$$\int kf(x)dx = k \cdot \int f(x)dx, \text{ onde } k = \text{constante}$$

Prova:
$$\int kf(x)dx = \int \underbrace{[f(x) + f(x) + \dots + f(x)]}_{k \text{ termos}} dx =$$

$$= \int f(x)dx + \int f(x)dx + \dots + \int f(x)dx = k \cdot \int f(x)dx$$

Exemplo: $\int 2x^3 dx = 2 \cdot \int x^3 dx = 2 \cdot \frac{x^4}{4} + C = \frac{x^4}{2} + C$

Observação: não existe regra do quociente para integração. Em alguns casos, você poderá resolver integrandos que envolvem quociente da seguinte forma:

$$\begin{aligned} \int \left(\frac{x^3 + 2x - 7}{x} \right) dx &= \int \left(\frac{x^3}{x} + \frac{2x}{x} - \frac{7}{x} \right) dx = \int \left(x^2 + 2 - \frac{7}{x} \right) dx \\ &= \frac{x^3}{3} + 2x - 7 \ln|x| + C \end{aligned}$$

1.3 Condições iniciais

Frequentemente, desejamos encontrar uma função (curva) específica, mas para isso é necessária uma informação adicional que nos permita definir o valor da constante arbitrária (C). Essa informação é chamada de condição inicial ou condição de fronteira, como destaca Leithold (1994). Veja como funciona no exemplo a seguir.

Exemplo 1: Considere que uma função custo marginal é dada por $C'(x) = 4x - 8$, onde $C(x)$ é o custo total da produção de x unidades. Se o custo da produção de 6 unidades é R\$ 40,00, encontre a função custo total.

$$C'(x) = 4x - 8$$

$$C(x) = \int (4x - 8)dx = \frac{4x^2}{2} - 8x + C = 2x^2 - 8x + C$$

Dada a condição inicial $C(6) = 40$, no resultado acima substituiremos $x = 6$ e, igualando o resultado a 40, teremos:

$$C(6) = 2 \cdot (6)^2 - 8 \cdot (6) + C = 40$$

$$C = 40 - 72 + 48 = 88 - 72 = 16$$

$$\therefore C(x) = 2x^2 - 8x + 16$$

Exemplo 2 ²: Um fabricante estima que o custo marginal para produzir q unidades de um certo produto é $C'(q) = 3q^2 - 24q + 48$ reais por unidade. Se o custo para produzir 10 unidades é de R\$ 5.000,00, qual é o custo para produzir 30 unidades?

$$\int (3q^2 - 24q + 48) dq = \int 3q^2 dq - \int 24q dq + \int 48 dq =$$

$$\frac{3q^3}{3} - 24 \cdot \frac{q^2}{2} + 48q + C = q^3 - 12q^2 + 48q + C = C(q) \text{ ou Função custo total}$$

Substituindo, na função encontrada, a condição inicial dada no enunciado:

$$C(10) = (10)^3 - 12 \cdot (10)^2 + 48 \cdot (10) + C = 5000 \rightarrow C = 4720$$

$$\therefore C(q) = q^3 - 12q^2 + 48q + 4720$$

Para encontrarmos o custo para produzir 30 unidades, façamos a substituição de q por 30 na função específica encontrada acima:

$$C(30) = (30)^3 - 12 \cdot (30)^2 + 48 \cdot (30) + 4720 = 22.360$$

E chegaremos ao custo total de R\$ 22.360,00.

² Exemplo retirado de Hoffmann e Bradley (2008).

1.4 Técnicas de integração

Muitas vezes não vamos conseguir calcular as integrais aplicando diretamente as regras que vimos anteriormente. Então, para calculá-las, podemos utilizar algumas técnicas que irão simplificar essas integrais.

1.4.1 Integração por substituição

A ideia é substituir uma expressão em x por uma variável u para transformar a integral original em uma integral mais simples.

Como observam Chiang e Wainwright (2006), essa regra é a contraparte da regra da cadeia no cálculo integral e, portanto, pode ser comprovada por meio da própria regra da cadeia. Dada a função $F(u)$, onde $u = u(x)$, a regra da cadeia determina que:

$$\frac{d}{dx} F(u) = \frac{d}{du} F(u) \cdot \frac{du}{dx} = F'(u) \cdot \frac{du}{dx} = f(u) \cdot \frac{du}{dx}$$

Sendo assim, então:

$$\int f(u) \frac{du}{dx} \cdot dx = F(u) + C$$

Esse resultado também é obtido pelo cancelamento das duas expressões dx à esquerda.

A regra de substituição diz que é permitido operar com dx e du após sinais de integração como se fossem diferenciais. Se $u = f(x)$, então $du = f'(x)dx$.

Veja a seguir o passo a passo para integrar por substituição, baseado em Hoffmann e Bradley (2008):

1. Introduzir a variável u para substituir uma expressão em x com o objetivo de simplificar uma integral;

2. Escrever a integral em termos de u . Para isso, calcule $\frac{du}{dx}$,
3. Calcular a integral resultante e substituir u por sua expressão em termos de x para obter a solução.

Observe o seguinte exemplo apresentado por esses autores: $\int 2x(x^2 + 1)dx$.

$$u = x^2 + 1 \rightarrow \frac{du}{dx} = 2x \text{ ou } du = 2x dx$$

$$\therefore \int 2x(x^2 + 1)dx = \int u du = \frac{u^2}{2} + C_1 = \frac{(x^2 + 1)^2}{2} + C_1 = \frac{x^4}{2} + \frac{2x^2}{2} + \frac{1}{2} + C_1$$

$$= \frac{x^4}{2} + x^2 + C$$

Ou, nesse caso, ela poderia ser resolvida multiplicando o integrando:

$$\int 2x(x^2 + 1)dx = \int (2x^3 + 2x)dx = \frac{2x^4}{4} + \frac{2x^2}{2} + C =$$

$$= \frac{x^4}{2} + x^2 + C$$

Contudo, nem sempre é fácil simplificar o integrando por multiplicação. Nesses casos, a regra de substituição é a melhor opção.

Veja ainda outro exemplo apresentado por Hoffmann e Bradley (2008). Observe que após calcular $\frac{du}{dx}$, o resultado foi multiplicado por $\frac{2}{2}$ (que é igual a 1 e não altera o resultado) na busca do termo em u que iria substituir $6x^2 dx$ na integral requerida.

$$\int 6x^2(x^3 + 2)^{99} dx$$

$$u = x^3 + 2 \rightarrow \frac{du}{dx} = 3x^2 = \frac{6x^2}{2} \rightarrow 2du = 6x^2 dx$$

Substituindo, temos:

$$\int 6x^2(x^3 + 2)^{99} dx = \int 2u^{99} du = 2 \cdot \frac{u^{100}}{100} + C = \frac{1}{50} (x^3 + 2)^{100} + C$$

A partir desses exemplos, é possível observar que essa regra é útil sempre que um dos fatores for a derivada do outro fator, ou ainda quando ele for a derivada multiplicada por uma constante.

Se o integrando é um produto ou quociente de dois termos e um dos termos é o produto de uma constante pela derivada de uma expressão que aparece no outro termo, essa expressão é provavelmente uma boa escolha para u . Veja no exemplo a seguir:

$$\int \frac{3x + 6}{\sqrt{2x^2 + 8x + 8}} dx$$

$$u = 2x^2 + 8x + 8 \rightarrow \frac{du}{dx} = 4x + 8 \rightarrow du = 4(x + 2)dx = \frac{4}{3}(3x + 6)dx$$

$$\therefore \frac{3}{4} du = (3x + 6)dx$$

Atente para a multiplicação por $\frac{3}{4}$ na busca pela expressão em u para substituir o integrando adequadamente.

$$\text{Substituindo: } \int \frac{3x+6}{\sqrt{2x^2+8x+8}} dx = \int \frac{3}{4} \cdot \frac{1}{\sqrt{u}} du = \frac{3}{4} \int u^{-1/2} du = \frac{3}{4} \cdot \frac{u^{1/2}}{1/2} + C$$

$$= \frac{3}{4} \cdot 2u^{1/2} + C = \frac{3}{2} \cdot \sqrt{2x^2 + 8x + 8} + C$$

Para verificar se a integral foi calculada corretamente basta derivar a função encontrada e comparar com o integrando, não é mesmo?

Vejam mais um exemplo:

$$\int x^3 \cdot e^{x^4+2} dx$$

$$u = x^4 + 2 \rightarrow \frac{du}{dx} = 4x^3 \rightarrow du = 4x^3 dx \rightarrow \frac{du}{4} = x^3 dx$$

$$\text{Substituindo, } \int x^3 \cdot e^{x^4+2} dx = \int \frac{1}{4} e^u du = \frac{1}{4} \cdot e^u + C = \frac{1}{4} \cdot e^{x^4+2} + C$$

Hoffmann e Bradley (2008) trazem algumas regrinhas que podem ajudar na escolha de $u(x)$:

1. Sempre que possível, escolha u de tal forma que a sua deriva $u'(x)$ seja parte da função que está sendo integrada, isto é, de $f(x)$;
2. Busque definir o seu u como a parte da função $f(x)$ a ser integrada que a torna difícil de integrar diretamente, ou seja, quando x aparece como radicando, ou em um denominador ou, ainda, como um expoente;
3. Não exagere nas substituições, pois a integral transformada pode ficar mais difícil de resolver que a original;
4. Caso a substituição que você fez não tenha resultado em uma integral fácil de calcular, tente uma substituição diferente.

Mas nem sempre a substituição funciona. Essa regra não será útil se um dos fatores for o produto de uma variável pela derivada de uma expressão que aparece no outro termo. Por exemplo:

$$\int x \cdot e^{2x} dx$$

$$u = 2x \rightarrow x = \frac{u}{2} \rightarrow \frac{du}{dx} = 2 \rightarrow \frac{1}{2} du = dx$$

Multiplicando os dois lados por x e, em seguida, substituindo x do lado esquerdo por $\frac{u}{2}$, teremos:

$$\frac{x}{2} du = x dx \rightarrow \frac{u}{4} du = x dx$$

$$\therefore \int x \cdot e^{2x} dx = \int \frac{u}{4} e^u du$$

Verifique que a substituição não simplificou o integrando. Nesses casos, podemos recorrer a outra técnica de integração.

1.4.2 Integração por partes

Essa técnica pode ser usada para integrar produtos da forma $f(x).g(x)$ nos quais um dos fatores, $g(x)$, pode ser facilmente **integrado** e o outro, $f(x)$, torna-se mais simples ao ser **derivado**.

Stewart (2017) apresenta a seguinte fórmula para a integração por partes, oriunda da regra do produto aplicada no cálculo de derivadas:

$$\int f(x).g(x)dx = f(x).G(x) - \int f'(x).G(x)dx$$

Na fórmula acima, $G(x)$ é uma integral de g .

Há ainda outra notação também apresentada pelo autor. Sejam $u = f(x)$ e $v = G(x)$. Então as diferenciais são $du = f'(x)dx$ e $dv = g(x)dx$ e, assim, pela regra da substituição, a fórmula para a integração por partes se torna:

$$\int u dv = u.v - \int v du$$

Prova: Partindo da regra da derivada do produto de duas funções temos:

$$\frac{d[f(x).G(x)]}{dx} = f'(x).G(x) + f(x).G'(x) = f'(x).G(x) + f(x).g(x)$$

Então:

$$\begin{aligned} \int [f'(x).G(x) + f(x).g(x)]dx &= f(x).G(x) \\ \int f'(x).G(x)dx + \int f(x).g(x)dx &= f(x).G(x) \\ \int f(x).g(x)dx &= f(x).G(x) - \int f'(x).G(x)dx \end{aligned}$$

Vamos agora tentar resolver o último exemplo da subseção anterior aplicando a integração por partes.

$$\int x \cdot e^{2x} dx$$

$$f(x) = x \rightarrow f'(x) = 1$$

$$g(x) = e^{2x} \rightarrow G(x) = \int e^{2x} dx = \frac{1}{2} e^{2x}$$

Temos então:

$$\int x e^{2x} dx = x \cdot \frac{1}{2} e^{2x} - \int 1 \cdot \frac{1}{2} e^{2x} dx = \frac{1}{2} \cdot x e^{2x} - \frac{1}{2} \int e^{2x} dx =$$

$$= \frac{1}{2} x e^{2x} - \frac{1}{2} \cdot \frac{1}{2} e^{2x} + C = \frac{1}{2} e^{2x} \cdot \left(x - \frac{1}{2} \right) + C$$

Para esse mesmo exemplo, usando a notação $\int u dv = u \cdot v - \int v du$, teríamos:

$$u = x \rightarrow du = dx \rightarrow \frac{du}{dx} = 1$$

$$dv = e^{2x} dx \rightarrow \int dv = \int e^{2x} dx \rightarrow v = \frac{1}{2} e^{2x} dx$$

$$\text{Portanto: } \int x e^{2x} dx = x \cdot \frac{1}{2} e^{2x} - \int \frac{1}{2} e^{2x} dx = \frac{1}{2} e^{2x} \cdot \left(x - \frac{1}{2} \right) + C = \frac{x e^{2x}}{2} - \frac{e^{2x}}{4} + C$$

Vejamos também o exemplo a seguir, resolvido por Chiang e Wainwright (2006).

$$\int x(x+1)^{1/2} dx$$

$$f(x) = x \rightarrow f'(x) = 1$$

$$g(x) = (x+1)^{1/2} \rightarrow G(x) = \int (x+1)^{1/2} dx$$

Resolvendo a integral de $G(x)$ por substituição:

$$u = x + 1 \rightarrow \frac{du}{dx} = 1 \rightarrow du = dx$$

$$G(x) = \int (x + 1)^{1/2} dx = \int u^{1/2} du = \frac{u^{3/2}}{3/2} = \frac{2}{3} u^{3/2} = \frac{2}{3} (x + 1)^{3/2}$$

Aplicando a fórmula da integração por partes:

$$\begin{aligned} \int x(x + 1)^{1/2} dx &= x \cdot \frac{2}{3} (x + 1)^{3/2} - \int 1 \cdot \frac{2}{3} (x + 1)^{3/2} dx \\ &= x \cdot \frac{2}{3} (x + 1)^{3/2} - \frac{2}{3} \cdot \frac{(x + 1)^{5/2}}{5/2} + C = \frac{2}{3} x(x + 1)^{3/2} - \frac{4}{15} (x + 1)^{5/2} + C \end{aligned}$$

Vamos agora calcular a $\int x^2 e^{2x} dx$. Para isso, precisaremos aplicar duas vezes a integração por partes. Considerando:

$$f(x) = x^2 \rightarrow f'(x) = 2x$$

$$g(x) = e^{2x} \rightarrow G(x) = \int e^{2x} dx = \frac{e^{2x}}{2}$$

E em seguida aplicando a integração por partes:

$$\int x^2 e^{2x} dx = x^2 \cdot \frac{e^{2x}}{2} - \int 2x \cdot \frac{e^{2x}}{2} dx = x^2 \cdot \frac{e^{2x}}{2} - \int x e^{2x} dx$$

Observe que resolvemos no início desta subseção a integral $\int x e^{2x} dx$ aplicando a mesma técnica de integração por partes e encontramos como resultado $\frac{x e^{2x}}{2} - \frac{e^{2x}}{4} + C$. Ele será utilizado substituindo o último termo da resultante da aplicação dessa técnica no presente exemplo e teremos:

$$\int x^2 e^{2x} dx = x^2 \cdot \frac{e^{2x}}{2} - \frac{x e^{2x}}{2} + \frac{e^{2x}}{4} + C = \frac{e^{2x}}{2} \cdot \left(x^2 - x + \frac{1}{2} \right) + C$$

Sobre o cálculo de integrais é importante estar ciente de que:

“ De fato, não existe nenhuma fórmula geral que dá a integral de um produto de duas funções em termos das integrais isoladas dessas funções; e também não temos uma fórmula geral que nos dê a integral de um quociente de duas funções em termos de suas integrais isoladas. E é aqui que está a razão porque a integração, no todo, é mais difícil do que a diferenciação e, porque, quando os integrandos são complicados, é mais conveniente consultar a resposta em tabelas de fórmulas de integração já preparadas, em vez de nós mesmos fazermos a integração.” (CHIANG; WAINWRIGHT, 2006, p. 433).

Revisando

- Integral indefinida, antiderivada, primitiva:

$$\int f(x)dx = F(x) + C \text{ se e somente se } F'(x) = f(x)$$

- Propriedade:

Se $F(x)$ é uma integral de uma função contínua $f(x)$, qualquer outra integral de $f(x)$ tem a forma $G(x) = F(x) + C$, onde C é uma constante.

- Regras de integração:

a) Regra da constante: $\int kdx = kx + C$, onde k é uma constante.

b) Regra da potência: $\int x^n dx = \frac{x^{n+1}}{n+1} + C$, para $n \neq -1$.

c) Regra do logaritmo: $\int \frac{1}{x} dx = \ln|x| + C$, para qualquer $x \neq 0$.

d) Regra da exponencial: $\int e^{kx} dx = \frac{1}{k} e^{kx} + C$, onde k é uma constante $\neq 0$.

e) Integral de uma soma/subtração: $\int [f(x) \pm g(x)] dx = \int f(x) dx \pm \int g(x) dx$.

f) Integral de um múltiplo: $\int kf(x) dx = k \int f(x) dx$, onde $k = \text{constante}$.

- Integração por substituição: Passo a passo, conforme Hoffmann e Bradley (2008): 1. Introduzir a variável u para substituir uma expressão em

x com o objetivo de simplificar uma integral; 2. Escrever a integral em termos de u . Para isso, calcule $\frac{du}{dx}$; e 3. Calcular a integral resultante e substituir u por sua expressão em termos de x para obter a solução.

- Integração por partes: Usada para integrar produtos da forma $f(x) \cdot g(x)$ nos quais um dos fatores, $g(x)$, pode ser facilmente **integrado** e o outro, $f(x)$, se torna mais simples ao ser **derivado**, através da fórmula:

$$\int f(x) \cdot g(x) dx = f(x) \cdot G(x) - \int f'(x) \cdot G(x) dx$$

Saiba mais

Capítulo de livro: **Integração**, de Laurence D. Hoffmann e Gerald L. Bradley

HOFFMANN, Laurence D.; BRADLEY, Gerald L. Integração. In: HOFFMANN, Laurence D.; BRADLEY, Gerald L. **Cálculo**: um curso moderno e suas aplicações. 9. ed. Rio de Janeiro: Livro Técnico e

Material de apoio: **Introdução ao estudo das integrais**, de Patrícia Pugliesi Carneiro

CARNEIRO, Patrícia Pugliesi. **Introdução ao estudo das integrais**. São Cristóvão, SE: Departamento de Economia da Universidade Federal de Sergipe, 2014. Disponível em: http://dee.ufs.br/uploads/page_attach/path/2261/INTEGRAIS - Material Did tico de Apoio_2 .pdf. Acesso em: 05 out. 2020.

Vídeo: **Integral por partes**, de Paulo Pereira

INTEGRAL por partes. Publicado pelo canal Equaciona com Paulo Pereira.
Disponível em: <https://youtu.be/E3ZILV7ER54>. Acesso em: 05 out. 2020.

Referências

CHIANG, Alpha C.; WAINWRIGHT, Kevin. **Matemática para economistas**. 4. ed. Rio de Janeiro: Elsevier, 2006.

HOFFMANN, Laurence D.; BRADLEY, Gerald L. **Cálculo**: um curso moderno e suas aplicações. 9. ed. Rio de Janeiro: Livro Técnico e Científico, 2008.

LEITHOLD, Louis. **O cálculo com geometria analítica**. 3. ed. São Paulo: Harbra, 1994.

STEWART, James. **Cálculo, v. 1**. 8. ed. São Paulo: Cengage Learning, 2017.

2

Integral definida, integrais impróprias e suas aplicações

Métodos Quantitativos Aplicados
às Ciências Contábeis 2

Lidia Maria Alves Rodella Amorim

Objetivos da aprendizagem

- 1 Compreender o conceito de integral definida como limite de uma soma e o seu uso para o cálculo de áreas;
- 2 Relacionar a integral definida ao processo de integração através do teorema fundamental do cálculo.

Introdução

Suponha que você esteja interessado em vender um terreno com 90 metros de largura, limitado por ruas em três lados e por um rio no quarto lado. Você foi informado que o rio pode ser descrito pela curva $y = x^3 + 1$, onde x e y são medidos em centenas de metros. Se a área do terreno é A metros quadrados e é sabido que o metro quadrado na região vale R\$ 15,00, o valor do terreno é $15A$ reais. Mas como calcular a área A do terreno e, assim, saber quanto ele vale no mercado, uma vez que o limite superior do terreno é curvo (dado por y) e, portanto, não seria possível calcular a área A usando uma fórmula da geometria?

Neste capítulo, vamos mostrar que a área sob uma curva, como a área A do nosso exemplo, pode ser calculada como o limite de uma soma de termos que recebe o nome de integral definida. Em seguida, vamos apresentar o teorema fundamental do cálculo, que permite calcular integrais definidas a partir de integrais indefinidas (Capítulo 1). Veremos algumas propriedades de integrais definidas e como usá-las para encontrar a área entre duas curvas. Por fim, serão apresentadas as integrais impróprias, o teorema do valor médio para integrais e algumas aplicações práticas das integrais definidas.

2.1 A área como limite de uma soma

Considerando uma função constante $f(x)$, como calcular a área sob o gráfico da função limitada pela curva e pelo eixo x entre a e b ?

Figura 1 – Gráfico da função $f(x)$ constante



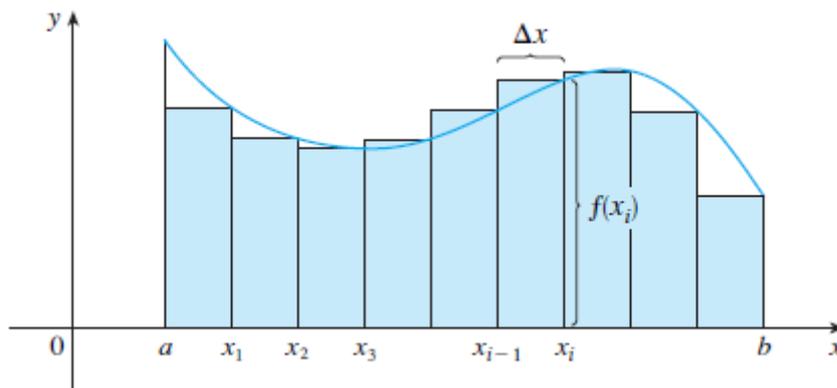
Fonte: Elaborado pela autora.

Na Figura 1, temos o gráfico de uma função $f(x)$ constante em que a área (A) sob a reta que representa essa função no intervalo a e b forma um retângulo. Portanto, é possível calcular a área A como visto em geometria, da seguinte forma:

$$A = (b - a) \cdot f(a)$$

Se tivéssemos um triângulo, um quadrado ou um trapézio, ainda seria fácil encontrarmos a área. Mas se tivermos uma função cujo gráfico é representado abaixo? Como calcular a área sob essa curva e entre dois valores de x ?

Figura 2 – Aproximação da área sob uma curva por retângulos



Fonte: Stewart (2017, p. 330).

Vamos analisar a área da região sob a curva $f(x)$ em um intervalo $a \leq x \leq b$, onde $f(x) \geq 0$ e f é contínua, como na Figura 2. Neste caso, podemos não saber como calcular a área sob a curva dada, mas sabemos como calcular a área de um retângulo. Hoffmann e Bradley (2008) dizem que, assim, podemos dividir a região dada em uma série de regiões retangulares e calculamos o valor aproximado da área A sob a curva $y = f(x)$ somando as áreas dessas regiões retangulares.

De acordo com os autores, inicialmente, dividimos o intervalo $a \leq x \leq b$ em n subintervalos iguais de largura $\Delta x = \frac{b-a}{n}$, e chamamos de x_j a extremidade esquerda do intervalo de ordem j , para $j = 1, 2, \dots, n$. Em seguida, traçamos n retângulos tais que o retângulo de ordem j tenha uma largura igual a Δx e uma altura igual a $f(x_j)$, como na Figura 2. A soma das áreas dos n retângulos, $A = \Delta x \cdot f(x_j)$, é

$$S_n = \Delta x \cdot f x_1 + \Delta x \cdot f x_2 + \dots + \Delta x \cdot f x_n = \Delta x \cdot [f x_1 + f x_2 + \dots + f x_n],$$

conhecida como soma de Riemann, que é aproximadamente igual à área total A sob a curva dada.

Observe na Figura 2 que essa aproximação vai existir porque, para alguns retângulos, estamos superestimando a área sob a curva e, para outros retângulos, estamos subestimando essa área. No entanto, quanto maior é o número n de subintervalos, mais a soma S_n se aproxima de A .

Você sabia?

A definição de integral que usamos se deve a Bernhard Riemann. Ele realizou seu doutorado sob a orientação do legendário Gauss na Universidade de Göttingen, e lá permaneceu para ensinar (STEWART, 2017).

Dessa forma, Hoffmann e Bradley (2008) afirmam o seguinte: considerando $f(x)$ uma função contínua tal que $f(x) \geq 0$ no intervalo $a \leq x \leq b$, a área sob a curva $y = f(x)$ no intervalo $a \leq x \leq b$ é dada por:

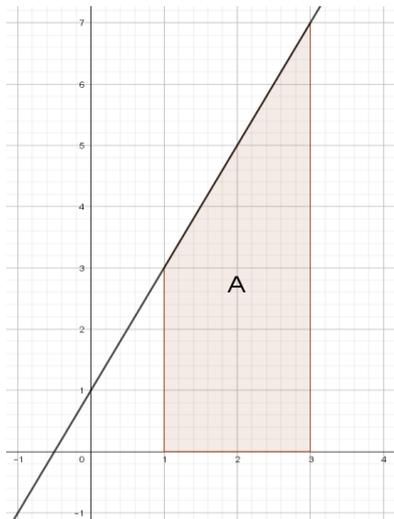
$$A = \lim_{n \rightarrow \infty} \Delta x \cdot [f x_1 + f x_2 + \dots + f x_n],$$

onde x_j é a extremidade esquerda do subintervalo de ordem j se o intervalo $a \leq x \leq b$ for dividido em partes iguais de comprimento $\Delta x = \frac{b-a}{n}$.

Vejamos o exemplo abaixo, retirado de Hoffmann e Bradley (2008)

Exemplo 1: Calcule a área sob a curva da função $f(x) = 2x + 1$ no intervalo $1 \leq x \leq 3$.

Figura 3 – Gráfico do exemplo 1



Fonte: Adaptado de Hoffmann e Bradley (2008, p. 322).

Nesse exemplo, a área A pode ser calculada pela forma do cálculo da área do trapézio (Figura 3) e, assim, teremos:

$$A = \frac{(B + b) \cdot h}{2} = \frac{(7 + 3) \cdot 2}{2} = 10$$

Mas vamos agora calcular essa área como a soma da área de n retângulos, para $n = 6$. Iniciamos calculando a base dos retângulos, Δx :

$$\Delta x = \frac{b - a}{n} = \frac{3 - 1}{6} = \frac{2}{6} = \frac{1}{3}$$

As extremidades dos subintervalos, que variam $\Delta x = \frac{1}{3}$ a partir de $x_1 = 1$, são:
 $x_1 = 1$; $x_2 = \frac{4}{3}$; $x_3 = \frac{5}{3}$; $x_4 = 2$; $x_5 = \frac{7}{3}$; $x_6 = \frac{8}{3}$. Já a altura dos retângulos será:
 $f(x_1 = 1) = 3$; $f(x_2 = \frac{4}{3}) = \frac{11}{3}$; $f(x_3 = \frac{5}{3}) = \frac{13}{3}$; $f(x_4 = 2) = 5$; $f(x_5 = \frac{7}{3}) = \frac{17}{3}$; $f(x_6 = \frac{8}{3}) = \frac{19}{3}$.

Assim, obteremos a área aproximada da região A por meio da soma:

$$S_n = \Delta x. [fx_1 + fx_2 + \dots + fx_6] = \frac{1}{3} \cdot \left[3 + \frac{11}{3} + \frac{13}{3} + 5 + \frac{17}{3} + \frac{19}{3} \right] = \frac{28}{3} \cong 9,333$$

No entanto, quanto maior for o n , mais próxima a soma estará da área exata A da região. Por exemplo, se $n = 100$, teremos $S_n = 9,96$. Já para $n = 500$, $S_n = 9,99$ e $\lim_{n \rightarrow \infty} S_n = 10$, que é o valor exato da área do trapézio.

2.2 Integral definida (ou integral de Riemann)

Como apresentam Chiang e Wainwright (2006), a integral definida é uma forma abreviada para a expressão de limite de uma soma, destacada a seguir:

$$\int_a^b f(x)dx = \lim_{n \rightarrow \infty} = \Delta x. [fx_1 + fx_2 + \dots + fx_n],$$

onde:

- $\Delta x = \frac{b-a}{n}$;
- $f(x)$ é o integrando;
- a é o limite inferior de integração;
- b é o limite superior de integração.

Então, a integral definida de $f(x)$ de a até b possui um valor numérico como resultado, diferentemente da integral indefinida, uma vez que representa a área sob a curva de $f(x)$ de a até b .

Portanto, usando a notação de integral definida, a área sob uma curva é:

$$A = \int_a^b f(x)dx$$

2.3 Teorema fundamental do cálculo

Como calcular o valor de uma integral definida sem calcular o limite de uma soma?

Hoffmann e Bradley (2008) afirmam que, se $f(x)$ é contínua no intervalo $a \leq x \leq b$, então:

$$\int_a^b f(x)dx = F(b) - F(a),$$

onde $F(x)$ é a integral (antiderivada) de $f(x)$ no intervalo $a \leq x \leq b$.

Atenção!

A notação utilizada para o cálculo da integral definida é:

$$\int_a^b f(x)dx = F(x)\Big|_a^b = F(b) - F(a).$$

Agora vamos resolver o **Exemplo 1**¹ utilizando o teorema fundamental do cálculo. Primeiro, vamos encontrar a integral indefinida e, depois, vamos substituir x pelos limites de integração e subtrair os resultados:

$$\begin{aligned} \int_1^3 (2x + 1)dx &= \left(\frac{2x^2}{2} + x + C\right)\Big|_1^3 = (x^2 + x + C)\Big|_1^3 \\ &= [3^2 + 3 + C] - [1^2 + 1 + C] = 12 + C - 2 - C = 10. \end{aligned}$$

Observe que a constante C não aparece no resultado final. Portanto, a constante C é omitida em todos os cálculos de integrais definidas.

¹ Calcule a área sob a curva da função $f(x) = 2x + 1$ no intervalo $1 \leq x \leq 3$.

Veremos agora outro exemplo dado por Hoffmann e Bradley (2008), dessa vez utilizando a integração por substituição.

Exemplo 2: $\int_0^1 8x(x^2 + 1)^3 dx$.

$$u = x^2 + 1 \rightarrow \frac{du}{dx} = 2x \rightarrow 4du$$

Como os limites de integração 0 e 1 se referem à variável x e não a u , existem duas formas alternativas de resolução:

a) Expressar a integral em termos de x ;

$$\int 8x(x^2 + 1)^3 dx = \int 4u^3 du = \frac{4u^4}{4} = u^4 = (x^2 + 1)^4$$

$$\int_0^1 8x(x^2 + 1)^3 dx = (x^2 + 1)^4 \Big|_0^1 = (1^2 + 1)^4 - (0^2 + 1)^4 = 16 - 1 = 15 \text{ ou}$$

b) Determinar os valores de u que correspondem a $x = 0$ e $x = 1$.

$$u(0) = 0^2 + 1 = 1 \text{ e } u(1) = 1^2 + 1 = 2$$

$$\int_0^1 8x(x^2 + 1)^3 dx = \int_1^2 4u^3 du = u^4 \Big|_1^2 = 2^4 - 1^4 = 16 - 1 = 15$$

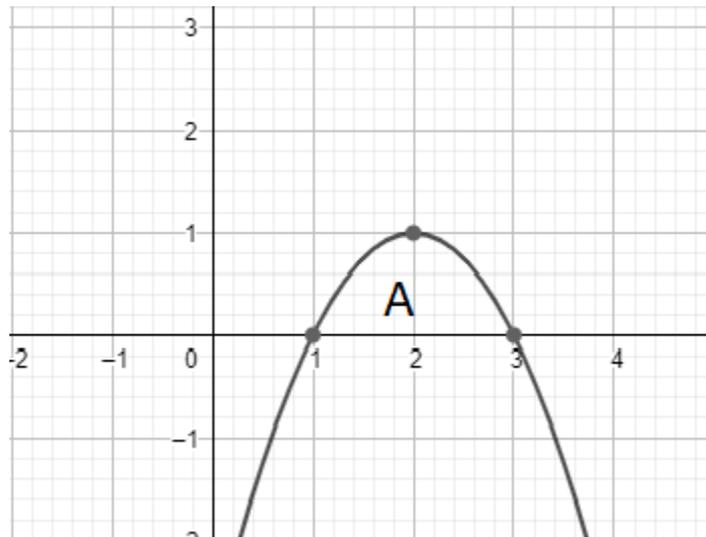
Exemplo 3: Determine a área limitada pela curva $y = -x^2 + 4x - 3$ e pelo eixo x .

Vamos começar encontrando as raízes de x (ou seja, os valores de x que fazem $y = 0$) para, em seguida, construirmos o gráfico da função (Figura 4).

$$\Delta = b^2 - 4ac = 16 - 12 = 4$$

$$x = \frac{-b \pm \sqrt{\Delta}}{2a} = \frac{-4 \pm \sqrt{4}}{-2} \rightarrow x_1 = 1 \rightarrow x_2 = 3$$

Figura 4 – Gráfico da função $y = -x^2 + 4x - 3$



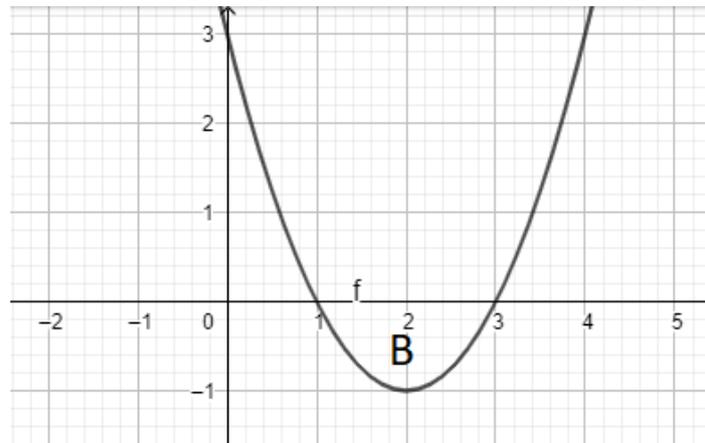
Fonte: Elaborado pela autora.

Observe que, no intervalo $[1, 3]$, temos $f(x) > 0$ e, portanto, a integral é positiva.

$$\begin{aligned}
 A &= \int_1^3 (-x^2 + 4x - 3)dx = \left(\frac{-x^3}{3} + 2x^2 - 3x \right) \Big|_1^3 \\
 &= (-9 + 18 + 9) - \left(-\frac{1}{3} + 2 - 3 \right) = 0 - \left(-\frac{4}{3} \right) = \frac{4}{3}
 \end{aligned}$$

Contudo, se invertermos o sinal da função y , teremos $y = x^2 - 4x + 3$ e, no intervalo $[1, 3]$, teremos $f(x) < 0$. Assim, conseqüentemente, a integral será negativa. Veja a Figura 5:

Figura 5 – Gráfico da função $y = x^2 - 4x + 3$



Fonte: Elaborado pela autora.

$$B = \int_1^3 (x^2 - 4x + 3)dx = -\frac{4}{3}$$

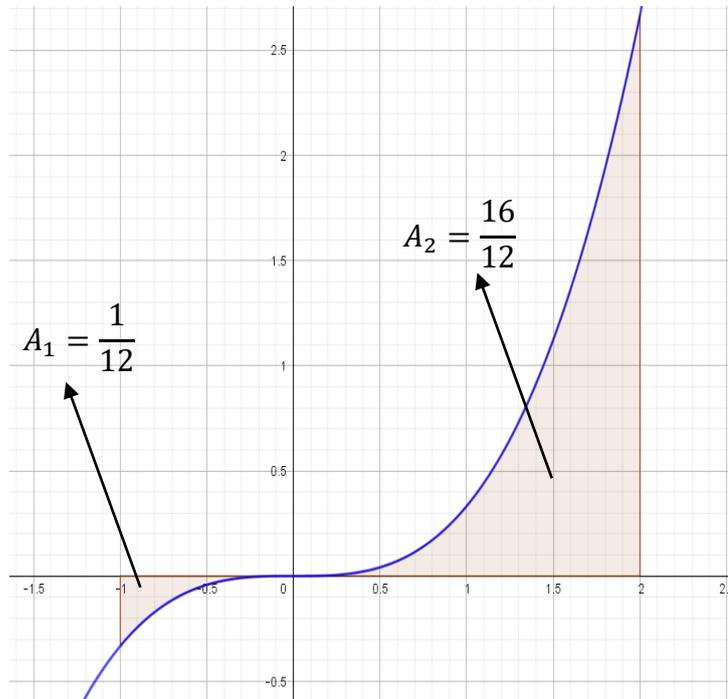
É importante lembrar que não existe área negativa, não é mesmo? O sinal negativo que apareceu no resultado da função y com sinal invertido indica que a área fica abaixo do eixo x e acima de uma curva dada (Figura 5). A resposta é negativa porque a altura de cada retângulo envolvido nessa área é negativa (HOFFMANN; BRADLEY, 2008). Portanto, se estivermos interessados no valor numérico, e não no valor algébrico de tal área, devemos tomar o valor absoluto da integral definida.

Temos, portanto, que se $f(x) \geq 0 \forall x \in [a, b] \therefore \int_a^b f(x)dx \geq 0$.

Por isso, quando formos calcular a área sob uma curva entre dois valores, precisamos saber se naquele intervalo $f(x)$ se mantém o mesmo sinal (positivo ou negativo) ou se ocorre alguma alteração de sinal. Vejamos no exemplo a seguir como proceder nesses casos.

Exemplo 4: Calcule a área A sob o gráfico da função $f(x) = \frac{1}{3}x^3$ (Figura 6) entre $x = -1$ e $x = 2$.

Figura 6 – Gráfico da função $f(x) = \frac{1}{3}x^3$



Fonte: Elaborado pela autora.

Observe que $\int_{-1}^2 \frac{1}{3}x^3 dx \neq A_1 + A_2 = A$. Portanto, para encontrar a área $A = A_1 + A_2$, vamos precisar utilizar o módulo da função $f(x)$:

$$|f(x)| = \begin{cases} f(x), & \text{se } f(x) > 0 \\ -f(x), & \text{se } f(x) < 0 \end{cases}$$

Assim:

$$\begin{aligned} A &= \int_{-1}^2 |f(x)| dx = \int_{-1}^0 -f(x) dx + \int_0^2 f(x) dx = \\ &= - \int_{-1}^0 f(x) dx + \int_0^2 f(x) dx \end{aligned}$$

$$\begin{aligned}
 &= -\int_{-1}^0 \frac{1}{3}x^3 dx + \int_0^2 \frac{1}{3}x^3 dx = -\frac{1}{3} \cdot \frac{x^4}{4} \Big|_{-1}^0 + \frac{1}{3} \cdot \frac{x^4}{4} \Big|_0^2 = \\
 &= \left[0 - \left(-\frac{1}{12} \right) \right] + \left[\frac{16}{12} - 0 \right] = \frac{1}{12} + \frac{16}{12} = \frac{17}{12}
 \end{aligned}$$

2.4 Propriedades de integrais definidas

Veremos a seguir, conforme Chiang e Wainwright (2006), algumas propriedades de integrais definidas que podem ajudar a simplificar o seu cálculo.

A permuta de limites de integração muda o sinal da integral definida.

$$i. \quad \int_b^a f(x) dx = -\int_a^b f(x) dx$$

Prova:

$$\int_b^a f(x) dx = F(a) - F(b) = -[F(b) - F(a)] = -\int_a^b f(x) dx$$

ii. Uma integral definida pode ser expressa como uma soma de um número finito de subintegrais. Essa propriedade às vezes é descrita como aditividade.

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx \quad (a < b < c)$$

Apesar de b estar incluído nas duas áreas, não temos uma contagem dupla porque $\int_b^b f(x) dx = 0$.

As próximas propriedades são as mesmas das integrais indefinidas.

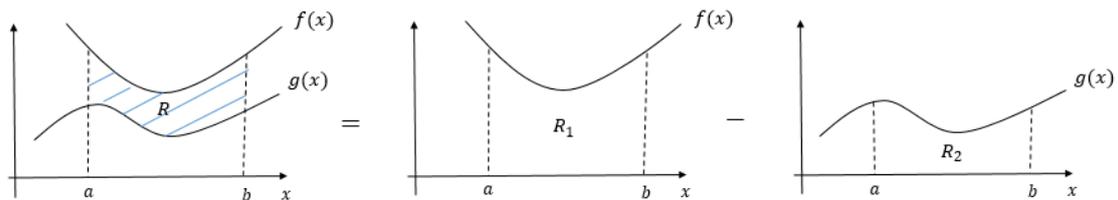
$$iii. \quad \int_a^b kf(x) dx = k \cdot \int_a^b f(x) dx$$

$$iv. \quad \int_a^b [f(x) + g(x)] dx = \int_a^b f(x) dx + \int_a^b g(x) dx$$

2.5 Área entre duas curvas

Podemos utilizar a integral definida para encontrar a área entre duas curvas.

Figura 7 – Área entre curvas



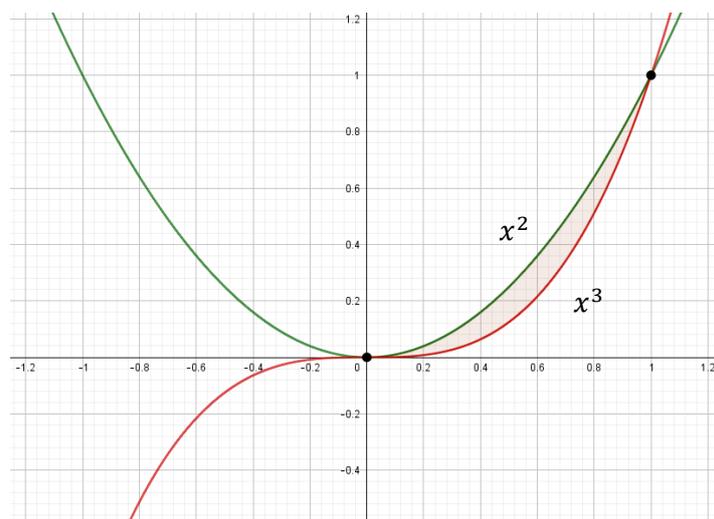
Fonte: Adaptado de Hoffmann e Bradley (2008, p. 334).

A partir da Figura 7, podemos verificar que a área de $R = \int_a^b f(x)dx - \int_a^b g(x)dx$. Ou, ainda, a área de $R = \int_a^b [f(x) - g(x)]dx$, onde $f(x) \geq g(x)$.

Observe os dois exemplos a seguir.

Exemplo 1: Determine a área da região limitada pelas curvas $y = x^3$ e $y = x^2$.

Figura 8 – Gráfico das funções do exemplo 1



Fonte: Adaptado de Hoffmann e Bradley (2008, p. 335).

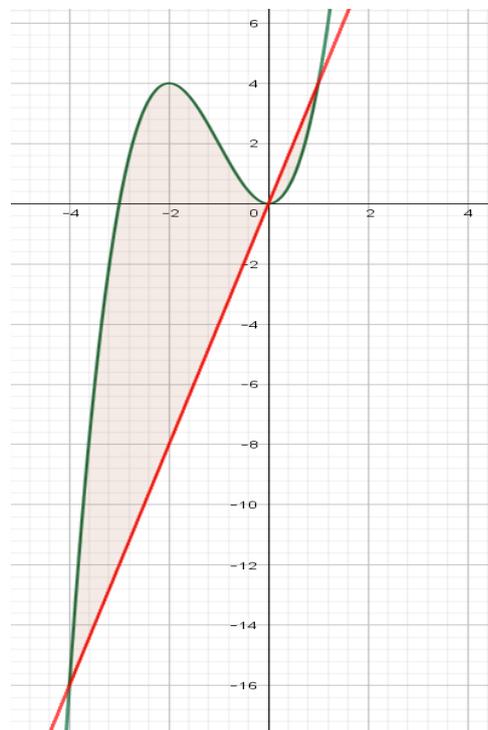
Observe que a função $y = x^2$ está acima da função $y = x^3$ no intervalo $[0, 1]$.

Assim:

$$\int_0^1 (x^2 - x^3)dx = \left(\frac{x^3}{3} - \frac{x^4}{4}\right)\Big|_0^1 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

Exemplo 2: Encontre a área limitada por duas curvas: $y_1 = 4x$ e $y_2 = x^3 + 3x^2$ (Figura 9).

Figura 9 – Área entre as curvas $y_1 = 4x$ e $y_2 = x^3 + 3x^2$



Fonte: Adaptado de Hoffmann e Bradley (2008, p. 336).

Através da Figura 9 podemos observar que há uma inversão entre as posições das curvas, uma vez que, para a área A_1 , definida no intervalo de x entre 4 e 0, temos $y_2 > y_1$. Logo:

$$A_1 = \int_{-4}^0 (x^3 + 3x^2 - 4x)dx = \left(\frac{x^4}{4} + \frac{3x^3}{3} - \frac{4x^2}{2}\right)\Big|_{-4}^0 =$$

$$= \left(\frac{x^4}{4} + x^3 - 2x^2 \right) \Big|_{-4}^0 = 0 - (64 - 64 - 32) = 32$$

Já para a área A_2 , definida no intervalo de x entre 0 e 1, temos $y_1 > y_2$. Assim:

$$\begin{aligned} A_2 &= \int_0^1 [4x - (x^3 + 3x^2)] dx = \int_0^1 (4x - x^3 - 3x^2) dx = \\ &= \left(2x^2 - \frac{x^4}{4} - x^3 \right) \Big|_0^1 = \left(2 - \frac{1}{4} - 1 \right) - (0) = \frac{8 - 1 - 4}{4} = \frac{3}{4} \\ A &= A_1 + A_2 = 32 + \frac{3}{4} = \frac{131}{4} \end{aligned}$$

2.6 Integrais impróprias

Nessa seção, estenderemos o conceito de integral definida para o caso em que o intervalo é infinito e também para o caso onde $f(x)$ tem uma descontinuidade infinita em $[a, b]$. Em ambos os casos, a integral é chamada integral imprópria (STEWART, 2017).

2.6.1 Tipo 1: Limites de integração infinitos

$$\int_a^\infty f(x) dx \text{ e } \int_{-\infty}^b f(x) dx$$

Segundo Chiang e Wainwright (2006), as integrais definidas dessa forma, isto é, com um limite de integração infinito, são chamadas de integrais impróprias. Nesses casos, não é possível avaliar as integrais como, respectivamente:

$$F(\infty) - F(a) \text{ e } F(b) - F(-\infty)$$

Isso porque ∞ não é um número e, portanto, não pode substituir x na função $F(x)$. Para resolver essas integrais, vamos recorrer ao conceito de limites.

De acordo com os autores, a integral imprópria $\int_a^\infty f(x) dx$ poderá ser definida como o limite de uma outra integral (própria) quando o limite superior (ou inferior) de integração tender a infinito, isto é:

$$\int_a^\infty f(x) dx = \lim_{b \rightarrow \infty} \int_a^b f(x) dx \text{ e } \int_{-\infty}^b f(x) dx = \lim_{a \rightarrow -\infty} \int_a^b f(x) dx$$

Chiang e Wainwright (2006) indicam que, se esse limite existir, diz-se que a integral imprópria é convergente (ou converge), e o processo de cálculo do limite dará o valor da integral. Se o limite não existir, diz-se que a integral imprópria é divergente e, na verdade, não tem significado.

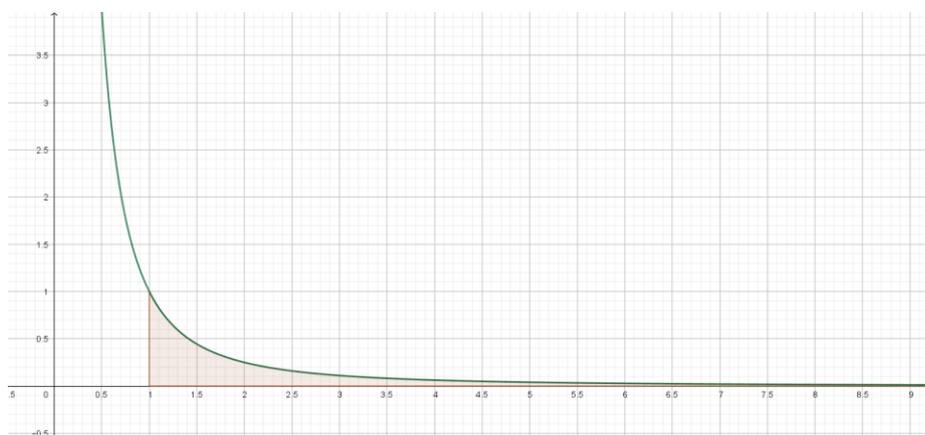
Para entender melhor, vamos ver alguns exemplos a seguir.

Exemplo 1: Calcule $\int_1^\infty \frac{1}{x^2} dx$.

$$\int_1^\infty \frac{1}{x^2} dx = \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x^2} dx = \lim_{b \rightarrow \infty} \left(\frac{-1}{x} \right) \Big|_1^b = \lim_{b \rightarrow \infty} \left(\frac{-1}{b} + 1 \right) = 1$$

Portanto, essa integral imprópria converge e seu valor é 1. Veja o gráfico dessa integral imprópria na Figura 10.

Figura 10 – Integral imprópria convergente.



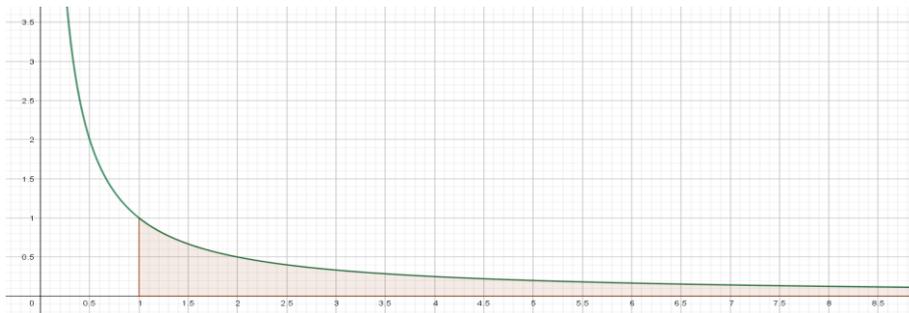
Fonte: Adaptado de Chiang e Wainwright (2006, p. 442).

Exemplo 2: Calcule $\int_1^{\infty} \frac{1}{x} dx$.

$$\int_1^{\infty} \frac{1}{x} dx = \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x} dx = \lim_{b \rightarrow \infty} \ln x \Big|_1^b = \lim_{b \rightarrow \infty} (\ln b - \ln 1) = \infty$$

Com esse resultado, temos que essa integral é divergente. Isso também é possível de ser visualizado no seu gráfico (Figura 11). Veja que a extensão da fronteira do lado direito na direção leste é indefinida e, por sua vez, resultará em uma área infinita, mesmo que a forma do gráfico apresente uma semelhança com o gráfico do exemplo 1 (Figura 10).

Figura 11 - Integral imprópria divergente.



Fonte: Adaptado de Chiang e Wainwright (2006, p. 442).

Há ainda os casos em que ambos os limites de integração são infinitos, então teremos

$$\int_{-\infty}^{\infty} f(x) dx = \lim_{\substack{b \rightarrow \infty \\ a \rightarrow -\infty}} \int_a^b f(x) dx,$$

e a integral converge se, e somente se, o limite existir.

2.6.2 Tipo 2: Integrando infinito

Mesmo com limites de integração finitos, uma integral ainda pode ser imprópria se o integrando, $f(x)$, tornar-se infinito em algum ponto do intervalo de integração $[a, b]$. Para calcular essa integral, vamos utilizar mais uma vez o

conceito de limite, conforme Chiang e Wainwright (2006). Vejamos alguns exemplos apresentados pelos autores para melhor compreensão.

Exemplo 3: Calcule $\int_0^9 x^{-1/2} dx$.

Observe que se trata de uma integral imprópria porque o integrando é infinito no limite inferior de integração: $\frac{1}{x^{1/2}} \rightarrow \infty$ quando $x \rightarrow 0^+$. Assim, teremos:

$$\int_0^9 x^{-1/2} dx \equiv \lim_{a \rightarrow 0^+} \int_a^9 x^{-1/2} dx = \lim_{a \rightarrow 0^+} (6 - 2\sqrt{a}) = 6$$

Portanto, a integral é convergente para 6.

Exemplo 4: Calcule $\int_0^1 \frac{1}{x} dx$.

$$\int_0^1 \frac{1}{x} dx \equiv \lim_{a \rightarrow 0^+} \int_a^1 \frac{1}{x} dx = \lim_{a \rightarrow 0^+} (-\ln a) = +\infty$$

Como esse limite não existe, a integral dada é divergente.

Chiang e Wainwright (2006) destacam que, se o integrando torna-se infinito no limite superior de integração, fazemos da mesma forma. No entanto, se o integrando torna-se infinito no intervalo aberto (a, b) e não em a ou b , então, supondo que $f(x) \rightarrow \infty$ quando $x \rightarrow p$, onde p é um ponto no intervalo (a, b) , pela propriedade da aditividade, temos:

$$\int_a^b f(x) dx = \int_a^p f(x) dx + \int_p^b f(x) dx.$$

A integral da esquerda pode ser considerada convergente se, e somente se, cada subintegral tiver um limite. Veja o exemplo abaixo.

Exemplo 5: Calcule $\int_{-1}^1 \frac{1}{x^3} dx$.

Observe que o integrando tende ao infinito quando x se aproxima de zero. Assim, devemos escrever a integral como a soma:

$$\int_{-1}^1 \frac{1}{x^3} dx = \int_{-1}^1 x^{-3} dx = \int_{-1}^0 x^{-3} dx + \int_0^1 x^{-3} dx \equiv i_1 + i_2$$

onde,

$$i_1 = \int_{-1}^b x^{-3} dx = -\frac{1}{2x^2} \Big|_{-1}^b = -\frac{1}{b^2} + \frac{1}{2} = -\infty.$$

Logo, podemos concluir que a integral é divergente, sem ter de avaliar i_2 .

2.7 Teorema do valor médio para integrais

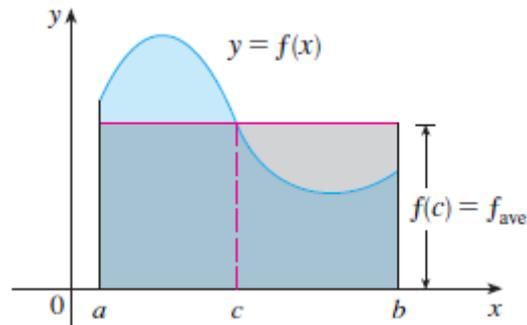
Se $f(x)$ é uma função contínua no intervalo $[a, b]$, então existe um número c no intervalo $[a, b]$ tal que:

$$f(c) \cdot (b - a) = \int_a^b f(x) dx$$

em que $f(c)$ é o valor médio (STEWART, 2017).

A interpretação geométrica do teorema do valor médio para integrais diz que, para funções positivas $f(x)$, existe um número c tal que o retângulo com base $[a, b]$ e altura $f(c)$ tem a mesma área que a região sob o gráfico de $f(x)$ desde a até b . Observe a Figura 12 abaixo.

Figura 12 – Teorema do valor médio.



Fonte: Stewart (2017, p. 410).

Exemplo 1: Calcule o valor médio da função $f(x) = x^2$ no intervalo $[1, 4]$ e encontre um valor c neste intervalo tal que $f(c)$ seja o valor médio.

Pela fórmula do teorema do valor médio temos:

$$f(c) \cdot (b - a) = \int_a^b f(x) dx \rightarrow f(c) = \frac{1}{(b - a)} \cdot \int_a^b f(x) dx$$

Fazendo as devidas substituições:

$$f(c) = \frac{1}{(4 - 1)} \cdot \int_1^4 x^2 dx = \frac{1}{3} \cdot \frac{x^3}{3} \Big|_1^4 = \frac{1}{3} \cdot \left(\frac{64}{3} - \frac{1}{3} \right) = \frac{1}{3} \cdot 21 = 7$$

Portanto, sendo $f(c) = 7$:

$$c^2 = 7 \rightarrow c = \sqrt{7}$$

2.8 Aplicações de integrais definidas

A seguir veremos alguns exercícios com aplicações práticas de integral definida retirados de Hoffman e Bradley (2008).

Exercício 1: Suponha que daqui a t anos um investimento esteja gerando lucro a uma taxa de $P_1'(t) = 50 + t^2$ centenas de reais por ano e um segundo investimento esteja gerando lucro a uma taxa de $P_2'(t) = 200 + 5t$ centenas de reais por ano.

a) Durante quantos anos o índice de rentabilidade do segundo investimento permanecerá maior que o do primeiro?

O índice de rentabilidade do segundo investimento é maior que o primeiro até que:

$$\begin{aligned} P_1'(t) &= P_2'(t) \\ 50 + t^2 &= 200 + 5t \\ t^2 - 5t - 150 &= 0 \\ (t - 15)(t + 10) &= 0 \\ t &= 15 \text{ ou } t = -10 \end{aligned}$$

Como estamos falando em anos, vamos desprezar o valor negativo, logo $t = 15$ anos.

b) Determine o excesso líquido de lucro para o período calculado no item anterior. Interprete o excesso líquido de lucro como uma área entre curvas.

O excesso de lucro do plano 2 em relação ao plano 1 é $E(t) = P_2(t) - P_1(t)$, e o excesso líquido de lucro EL no período $0 \leq t \leq 15$, calculado no item anterior, é dado pela integral definida:

$$\begin{aligned} EL &= \int_0^{15} E'(t) dt = \int_0^{15} [P_2'(t) - P_1'(t)] dt \\ &= \int_0^{15} [(200 + 5t) - (50 + t^2)] dt = \int_0^{15} [150 + 5t - t^2] dt = \\ &= \left[150t + \frac{5}{2}t^2 - \frac{t^3}{3} \right] \Big|_0^{15} = \left[150(15) + \frac{5}{2}(15)^2 - \frac{(15)^3}{3} \right] - [0] = \\ &= 1687,50 \text{ centenas de reais } \therefore R\$ 168.750,00 \text{ (a área entre as duas curvas)} \end{aligned}$$

Exercício 2: Em uma certa fábrica, o custo marginal é $3(q - 4)^2$ reais por unidade quando o nível de produção é q unidades. Qual é o aumento do custo de fabricação quando o nível de produção aumenta de 6 para 10 unidades?

Para encontrarmos a solução, precisamos calcular a $\int_6^{10} 3(q - 4)^2 dq$.

Por substituição, temos:

$$u = q - 4 \rightarrow \frac{du}{dq} = 1 \rightarrow du = dq.$$

Logo,

$$\begin{aligned} \int_6^{10} 3(q - 4)^2 dq &= 3 \int_{q=6}^{q=10} u^2 du = \\ &= 3 \cdot \frac{u^3}{3} \Big|_{q=6}^{q=10} = (q - 4)^3 \Big|_6^{10} = (10 - 4)^3 - (6 - 4)^3 = 208. \end{aligned}$$

Exercício 3: Quando possui t anos de serviço, uma certa máquina industrial gera receita a uma taxa $R'(t) = 5000 - 20t^2$ reais por ano, e os custos de operação e manutenção da máquina aumentam a uma taxa $C'(t) = 2000 + 10t^2$ reais por mês.

a) Qual é a vida útil da máquina?

$$\begin{aligned} R'(t) &= C'(t) \\ 5000 - 20t^2 &= 2000 + 10t^2 \\ 5000 - 2000 &= 10t^2 + 20t^2 \rightarrow 3000 = 30t^2 \rightarrow t^2 = \frac{3000}{30} \\ t^2 &= 100 \rightarrow t = \sqrt{100} \rightarrow t = 10 \text{ anos} \end{aligned}$$

b) Calcule o lucro gerado pela máquina durante sua vida útil.

$$\int_0^{10} [R'(t) - C'(t)] dt = \int_0^{10} [(5000 - 20t^2) - (2000 + 10t^2)] dt =$$

$$\begin{aligned}
 &= \int_0^{10} (3000 - 30t^2) dt = \left(3000t - \frac{30t^3}{3} \right) \Big|_0^{10} = (3000t - 10t^3) \Big|_0^{10} = \\
 &= (3000 \cdot 10 - 10 \cdot 10^3) - 0 = 20000
 \end{aligned}$$

Logo, um lucro de R\$ 20.000,00 será gerado pela máquina em sua vida útil.

Exercício 4: O gerente de uma sapataria observa que o preço (em reais) de um par de sapatos de uma marca popular varia a uma taxa de $p'(x) = \frac{-300x}{(x^2+9)^{3/2}}$ quando x centenas de pares são demandados (comprados) pelos consumidores. Quando o preço do par é de R\$ 75,00, 400 pares são demandados pelos consumidores.

a) Determine a função demanda $p(x)$.

$$\int \frac{-300x}{(x^2 + 9)^{3/2}} dx = -300 \int \frac{x}{(x^2 + 9)^{3/2}} dx$$

Por substituição, teremos:

$$u = x^2 + 9 \rightarrow \frac{du}{dx} = 2x \rightarrow x dx = \frac{du}{2}$$

Assim, a integral passará a ser da função:

$$\int \frac{1}{2} \cdot u^{-3/2} du = \frac{-300}{2} \cdot \frac{u^{-1/2}}{-1/2} + C = -\frac{300}{2} \cdot -2 \cdot \frac{1}{u^{1/2}} + C = \frac{300}{(x^2 + 9)^{1/2}} + C = p(x)$$

Se $p(4) = 75$, então:

$$\frac{300}{(4^2 + 9)^{1/2}} + C = 75 \rightarrow \frac{300}{5} + C = 75 \rightarrow C = 75 - 60 \rightarrow C = 15$$

$$\therefore p(x) = \frac{300}{(x^2 + 9)^{1/2}} + 15$$

b) Qual o preço corresponde à demanda de 500 pares?

$$p(5) = \frac{300}{(5^2 + 9)^{\frac{1}{2}}} + 15 = \frac{300}{34^{1/2}} + 15 \cong R\$ 66,36$$

c) Qual é a demanda se o preço do par de sapatos é de R\$ 90,00?

$$p(x) = 90 \rightarrow x = ?$$

$$90 = \frac{300}{(x^2 + 9)^{\frac{1}{2}}} + 15 \rightarrow \frac{300}{(x^2 + 9)^{\frac{1}{2}}} = 75 \rightarrow [(x^2 + 9)^{\frac{1}{2}}]^2 = \left(\frac{300}{75}\right)^2$$

$$x^2 + 9 = 16 \rightarrow x^2 = 7 \rightarrow x = \sqrt{7} \rightarrow x \cong 2,64$$

Como x se refere a centenas de pares, é preciso multiplicar o valor de x encontrado por 100. Assim, a demanda ao preço de R\$ 90,00 será de aproximadamente 264 pares de sapatos.

Revisando

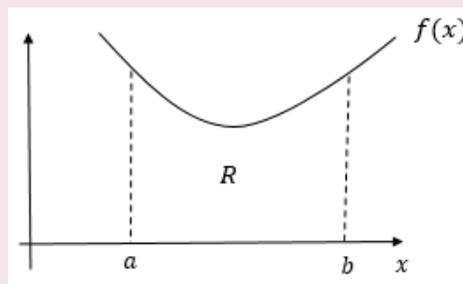
- Integral definida:

$$\int_a^b f(x)dx = \lim_{n \rightarrow \infty} = \Delta x \cdot [f x_1 + f x_2 + \dots + f x_n]$$

- Área sob uma curva:

$$\text{Área de } R = \int_a^b f(x)dx$$

Figura 13 – Área R sob uma curva



Fonte: Adaptado de Hoffmann e Bradley (2008, p. 370)

- Propriedades de integrais definidas:

$$\int_b^a f(x)dx = - \int_a^b f(x)dx$$

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx \quad (a < b < c)$$

$$\int_a^b kf(x)dx = k \cdot \int_a^b f(x)dx$$

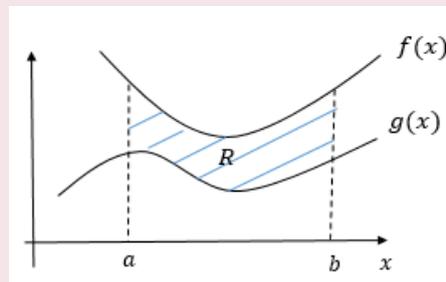
$$\int_a^b [f(x) + g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx$$

- Teorema fundamental do cálculo:

$$\int_a^b f(x)dx = F(b) - F(a) \quad \text{onde } F'(x) = f(x)$$

Área entre duas curvas:

Figura 14 – Área R entre duas curvas



Fonte: Adaptado de Hoffmann e Bradley (2008, p. 334).

$$\text{Área de } R = \int_a^b [f(x) - g(x)]dx$$

- Integrais impróprias:

$$\int_a^\infty f(x) dx = \lim_{b \rightarrow \infty} \int_a^b f(x) dx \quad \text{e} \quad \int_{-\infty}^b f(x) dx = \lim_{a \rightarrow -\infty} \int_a^b f(x) dx$$

$$\int_{-\infty}^\infty f(x) dx = \lim_{\substack{b \rightarrow \infty \\ a \rightarrow -\infty}} \int_a^b f(x) dx$$

- Valor médio para integrais:

$$f(c) = \frac{1}{(b-a)} \int_a^b f(x) dx$$

Saiba mais

Vídeo: **Integral imprópria com limites de integração infinitos**, por Paulo Pereira.

INTEGRAL IMPRÓPRIA com limites de integração infinitos - Cálculo 1 (#48) (Exemplos resolvidos). Publicado pelo canal Equaciona com Paulo Pereira. Disponível em: <https://youtu.be/nCbocso-Pbw>. Acesso em: 29 out. 2020.

Vídeo: **Integral imprópria com integrandos infinitos**, por Paulo Pereira.

INTEGRAL IMPRÓPRIA com integrandos infinitos – Cálculo 1 (#49). Publicado pelo canal Equaciona com Paulo Pereira. Disponível em: <https://youtu.be/2kyjGsqPlko> Acesso em: 29 out. 2020.

Material didático de apoio: **Introdução ao estudo das integrais**, de Patrícia Pugliesi.

CARNEIRO, Patrícia Pugliesi. **Introdução ao estudo das integrais**. São Cristóvão, SE: Departamento de Economia da Universidade Federal de Sergipe, 2014. Disponível em: http://dee.ufs.br/uploads/page_attach/path/2261/INTEGRAIS_-_Material_Did_tico_de_Apoio_2_.pdf. Acesso em: 05 out. 2020.

Referências

CHIANG, Alpha C.; WAINWRIGHT, Kevin. **Matemática para economistas**. 4. ed. Rio de Janeiro: Elsevier, 2006.

HOFFMANN, Laurence D.; BRADLEY, Gerald L. **Cálculo**: um curso moderno e suas aplicações. 9. ed. Rio de Janeiro: Livro Técnico e Científico, 2008.

STEWART, James. **Cálculo, v. 1**. 8. ed. São Paulo: Cengage Learning, 2017.

3

Estatística descritiva

Métodos Quantitativos Aplicados
às Ciências Contábeis 2

Lidia Maria Alves Rodella Amorim

Objetivos da aprendizagem

- 1 Criar uma distribuição de frequência para um conjunto de dados;
- 2 Calcular e interpretar medidas de tendência central e de dispersão;
- 3 Calcular e interpretar a covariância e um coeficiente de correlação.

Introdução

A estatística é o conjunto de técnicas que permite, de forma sistemática, coletar, organizar, descrever, analisar e interpretar dados oriundos de estudos ou experimentos realizados em qualquer área do conhecimento.

Um exemplo de aplicação da estatística na área de auditoria, apresentado por Doane e Seward (2014), é o de uma empresa que realiza o pagamento de mais de 12 mil faturas por mês e sabe que algumas faturas são pagas incorretamente. Os auditores não podem checar todas as faturas, então selecionam uma amostra para estimar a proporção das faturas pagas indevidamente. Qual deve ser o tamanho da amostra para que os auditores confiem que a estimativa esteja suficientemente próxima da real proporção?

Além disso, a maioria das informações estatísticas publicadas nos jornais, revistas, relatórios de empresas e outras publicações consiste em dados resumidos e apresentados de forma que sejam de fácil compreensão para o público, como destacam Anderson *et al* (2011). Esses sumários de dados, que podem ser gráficos, tabelas ou números resumidos, são conhecidos como **estatística descritiva**, assunto que veremos neste capítulo após a introdução de alguns conceitos básicos.

3.1 Conceitos básicos de estatística

Para saber resolver o problema apresentado acima é preciso aprender ferramentas das duas espécies principais de estatística, como destacam Doane e Seward (2014, p. 6):

“Estatística descritiva corresponde à coleta, organização, apresentação e resumo de dados (com diagramas e gráficos ou utilizando um valor numérico resumido). Inferência estatística refere-se a generalizar resultados de uma amostra para uma população, estimar parâmetros desconhecidos, chegar a conclusões e tomar decisões.”

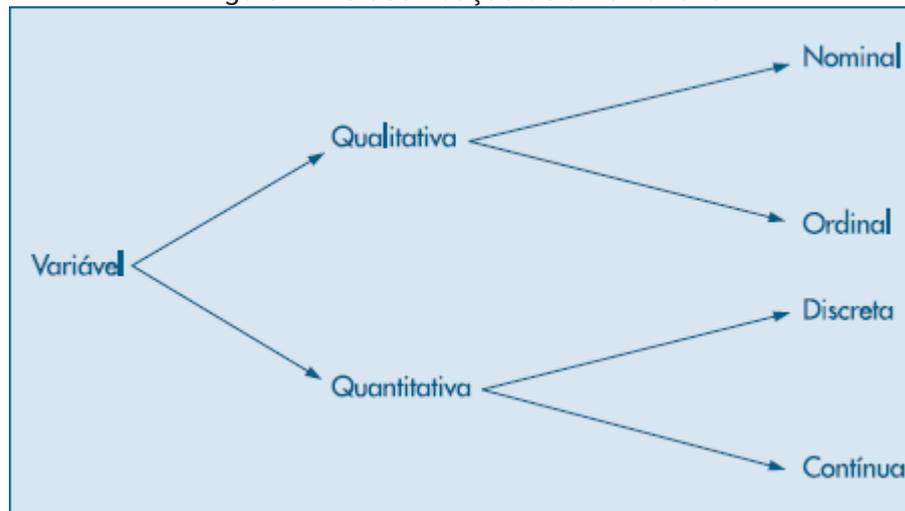
Vamos iniciar aprendendo algumas definições básicas, ok?

Há dois tipos de conjuntos de dados usados em estatística: população e amostra. Uma **população** é o conjunto de objetos, indivíduos ou resultados experimentais acerca do qual se pretende estudar alguma característica comum. Uma **amostra** é uma parte da população que é observada com o objetivo de obter informação para estudar a característica pretendida.

A partir de uma amostra de n itens escolhidos de uma população, calculamos estatísticas que podem ser usadas como estimativas de parâmetros encontrados na população (DOANE; SEWARD, 2014). Assim, temos que **parâmetro** é a descrição numérica de uma característica da população e **estatística** é a descrição numérica de uma característica da amostra.

As características dos elementos da população ou da amostra são chamadas de **variáveis** e são classificadas conforme o esquema apresentado por Bussab e Morettin (2010) na Figura 1. Por outro lado, **dados** são os valores observados das variáveis.

Figura 1 – Classificação de uma variável



Fonte: Bussab e Morettin (2010, p. 10).

Como destacam com os autores, algumas variáveis, chamadas de **qualitativas** (como sexo, educação e estado civil), apresentam como possíveis resultados uma qualidade (ou atributo) do indivíduo pesquisado. Já outras, as **variáveis quantitativas** (como número de filhos, salário, idade), apresentam como possíveis realizações os números decorrentes de uma contagem ou mensuração.

Observe na Figura 1 que ainda é possível dividir as variáveis qualitativas em dois tipos, apresentados por Bussab e Morettin (2010): a variável **qualitativa nominal**, para a qual não existe nenhuma ordenação nas possíveis realizações (por exemplo, cor do cabelo e estado civil), e a variável **qualitativa ordinal**, para a qual existe uma ordem nos seus resultados (mês de observação, grau de instrução etc.).

De modo semelhante, esses autores destacam que as variáveis quantitativas podem ser classificadas como: **discretas**, quando seus possíveis valores formam um conjunto de números finito ou enumerável, e que resultam, geralmente, de uma contagem (número de filhos); ou **contínuas**, cujos possíveis valores pertencem a um intervalo de números reais e que resultam de uma mensuração (altura, valores monetários).

Voltando à distinção entre população e amostra, temos mais dois conceitos relevantes: o **censo** é realizado para o levantamento de dados de variáveis referentes a uma população, já a **amostragem** se refere a uma contagem ou medição de parte de uma população e é mais comumente usada nos estudos estatísticos.

Conforme Doane e Seward (2014) explicam, há duas categorias principais de métodos de amostragem: a **probabilística**, em que os itens são selecionados aleatoriamente (ou por um procedimento que envolve o acaso), buscando produzir uma amostra representativa de uma população; e a **não probabilística**, que é não aleatória e, portanto, é menos científica, mas, algumas vezes, é usada por conveniência.

Focando na amostragem probabilística, mais comumente utilizada nos estudos científicos, temos as quatro técnicas apresentadas a seguir, definidas conforme Larson e Farber (2010):

- **Amostra aleatória simples:** é aquela em que toda amostra de mesmo tamanho tem igual possibilidade de ser selecionada. Ou seja, todos os itens na população de N itens têm a mesma chance de serem escolhidos na amostra de n itens. Uma forma de selecionar uma amostra aleatória simples é definir um número diferente para cada membro da população e, então, fazer um sorteio ou usar uma tabela numérica aleatória. A amostragem pode ser (DOANE; SEWARD, 2014):

Sem reposição: uma vez que um item seja incluído em uma amostra, ele não poderá ser selecionado novamente para a mesma amostra.

Com reposição: o mesmo número aleatório pode aparecer mais de uma vez.

- **Amostra sistemática:** é aquela na qual é atribuído um número a cada elemento da população. Os itens que compõem a população são ordenados de alguma forma, depois um número inicial é selecionado aleatoriamente e, então, itens da amostra são selecionados em intervalos regulares a partir do número inicial.

- **Amostra estratificada:** quando é importante que uma amostra tenha membros de cada segmento da população. Esse método é aplicável quando a população pode ser dividida em subgrupos homogêneos de tamanhos conhecidos (os estratos). Dentro de cada estrato, pode-se tomar uma amostra aleatória simples do tamanho desejado.

- **Amostra por conglomerado (ou por agrupamento):** quando a população está em subgrupos que ocorrem naturalmente, cada um tendo características similares. Para selecionar uma amostra por agrupamento, divide a população em grupos, chamados clusters, e selecione todos os membros em um ou mais (mas não em todos) clusters.

São, essencialmente, estratos consistindo em regiões geográficas, e cada cluster deve consistir de membros com todas as características de interesse.

3.2 Distribuições de frequências

Segundo Bussab e Morettin (2010), quando se estuda uma característica, ou melhor, uma variável, o maior interesse do pesquisador é conhecer o comportamento dessa variável, analisando a ocorrência de suas possíveis realizações. Para isso, veremos nesta seção uma forma de se organizar um conjunto de realizações para termos uma visão global sobre elas, ou seja, uma visão global de sua distribuição.

Quando um dado tem muitas entradas, pode ser difícil de observar padrões, como observam Larson e Farber (2010). Para auxiliar na organização e análise de um conjunto de dados, eles podem ser agrupados em intervalos chamados de *classes* e formar, assim, uma distribuição de frequência. Ela também poderá ser utilizada para a construção de gráficos.

Portanto, Larson e Farber (2010) definem a **distribuição de frequência** como uma tabela que mostra classes (blocos) ou intervalos das entradas de dados com uma contagem do número de entradas (observações) em cada classe. Definimos **frequência f** (ou **frequência absoluta**) de um valor de uma variável como sendo o número de vezes que aquele valor se repete no conjunto de dados observados. Os passos básicos destacados por Doane e Seward (2014) para a construção de uma distribuição de frequência são:

1) Organizar os dados em ordem crescente.

2) Definir o número de classes. Como uma distribuição de frequência busca sintetizar muitos pontos dos dados em uma tabela relativamente pequena, espera-se que o número k de classes seja muito menor que o tamanho da amostra n . É importante destacar que, quanto menos classes há, mais informações perdemos. No entanto, quanto mais classes há, menos os dados estarão resumidos. No caso de variáveis quantitativas, a regra de Sturges, apresentada abaixo, pode auxiliar na definição do número de classes.

$$k = 1 + 3,3 \log n$$

No entanto, a quantidade e os limites das classes têm que ser razoáveis e apropriados. Isso é mais importante que a regra de Sturges. Lembre-se que, no caso das variáveis qualitativas, as classes são definidas pelos possíveis atributos.

3) Estabelecer os limites das classes. Estabelecer os limites das classes requer certo discernimento. Como orientação, encontre o comprimento aproximado de cada classe por meio da divisão da amplitude dos dados (maior valor observado menos o menor valor observado) pelo número de classes:

$$\text{Comprimento da classe} \approx \frac{x_{max} - x_{min}}{k}$$

4) Alocar os valores dos dados nas classes apropriadas. Geralmente, o limite inferior da classe está incluído, enquanto o limite superior está excluído. Lembre-se que nenhum dos blocos deve se sobrepor e que cada valor é computado em apenas uma das classes.

5) Criar a tabela. Na tabela de distribuição de frequência podem ser incluídas diversas características adicionais (como colunas adicionais) que ajudarão a fornecer um melhor entendimento dos dados. Larson e Farber (2010) sugerem os seguintes acréscimos:

- **Ponto médio da classe:** é o quociente da soma dos limites inferiores e superiores da classe por 2.

$$\text{Ponto médio} = \frac{(\text{limite inferior da classe}) + (\text{limite superior da classe})}{2}$$

- **Frequência relativa da classe:** é a proporção ou porcentagem de dados que está em determinada classe. Para encontrá-la, divida a frequência f pelo tamanho da amostra, n .

$$\text{Frequência relativa } (f_{r_i}) = \frac{\text{Frequência da classe } (f_i)}{n}$$

- **Frequência acumulada da classe:** é a soma da frequência daquela classe com todas as anteriores. Conseqüentemente, a frequência acumulada da última classe é igual ao tamanho da amostra, n .

Para entender melhor, vamos construir uma tabela de distribuição de frequência referente à variável *grau de instrução* a partir dos dados de Bussab e Morettin (2010) apresentados na Tabela 1.

Tabela 1 – Informações de 36 empregados da seção de orçamentos da Companhia MB.

Nº	Estado civil	Grau de instrução	Nº de filhos	Salário (x sal. mín.)	Idade		Região de procedência
					anos	meses	
1	solteiro	ensino fundamental	—	4,00	26	03	interior
2	casado	ensino fundamental	1	4,56	32	10	capital
3	casado	ensino fundamental	2	5,25	36	05	capital
4	solteiro	ensino médio	—	5,73	20	10	outra
5	solteiro	ensino fundamental	—	6,26	40	07	outra
6	casado	ensino fundamental	0	6,66	28	00	interior
7	solteiro	ensino fundamental	—	6,86	41	00	interior
8	solteiro	ensino fundamental	—	7,39	43	04	capital
9	casado	ensino médio	1	7,59	34	10	capital
10	solteiro	ensino médio	—	7,44	23	06	outra
11	casado	ensino médio	2	8,12	33	06	interior
12	solteiro	ensino fundamental	—	8,46	27	11	capital
13	solteiro	ensino médio	—	8,74	37	05	outra
14	casado	ensino fundamental	3	8,95	44	02	outra
15	casado	ensino médio	0	9,13	30	05	interior
16	solteiro	ensino médio	—	9,35	38	08	outra
17	casado	ensino médio	1	9,77	31	07	capital
18	casado	ensino fundamental	2	9,80	39	07	outra
19	solteiro	superior	—	10,53	25	08	interior
20	solteiro	ensino médio	—	10,76	37	04	interior
21	casado	ensino médio	1	11,06	30	09	outra
22	solteiro	ensino médio	—	11,59	34	02	capital
23	solteiro	ensino fundamental	—	12,00	41	00	outra
24	casado	superior	0	12,79	26	01	outra
25	casado	ensino médio	2	13,23	32	05	interior
26	casado	ensino médio	2	13,60	35	00	outra
27	solteiro	ensino fundamental	—	13,85	46	07	outra
28	casado	ensino médio	0	14,69	29	08	interior
29	casado	ensino médio	5	14,71	40	06	interior
30	casado	ensino médio	2	15,99	35	10	capital
31	solteiro	superior	—	16,22	31	05	outra
32	casado	ensino médio	1	16,61	36	04	interior
33	casado	superior	3	17,26	43	07	capital
34	solteiro	superior	—	18,75	33	07	capital
35	casado	ensino médio	2	19,40	48	11	capital
36	casado	superior	3	23,30	42	02	interior

Fonte: Bussab e Morettin (2010, p. 11).

A partir desses dados, temos a seguinte distribuição de frequência segundo o grau de instrução (Tabela 2).

Observando os dados da segunda coluna da Tabela 2, verifica-se que, dos 36 empregados do setor, 12 têm o ensino fundamental, 18 o ensino médio e 6 o superior. Na penúltima coluna são apresentadas as porcentagens para cada possível resultado da variável *grau de instrução*. Elas são muito importantes quando se quer comparar resultados de duas pesquisas diferentes.

Tabela 2 – Frequências e porcentagens dos 36 funcionários da seção de orçamento da companhia MB segundo o grau de instrução.

Grau de instrução	Frequência	Frequência relativa	Porcentagem	Frequência acumulada
Fundamental	12	0,3333	33,33	12
Médio	18	0,5000	50,00	30
Superior	6	0,1667	16,67	36
Total	36	1	100	-

Fonte: Bussab e Morettin (2010, p. 12).

A construção de tabelas de frequência para variáveis contínuas requer certa atenção. Por exemplo, a tabela de frequências para a variável *salário*, baseada nos dados da Tabela 1, usando o mesmo procedimento acima, não poderá resumir todas as 36 observações num grupo menor, pois não existem dados iguais. A solução utilizada é agrupar os dados por faixas de salário, ou seja, intervalos de classes, conforme a Tabela 3.

Tabela 3 – Frequências e porcentagens dos 36 funcionários da seção de orçamento da companhia MB por faixa de salário

Classe de salários	Frequência	Porcentagem	Ponto médio
4,00 –8,00	10	27,78	6,00
8,00 –12,00	12	33,33	10,00
12,00 –16,00	8	22,22	14,00
16,00 –20,00	5	13,89	18,00
20,00 –24,00	1	2,78	22,00
Total	36	100	-

Fonte: Bussab e Morettin (2010, p. 13).

Diante do exposto, vemos que ao resumir os dados relativos a uma variável contínua, perde-se alguma informação. Por exemplo, não sabemos quais são os oito salários da classe 12 a 16, a não ser que seja analisada a Tabela 1. Observe, ainda, que usamos a notação $a|–b$ para o intervalo de número contendo o limite inferior a , mas não contendo o limite superior b .

3.3 Gráficos

Como destacam Bussab e Morettin (2010), a representação gráfica da distri-

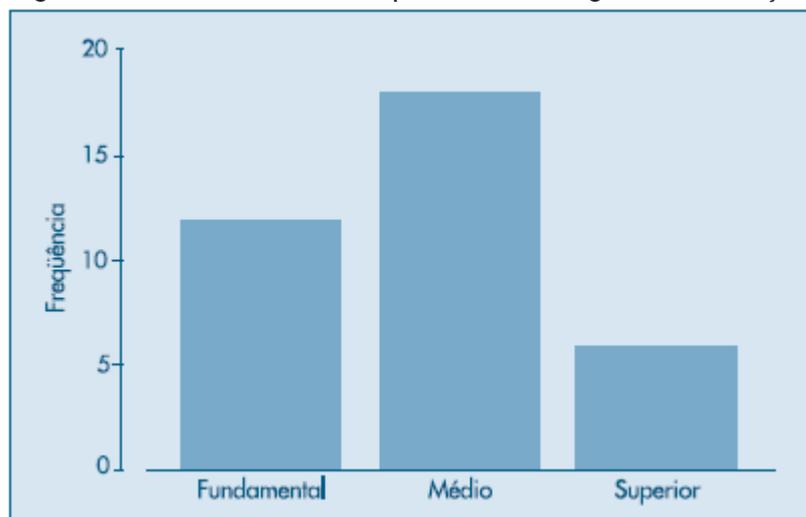
buição de uma variável tem a vantagem de informar sobre sua variabilidade rápida e resumidamente. Em virtude de suas características, nem sempre os gráficos que se aplicam às variáveis quantitativas podem ser utilizados para as variáveis qualitativas.

3.3.1 Gráficos para variáveis qualitativas

Há uma diversidade de tipos de gráficos para representar variáveis qualitativas. Alguns são versões diferentes do mesmo princípio, de modo que irei apresentar dois deles: gráficos em barras e gráficos de composição de setores (“pizza”).

Conforme descrevem Bussab e Morettin (2010), o **gráfico em barras** consiste em construir retângulos ou barras com uma das dimensões proporcional à magnitude a ser representada (frequência ou proporção), e a outra dimensão arbitrária, porém igual para todas as barras. Essas barras são apresentadas paralelamente umas às outras, de forma horizontal ou vertical. Na Figura 2 temos o gráfico em barras verticais para a variável grau de instrução, de acordo com a Tabela 2 de distribuição de frequência.

Figura 2 – Gráfico em barras para a variável *grau de instrução*.

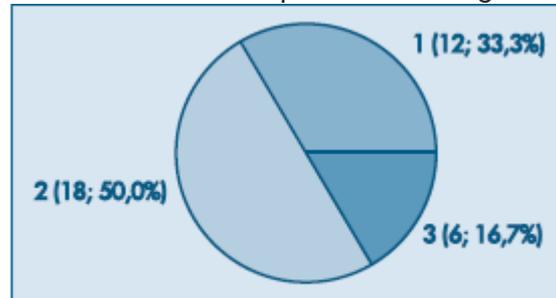


Fonte: Bussab e Morettin (2010, p. 15).

Por outro lado, de acordo com os autores, o **gráfico de composição em setores** visa representar a composição de partes de um todo, usualmente em porcentagem, sendo o formato de “pizza” o mais conhecido. É composto por um círculo de raio

arbitrário, representando o todo, dividido em setores, que correspondem às classes de forma proporcional. Veja a Figura 3 para a variável grau de instrução.

Figura 3 – Gráfico em setores para a variável grau de instrução



1 = Fundamental, 2 = Médio e 3 = Superior

Fonte: Bussab e Morettin (2010, p. 16).

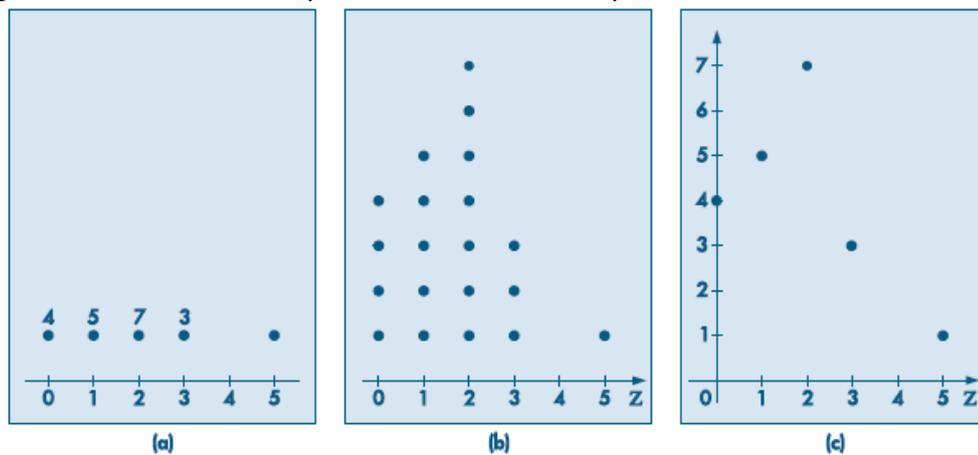
3.3.2 Gráficos para variáveis quantitativas

Para variáveis quantitativas há uma variedade maior de tipos de gráficos. Além dos gráficos usados para as variáveis qualitativas, que também podem ser utilizados para as variáveis quantitativas, veremos nesta seção o gráfico de dispersão unidimensional e o histograma.

No **gráfico de dispersão unidimensional**, a frequência é representada por pontos, como vemos na Figura 4, que representa dados sobre o número de filhos da Tabela 1. Esses pontos podem:

- (a) ser dispostos ao longo de uma reta, acompanhados por um número que indica as repetições, como na Figura 4 (a);
- (b) ter os valores repetidos “empilhados” um sobre o outro, como na Figura 4 (b);
- (c) apresentar o ponto mais alto da pilha, como na Figura 4 (c), que é o modo de representação mais comumente utilizado.

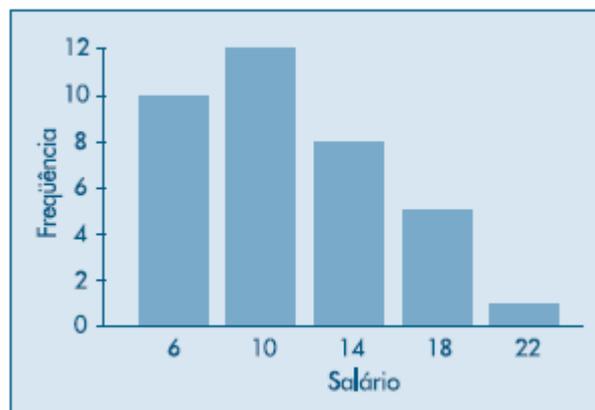
Figura 4 – Gráficos de dispersão unidimensional para a variável *número de filhos*.



Fonte: Bussab e Morettin (2010, p. 17).

Para variáveis quantitativas contínuas, são necessárias algumas adaptações. Vejamos o caso da distribuição de frequência da variável *salário* (Tabela 3). Para fazer um gráfico em barras, devemos usar o artifício de aproximar a variável contínua de uma variável discreta, o que pode ser feito supondo-se que todos os salários em determinada classe são iguais ao ponto médio desta classe. Assim, os dez salários da primeira classe serão considerados iguais a 6,00, os 12 salários da segunda classe iguais a 10,00 e assim por diante, conforme Figura 5.

Figura 5 – Gráfico em barras para a variável salários.



Fonte: Bussab e Morettin (2010, p. 18).

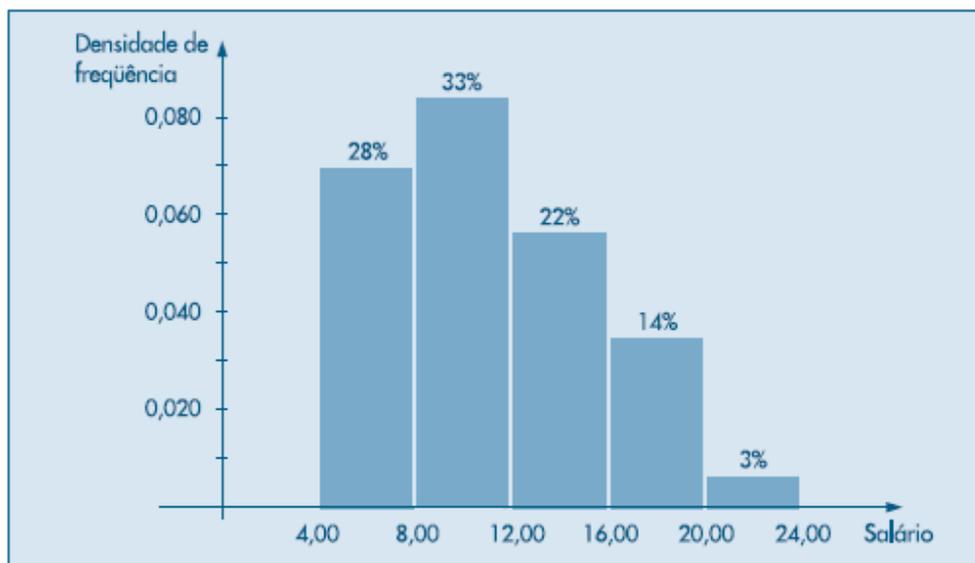
Contudo, os autores destacam que essa forma de representar uma variável contínua faz com que se percam muitas das informações nela contidas. Nesses casos, recomenda-se o uso do gráfico conhecido como *histograma*, que é definido por

Bussab e Morettin (2010) como um gráfico de barras contíguas, ou seja, sem espaço entre as barras, com as bases proporcionais aos intervalos das classes e a área de cada retângulo proporcional à respectiva frequência, que pode ser tanto a frequência absoluta quanto a frequência relativa.

Considerando a amplitude (limite superior menos o limite inferior) do i -ésimo intervalo por Δ_i , para que a altura do retângulo respectivo seja proporcional a f_i , a sua altura deve ser proporcional à sua **densidade de frequência** (f_i/Δ_i , ou seja, frequência absoluta dividida pela respectiva amplitude). Com essa convenção, a área total do histograma será igual a um. Veja o gráfico do histograma da variável salários na Figura 6.

Quando os intervalos de classe (amplitude) são iguais, podemos utilizar, no eixo das ordenadas do histograma, tanto os valores das frequências absolutas, quanto os das frequências relativas, ou os da proporção ou, ainda, os das densidades de frequência. Contudo, se os intervalos de classe (amplitudes) não forem iguais, teremos que utilizar as densidades de frequência ou a proporção no eixo Y.

Figura 6 – Histograma da variável *salários*.



Fonte: Bussab e Morettin (2010, p. 19).

Atenção!

O histograma é importante porque sugere a forma da distribuição da população que estamos amostrando. Ele também ajuda a explicar um importante aspecto de probabilidade e dá uma ideia da localização do centro dos dados e da variabilidade do conjunto de dados.

3.4 Medidas descritivas

Na estatística descritiva são utilizadas técnicas chamadas de medidas descritivas, que são funções de valores de uma variável numérica e que, segundo Bussab e Morettin (2010), reduzem um conjunto de dados numéricos a um pequeno grupo de valores representativos de toda informação relevante a respeito desses dados.

A descrição numérica dos dados possui grupos característicos de medidas: tendência central, dispersão ou variabilidade, forma, posição e associação. Veremos a seguir as medidas que compõem cada um desses grupos, lembrando que as medidas descritivas derivadas de uma amostra são denominadas **estatísticas** e as de uma população são chamadas **parâmetros**, como destacam Doane e Seward (2014).

3.4.1 Medidas de tendência central

De acordo com esses autores, as medidas de tendência central buscam descrever o valor do meio ou os valores típicos de um conjunto de dados. Para se ter uma ideia clara a respeito da tendência central dos dados, é preciso olhar para várias medidas. As três mais utilizadas são a média, a mediana e a moda.

A **média** é a medida mais conhecida e mais utilizada, em virtude da sua facilidade de cálculo e de compreensão. Ela pode se apresentar de duas formas:

- **Média aritmética simples** – É a soma das observações dividida pelo número delas (BUSSAB; MORETTIN, 2010). Todos os valores participam do

cálculo com o mesmo peso. Assim, para um conjunto de n valores da variável $X (x_1, x_2, x_3, \dots, x_n)$, temos:

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

- **Média ponderada** – É a soma das observações dividida pelo número delas na qual pelo menos um dos valores possui um peso diferente dos demais (BUSSAB; MORETTIN, 2010). Portanto, teremos um conjunto de valores da variável $X (x_1, x_2, x_3, \dots, x_n)$ e um conjunto de pesos $(p_1, p_2, p_3, \dots, p_n)$:

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

A **mediana** (md), segundo Bussab e Morettin (2010) é o valor que ocupa a posição central da série de observações, quando elas estão ordenadas em ordem crescente. Metade dos dados é igual ou inferior à mediana e metade é igual ou superior. A mediana é a observação do meio em uma série de dados ordenada se n for um número ímpar, mas é a média das duas observações centrais se n for um número par, conforme salientam Doane e Seward (2014). Assim, considerando $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n-1)} \leq x_{(n)}$, podemos apresentá-la formalmente como Bussab e Morettin (2010):

$$md(X) = \begin{cases} x_{(\frac{n+1}{2})}, se\ n\ ímpar \\ \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2}, se\ n\ par \end{cases}$$

Por exemplo, nessa série de cinco observações $A = \{3, 4, 7, 8, 8\}$, temos $md = x_{(3)} = 7$. Já na série de seis observações $B = \{3, 4, 7, 8, 8, 9\}$:

$$md = \frac{x_{(3)} + x_{(4)}}{2} = \frac{7+8}{2} = 7,5.$$

Como definem Bussab e Morettin (2010), a **moda** é o valor mais frequente do conjunto de valores observados. Além disso, é a única medida que pode não existir e, existindo, pode não ser única.

Vamos encontrar essas medidas de tendência central (média, mediana e moda) para o seguinte conjunto de dados: 25, 10, 9, 12, 21, 16, 18, 23, 13, 5, 23, 22, 9, 20, 17 e 9. A primeira coisa que devemos fazer é ordená-los de forma crescente: 5, 9, 9, 9, 10, 12, 13, 16, 17, 18, 20, 21, 22, 23, 23 e 25. A sua média será dada por:

$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n} = \frac{252}{16} = 15,75$$

Considerando que é uma série de tamanho par, com 16 observações, sua mediana é igual a:

$$\frac{x_{(8)} + x_{(9)}}{2} = \frac{16 + 17}{2} = 16,5$$

E por fim, com os dados ordenados, facilmente é possível identificar que o número que mais se repete, ou seja, a moda, é 9.

3.4.2 Medidas de dispersão

O resumo de uma série de dados por uma única medida de tendência central esconde toda a informação sobre a variabilidade do conjunto de observações. Assim, como também salientam Bussab e Morettin (2010), as medidas mais usadas que irão nos ajudar a mensurar a dispersão dos dados em torno da média são a variância e o desvio padrão. Essas medidas de dispersão servem para medir a variação do conjunto de dados (LARSON; FARBER, 2010).

A primeira delas é a **variância**, calculada através da soma do quadrado dos desvios das observações em relação à média dividida por N , no caso de uma

população, ou, em se tratando de uma amostra, dividida por $n - 1$, segundo Doane e Seward (2014), conforme as respectivas fórmulas apresentadas a seguir:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N} \qquad S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

Contudo, de acordo com os autores citados acima, utilizamos na maioria das vezes o **desvio padrão** para representar a variabilidade do conjunto de dados, uma vez que ele é a raiz quadrada positiva da variância (veja as fórmulas abaixo) e, sendo assim, sua unidade de medida é a mesma da variável em questão. Dessa forma, ele nos ajuda a compreender como os valores individuais em um conjunto de dados variam em torno da média.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N}} \qquad S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

Vamos calcular essas medidas de dispersão juntos? Suponha que temos uma amostra-piloto sobre a idade dos funcionários de uma empresa: 35, 27, 21, 55, 18, 27, 30, 21, 24. Para calcular a variância desse conjunto de dados, sugiro a construção de uma tabela (veja a Tabela 4), iniciando com a ordenação dos dados (coluna 2) e calculando a média $\bar{x} = \frac{258}{9} = 28,67$:

Tabela 4 – Cálculo dos desvios e dos desvios ao quadrado da série *idade dos funcionários*.

i	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
1	18	$18 - 28,67 = -10,67$	$(-10,67)^2 = 113,85$
2	21	$21 - 28,67 = -7,67$	$(-7,67)^2 = 58,83$
3	21	$21 - 28,67 = -7,67$	$(-7,67)^2 = 58,83$
4	24	$24 - 28,67 = -4,67$	$(-4,67)^2 = 21,81$
5	27	$27 - 28,67 = -1,67$	$(-1,67)^2 = 2,79$
6	27	$27 - 28,67 = -1,67$	$(-1,67)^2 = 2,79$
7	30	$30 - 28,67 = 1,33$	$(1,33)^2 = 1,77$
8	35	$35 - 28,67 = 6,33$	$(6,33)^2 = 40,07$
9	55	$55 - 28,67 = 26,33$	$(26,33)^2 = 693,27$
Σ	258		994

Fonte: Elaborada pela autora.

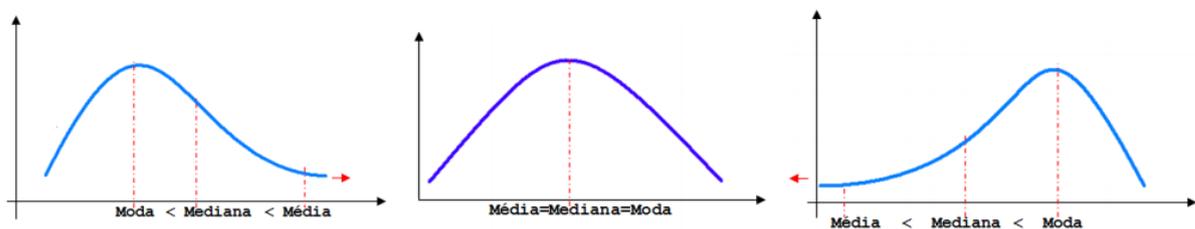
Portanto, temos:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{994}{9-1} = 124,25 \quad \text{e} \quad s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} = \sqrt{124,25} = 11,15 \text{ anos.}$$

3.4.3 Forma

A forma da distribuição pode ser avaliada por meio da observação do histograma ou da comparação entre a média e a mediana. Além disso, Doane e Seward (2014) acrescentam o seguinte: para *dados simétricos*, a média e a mediana são idênticas; para *dados assimétricos à direita* (ou de assimetria positiva), a média é superior à mediana; já para *dados assimétricos à esquerda* (ou de assimetria negativa), a média é um valor menor do que a mediana. Podemos observar essas formas de assimetria na Figura 7.

Figura 7 – Assimetria de distribuições: assimétrica à direita, simétrica e assimétrica à esquerda, respectivamente.



Fonte: Estatística (entre 2000 e 2020).

Vemos no primeiro gráfico da Figura 7 que a cauda longa do histograma aponta para a direita (a maioria dos dados se apresenta à esquerda e há poucos valores altos), justificando assim a nomenclatura “assimétrica à direita”. No último gráfico da Figura 7, que se refere à assimétrica à esquerda, a cauda aponta para a esquerda (há poucos valores baixos e grande parte dos dados se concentra à direita). Já na simétrica (gráfico central), as caudas do histograma apresentam certo equilíbrio, com valores baixos e altos balanceados. Essa descrição da Figura 7 está em consonância com o apresentado por Doane e Seward (2014).

3.4.4 Medidas de posição

As medidas de posição fornecem informações sobre a posição de valores particulares em relação ao conjunto de dados, indicando limites para proporções de observações nesse conjunto.

A **mediana**, vista anteriormente na seção 3.4.1, é a que divide o conjunto ordenado em duas partes iguais, conforme Larson e Farber (2010). Ainda segundo eles, os três quartis, Q_1 , Q_2 e Q_3 , dividem aproximadamente um conjunto de dados ordenado em quatro partes iguais. Aproximadamente $\frac{1}{4}$ dos dados está acima ou abaixo do primeiro quartil Q_1 . Em torno de metade dos dados está acima ou abaixo do segundo quartil Q_2 (que coincide com a mediana). Aproximadamente $\frac{3}{4}$ dos dados estão acima ou abaixo do terceiro quartil Q_3 .

Para especificar uma medida de posição, temos ainda os percentis e os decis, resumidos de acordo com Larson e Farber (2010):

Decis: Dividem um conjunto de dados em 10 partes iguais e são representados por $D_1, D_2, D_3, \dots, D_9$.

Percentis: Dividem um conjunto de dados em 100 partes iguais e são apresentados como $P_1, P_2, P_3, \dots, P_{99}$.

3.4.5 Medidas de associação

Para analisar o comportamento conjunto de duas variáveis quantitativas, com o objetivo de encontrar as possíveis relações ou associações entre elas, nós temos a *covariância* e o *coeficiente de correlação*. Como destacam Bussab e Morettin (2010), iremos entender a existência de uma associação como a mudança no comportamento de uma variável na presença ou não de informação sobre a segunda variável.

Você sabia?

O gráfico de dispersão, visto na seção 3.3.2, é uma ferramenta muito útil para verificar a associação entre duas variáveis quantitativas, ou entre dois conjuntos de dados.

De acordo com Doane e Seward (2014), a **covariância** mede o grau em que os valores de X e Y se alteram mutuamente e é denotada por $Cov(X, Y)$ ou σ_{XY} . Se as duas variáveis tendem a caminhar na mesma direção, sua covariância é positiva ($\sigma_{XY} > 0$) e, caso contrário, se tendem a evoluir em direções opostas, sua covariância é negativa ($\sigma_{XY} < 0$). No entanto, se as variáveis X e Y não estão relacionadas, sua covariância é zero ($\sigma_{XY} = 0$).

Para calcular a covariância para uma amostra, dados n pares de valores $(x_1, y_1), \dots, (x_n, y_n)$, Doane e Seward (2014) apresentam a seguinte fórmula:

$$Cov(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n - 1}$$

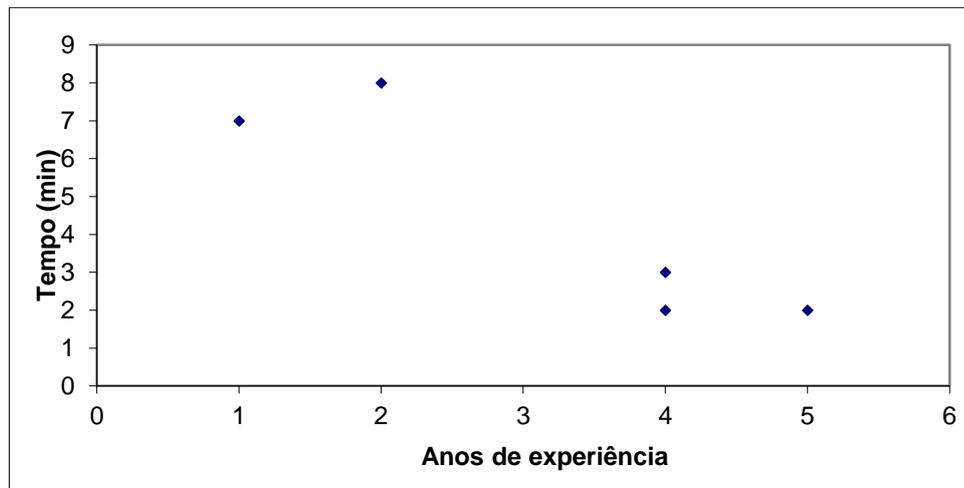
Esses autores salientam que as unidades de medida para a covariância são imprevisíveis porque as unidades de medida de X e Y podem diferir. Por isso, geralmente se utiliza o **coeficiente de correlação**, que é um valor de covariância padronizado, pois é a covariância dividida pelo produto dos desvios padrão de X e Y , e assume valores entre -1 e +1. Assim, sua fórmula de cálculo é a seguinte:

$$\rho_{XY} = \frac{Cov(X, Y)}{s_x \cdot s_y}$$

Por exemplo, suponha que há uma amostra de cinco operários de uma linha de produção e que a chefia quer saber se há uma associação entre os anos de experiência (X) e o tempo mínimo gasto na execução de certa tarefa (Y). Para isso foram levantados os seguintes dados dessa amostra: (1, 7), (2, 8), (4, 3), (4, 2), (5, 2). Faça o diagrama de dispersão e calcule o coeficiente de correlação entre X e Y .

Graficamente, esses dados são representados na Figura 8, em que se sugere uma relação negativa entre as variáveis.

Figura 8 – Gráfico de dispersão dos dados dos operários do exemplo.



Fonte: Elaborado pela autora.

Para confirmar essa possível relação negativa entre as variáveis, vamos calcular o coeficiente de correlação entre X e Y , conforme demandado no exemplo. Inicialmente sugiro organizarmos os dados e calcularmos os desvios em relação à média necessários para o cálculo, conforme Tabela 5.

Tabela 5 – Dados dos operários do exemplo e cálculo dos seus desvios, desvios ao quadrado e produto dos desvios de X e Y .

i	x_i	y_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x}) \cdot (y_i - \bar{y})$
1	1	7	-2,2	4,84	2,6	6,76	$(-2,2) \cdot (2,6) = -5,72$
2	2	8	-1,2	1,44	3,6	12,96	$(-1,2) \cdot (3,6) = -4,32$
3	4	3	0,8	0,64	-1,4	1,96	$(0,8) \cdot (-1,4) = -1,12$
4	4	2	0,8	0,64	-2,4	5,76	$(0,8) \cdot (-2,4) = -1,92$
5	5	2	1,8	3,24	-2,4	5,76	$(1,8) \cdot (-2,4) = -4,32$
Σ	16	22		10,8		33,2	-17,4

Fonte: Elaborada pela autora.

Utilizemos os valores calculados e apresentados na Tabela 5 para o coeficiente de correlação, conforme vemos a seguir:

$$\rho_{XY} = \frac{Cov(X, Y)}{s_x \cdot s_y} = \frac{\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n-1}}{\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \cdot \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}}} = \frac{\frac{-17,4}{4}}{\sqrt{\frac{10,8}{4}} \cdot \sqrt{\frac{33,2}{4}}} = \frac{-4,35}{\sqrt{2,7} \cdot \sqrt{8,3}} \cong -0,9189$$

O resultado acima demonstra uma relação inversa entre anos de experiência e tempo mínimo gasto para a execução de uma tarefa na linha de produção analisada no exemplo, confirmando assim a indicação dada pelo gráfico de dispersão (Figura 8).

Revisando

- As variáveis podem ser classificadas como qualitativas (dados por atributos) ou quantitativas (dados numéricos). As qualitativas são ainda classificadas como nominais (quando não é possível sua ordenação) e ordinais (quando existe uma ordem nos seus resultados).
- Uma distribuição de frequência ordena os dados em classes e pode incluir ainda frequências relativas, frequências acumuladas, ponto médio, densidade de frequência e densidade de frequência relativa.
- Para variáveis qualitativas devem ser utilizados os gráficos em barra e em composição de setores (“pizza”). Na apresentação de dados relativos a variáveis quantitativas são recomendados os gráficos de dispersão, em barras e o histograma.
- Fórmulas das medidas descritivas mais utilizadas:

$$\text{Média: } \bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

$$\text{Média ponderada: } \bar{X} = \frac{\sum_{i=1}^n x_i p_i}{\sum_{i=1}^n p_i}$$

$$\text{Variância da amostra: } s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

Desvio padrão da amostra: É a raiz quadrada da variância, portanto

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

$$\text{Covariância: } Cov(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

$$\text{Coeficiente de correlação: } \rho_{XY} = \frac{Cov(X, Y)}{s_x \cdot s_y}$$

Saiba mais

Livro: **Estatística aplicada à administração e economia**, de David Anderson, Dennis Sweeney e Thomas Williams.

ANDERSON, David. R.; SWEENEY, Dennis J.; WILLIAMS, Thomas A. **Estatística aplicada à administração e economia**. 2. ed. São Paulo: Cengage Learning, 2011.

Vídeo: **Distribuição de frequência estatística**, de Marcos Murakami.

DISTRIBUIÇÃO de frequência estatística. Publicado pelo canal Matemática Rapidola com Marcos Murakami. Disponível em <https://youtu.be/451SxriwvM>. Acesso em: 20 nov. 2020.

Livro: **Statistics for management and economics**, de Gerald Keller.

KELLER, Gerald. **Statistics for management and economics**. 7. ed. Mason, OH, USA: Thomson South-Western, 2005.

Referências

ANDERSON, David. R.; SWEENEY, Dennis J.; WILLIAMS, Thomas A. **Estatística aplicada a administração e economia**. 2. ed. São Paulo: Cengage Learning, 2011.

BUSSAB, Wilton de O.; MORETTIN, Pedro A. **Estatística básica**. 6. ed. São Paulo: Saraiva, 2010.

DOANE, David P.; SEWARD, Lori E. **Estatística aplicada à administração e economia**. 4. ed. Porto Alegre: Mc Graw Hill, 2014.

ESTATÍSTICA computacional – Aula 5. [Ponta Grossa, PR]: [UEPG], [entre 2000 e 2020]. Disponível em: <https://ead.uepg.br/apl/sigma/assets/editais/PS0027E0035.pdf>. Acesso em: 20 nov. 2020.

LARSON, Ron; FARBER, Betsy. **Estatística aplicada**. 4. ed. São Paulo: Pearson, 2010.

4

Probabilidade e variáveis aleatórias

Métodos Quantitativos Aplicados
às Ciências Contábeis 2

Lidia Maria Alves Rodella Amorim

Objetivos da aprendizagem

- 1 Aplicar as definições e as regras de probabilidades;
- 2 Usar o teorema de Bayes para calcular probabilidades;
- 3 Definir variável aleatória discreta e contínua;
- 4 Calcular valor esperado e variância.

Introdução

Vimos no capítulo anterior que uma estatística é uma medida que descreve um conjunto de observações. Larson e Farber (2010) salientam que, com os dados coletados e descritos, você pode usar o resultado para escrever resumos, elaborar suas conclusões e tomar decisões. Contudo, com as regras de probabilidade é possível ir além e responder a perguntas como:

- Quais são as chances de que a receita do próximo mês seja maior que a do mês anterior?
- Quais são as chances de que o novo sistema de produção proposto diminua a taxa de defeitos de fabricação de determinados produtos?
- Qual é a chance de que uma pessoa ganhe sozinha na Mega-Sena?

Com a probabilidade é possível compreender e quantificar a incerteza sobre o futuro. Assim, veremos a seguir regras/propriedades sobre probabilidade que nos ajudarão no processo de **inferência estatística**.

Aprenderemos também neste capítulo sobre variáveis aleatórias, a fim de auxiliar na compreensão e estudo dos modelos probabilísticos discretos e contínuos (assunto do próximo capítulo), que são utilizados para descrever as características essenciais de um processo estocástico e para orientar na tomada de decisões ou fazer previsões.

4.1 Probabilidade

A **probabilidade** é a área de estudo que nos permite compreender e quantificar a incerteza sobre o futuro. As regras da probabilidade são úteis para estabelecer uma ligação entre o que sabemos e o que é desconhecido em relação ao futuro, como afirmam Doane e Seward (2014).

A probabilidade também desempenha um papel crítico na inferência estatística. Ela fornece uma conexão entre a população e a amostra. Para tirarmos conclusões a respeito de um parâmetro, precisamos ter informações a respeito de sua **distribuição de probabilidade**.

Para discutir probabilidade, é preciso considerar que um **experimento** é qualquer processo que permite ao pesquisador fazer observações. Além disso, os experimentos podem ser **determinísticos** (quando repetidos nas mesmas condições, conduzem ao mesmo resultado, como a temperatura do ponto de ebulição da água, por exemplo) ou **aleatórios** (mesmo que realizados sob condições idênticas, não se pode prever, a princípio, o resultado particular que irá ocorrer, embora se saiba o conjunto dos possíveis resultados, como ocorre no lançamento de uma moeda ou de um dado, por exemplo).

Em se tratando de probabilidade, estamos interessados no **experimento aleatório**, que é uma ação ou processo que leva a um dos vários resultados possíveis, mas cujo resultado exato nós não conhecemos previamente. O conjunto de todos os resultados possíveis (representado por S) é o **espaço amostral**, como destacam Doane e Seward (2014). Veja alguns exemplos no quadro abaixo:

Quadro 1 – Exemplos de experimentos e seus respectivos espaços amostrais

Experimento	Espaço amostral
Lançar uma moeda	$S: \{cara, coroa\}$
Lançar um dado	$S: \{1, 2, 3, 4, 5, 6\}$
Duração de uma lâmpada	$S: \{t \mid t \geq 0\}$

Fonte: Elaborado pela autora.

De acordo com os supracitados autores, um **evento** é qualquer subconjunto do espaço amostral. Sendo assim, para o lançamento de um dado (experimento aleatório), um exemplo de evento pode ser “números ímpares” (evento A), ou seja, $A: \{1, 3, 5\}$.

Assim, a **probabilidade** de um evento – $P(E)$ – é um número que mede a possibilidade relativa de sua ocorrência (DOANE; SEWARD, 2014) e é encontrada a partir da soma das probabilidades de cada resultado do evento. No exemplo do evento “números ímpares”, apresentado acima, temos: $P(A) = P(1) + P(3) + P(5)$.

Qualquer probabilidade apresenta as seguintes propriedades:

- a) As probabilidades são sempre valores entre 0 e 1;
- b) A soma das probabilidades de todos os resultados do espaço amostral é igual a 1.

Como atribuir probabilidades aos elementos que compõem o espaço amostral? Doane e Seward (2014) afirmam que existem três modos distintos de se atribuir probabilidade e que devem ser considerados separadamente:

Quadro 2 – Três abordagens da probabilidade

Abordagem	Como é atribuída?	Exemplo
Empírica	Estimada a partir da frequência observada do resultado	Existe uma probabilidade de 2% de se encontrar gêmeos em um parto escolhido aleatoriamente.
Clássica	Conhecida <i>a priori</i> pela natureza do experimento	Existe uma probabilidade de 50% de se obter cara no lançamento de moedas.
Subjetiva	Baseada em opinião informada ou julgamento	Existe uma probabilidade de 75% de que a Inglaterra adote o Euro como unidade monetária até 2012.

Fonte: Doane e Seward (2014, p. 173).

Segundo os autores, a abordagem empírica pode ser utilizada para atribuir probabilidades aos dados empíricos coletados por meio de observações ou experimentos. Para isso, basta contar as frequências (f) dos resultados observados

e dividi-las pelo número de observações (n). Portanto, a probabilidade será estimada por f/n .

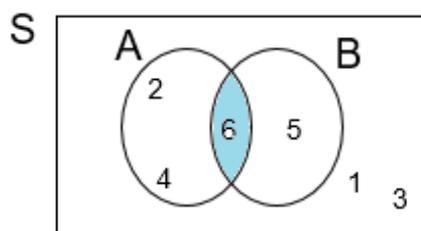
Atenção!

Probabilidades podem ser apresentadas em forma de frações, números decimais ou porcentagens.

Vejamos agora como estabelecer **relações** entre os eventos e como determinar suas probabilidades, considerando: um espaço amostral S ; A e B como dois eventos do espaço amostral S ; e o experimento de lançamento de um dado.

- **Interseção:** É o conjunto formado por elementos que pertencem aos dois eventos (A e B), em geral denotado por $A \cap B$. Exemplo: $A = \text{números pares}$ e $B = \text{números maiores que 4}$, então $A \cap B = \{6\}$. Logo, a probabilidade dos dois eventos ocorrerem simultaneamente é a probabilidade de que, após o lançamento de um dado, o resultado seja igual a 6. Assim, $P(A \cap B) = P(6) = 1/6$. O diagrama da Figura 1 representa essa relação.

Figura 1 – Diagrama de Venn da interseção dos eventos A e B

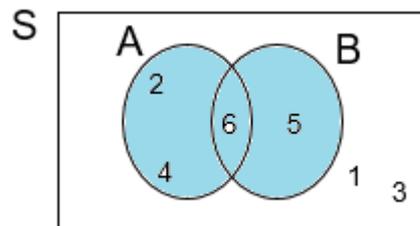


Fonte: Elaborado pela autora.

- **União:** É o conjunto formado por elementos que pertencem a pelo menos um dos dois eventos (A **ou** B). Exemplo: $A = \text{números pares}$; $B = \text{números maiores que 4}$. A união desses dois eventos, representada por $A \cup B$, é formada pelo conjunto $\{2, 4, 5, 6\}$, conforme visualizado na Figura 2. A probabilidade de ocorrer pelo menos

um dos dois eventos – ou seja, $P(A \cup B)$ – é dada pela soma das probabilidades de que cada um desses resultados ocorra após o lançamento de um dado. Assim, $P(A \cup B) = 4/6$.

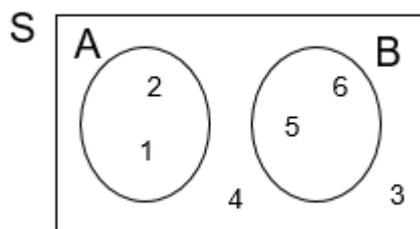
Figura 2 – Diagrama de Venn da união dos eventos A e B



Fonte: Elaborado pela autora.

- **Eventos mutuamente excludentes ou disjuntos:** Quando não há elementos em comum entre os eventos e, portanto, eles nunca ocorrem simultaneamente, de modo que a interseção dos eventos é um conjunto vazio. Exemplo: $A = \text{números menores que } 3$; $B = \text{números maiores que } 4$. O resultado da relação desses eventos é visto na Figura 3, na qual podemos observar que a ocorrência de um evento exclui a do outro. (DOANE; SEWARD, 2014).

Figura 3 – Diagrama de Venn dos eventos disjuntos A e B

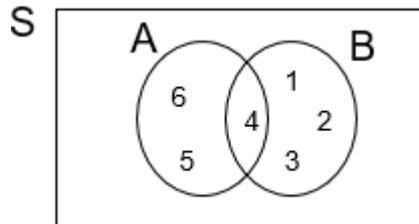


Fonte: Elaborado pela autora.

- **Eventos coletivamente exaustivos:** Quando todos os elementos do espaço amostral pertencem ou a A , ou a B , ou a ambos, isto é, quando a união dos eventos corresponde a todo o espaço amostral S . Nesse caso, pelo menos um dos eventos

tem necessariamente que ocorrer. Exemplo: $A = \text{números maiores que } 3$; $B = \text{números menores que } 5$. Aqui, todos os possíveis resultados do lançamento de um dado estão compreendidos por um desses eventos, conforme pode ser claramente visualizado na Figura 4.

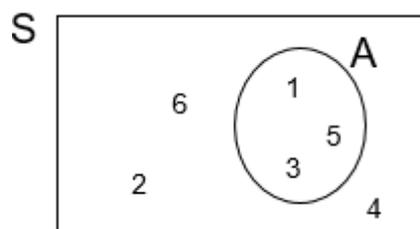
Figura 4 – Diagrama de Venn dos eventos coletivamente exaustivos A e B



Fonte: Elaborado pela autora.

- **Complemento de um evento:** É o conjunto formado pelos elementos que pertencem ao espaço amostral S , mas não pertencem ao evento A . Esse conjunto é denotado por \bar{A} . Exemplo: $A = \text{conjunto dos números ímpares} = \{1, 3, 5\}$. Dessa forma, $\bar{A} = \{2, 4, 6\}$, ou seja, o complemento do evento A é formado pelos números pares, conforme vemos na Figura 5. Como A e \bar{A} juntos constituem o espaço amostral, suas probabilidades somam 1. Portanto, como destacam Doane e Seward (2014), a probabilidade do complemento de A é encontrada subtraindo-se a probabilidade de A de 1, ou seja, $P(\bar{A}) = 1 - P(A)$. No exemplo dado, $P(\bar{A}) = 1 - \frac{1}{2} = \frac{1}{2}$.

Figura 5 – Diagrama de Venn do complemento do evento A



Fonte: Elaborado pela autora.

Para apresentar os conceitos de probabilidade marginal, probabilidade condicional, independência entre eventos e aplicar a interseção e a união, nós vamos supor que: **a)** estamos estudando duas características (variáveis) dos indivíduos de

uma população, a cor dos olhos e a cor do cabelo; e **b)** foram coletados dados de uma amostra de 200 pessoas, conforme a Tabela 1.

Tabela 1 – Dados da amostra sobre cor do cabelo e cor dos olhos

Cabelo (X)/Olhos (Y)	Azul	Castanho	Verde	Total
Preto	3	62	13	78
Loiro	32	9	13	54
Ruivo	11	33	24	68
Total	46	104	50	200

Fonte: Elaborada pela autora.

Para estimar as probabilidades utilizando a abordagem empírica, vamos agora dividir a frequência pela quantidade total de pessoas observadas (200), resultando na **probabilidade conjunta** (referente à interseção de dois eventos) apresentada na Tabela 2.

Tabela 2 – Probabilidade conjunta sobre cor do cabelo e cor dos olhos da amostra

Cabelo (X)/Olhos (Y)	Azul	Castanho	Verde	P(X)
Preto	0,015	0,310	0,065	0,390
Loiro	0,160	0,045	0,065	0,270
Ruivo	0,055	0,165	0,120	0,340
P(Y)	0,230	0,520	0,250	1

Fonte: Elaborada pela autora.

Observe que as células correspondentes a cada duas características (cor dos olhos e cor do cabelo) se referem à probabilidade da interseção. Portanto, a probabilidade de se selecionar alguém com cabelo ruivo e olhos castanhos, ou seja, $P(\text{ruivo} \cap \text{castanho})$, é igual a 0,165 ou 16,5% (multiplicamos por cem para obter o valor em termos percentuais). Da mesma forma, $P(\text{verde} \cap \text{loiro}) = 0,065$ ou 6,5% e $P(\text{preto} \cap \text{verde}) = 0,065$ ou 6,5%.

Quando só estamos interessados em uma característica, observamos a **probabilidade marginal**, isto é, a soma de todas as probabilidades de interseção entre a característica de interesse e todas as possibilidades da outra variável. Na Tabela 2, as probabilidades marginais aparecem na última coluna, referente às cores dos cabelos, e na última linha, referente às cores dos olhos. Portanto, de acordo com a referida tabela, a probabilidade de selecionarmos, por exemplo, alguém de olhos verdes, é de 0,25 ou 25%, e de cabelo ruivo, é de 0,34 ou 34%.

E se, a partir dos dados da Tabela 2, quisermos saber a probabilidade de selecionarmos alguém com cabelo preto ou olhos castanhos? Nesse caso, estamos tratando da probabilidade da união de dois eventos, ou seja, de que pelo menos um dos eventos ocorra. Portanto, nesse caso, ou a pessoa tem cabelo preto ou olhos castanhos ou os dois. Há dois modos de encontrar essa probabilidade, $P(\text{preto} \cup \text{castanho})$: **a)** somar a probabilidade da interseção entre essas duas características e todas as possibilidades da outra variável; ou **b)** somar as probabilidades marginais e subtrair a probabilidade da interseção entre elas, conforme apresentado a seguir, respectivamente.

$$\mathbf{a)} \quad P(\text{preto} \cup \text{castanho}) = 0,015 + 0,31 + 0,065 + 0,045 + 0,165 = 0,6$$

ou

$$\mathbf{b)} \quad P(\text{preto} \cup \text{castanho}) = P(\text{preto}) + P(\text{castanho}) - P(\text{preto} \cap \text{castanho})$$

isto é,

$$P(\text{preto} \cup \text{castanho}) = 0,39 + 0,52 - 0,31 = 0,6$$

Outro tipo de probabilidade muito utilizado é a **probabilidade condicional**, definida por Larson e Farber (2010, p. 119) como “a probabilidade de um evento ocorrer, dado que outro evento já tenha ocorrido”. Segundo os mesmos autores, “a probabilidade condicional de o evento B ocorrer, dado que o evento A tenha ocorrido, é denotada por $P(B|A)$ e lê-se ‘probabilidade de B , dado A ’” (2010, p. 119). Para calcular tal

probabilidade, divide-se a probabilidade da interseção entre os eventos pela probabilidade marginal da condição, ou seja:

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

Um exemplo: diante da Tabela 2, se sabemos que uma pessoa tem olhos azuis, a probabilidade de ela ter cabelos loiros será:

$$P(\text{loiro}|\text{azul}) = \frac{P(\text{loiro} \cap \text{azul})}{P(\text{azul})} = \frac{0,16}{0,23} = 0,6957 \text{ ou } 69,57\%$$

Observe que, nesse caso, não olhamos para todas as pessoas, apenas para as que possuem olhos azuis.

Diante do exposto, podemos resumir as relações da seguinte forma:

- Probabilidade da união: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Probabilidade condicional: $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- Probabilidade da interseção, a partir da probabilidade condicional:

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

Acontecem, ainda, alguns experimentos em que “um evento não afeta a probabilidade de outro”, conforme destacam Larson e Farber (2010, p. 120). Por exemplo, se você joga um dado e uma moeda, o resultado do lançamento do dado não influencia a probabilidade de que a moeda caia em coroa. Esses dois eventos são independentes. De acordo com esses autores, dois eventos são **independentes** se a realização de um deles não afeta a probabilidade do outro evento ocorrer. Dois eventos, A e B , são independentes se, e somente se:

$$P(B|A) = P(B)$$

ou

$$P(A|B) = P(A)$$

ou, conforme Bussab e Morettin (2010, p. 113), se

$$P(A \cap B) = P(A).P(B)$$

Exemplo: Retiramos uma carta de um baralho com 52 cartas, contendo 13 cartas de cada naipe (copas, espadas, ouro e paus). Sejam os eventos A (a carta é um ás) e B (a carta é de espadas). Podemos afirmar que A e B são independentes?

Passo 1

Encontrar as probabilidades marginais dos eventos A e B :

- Evento A : a carta é um ás. Temos que são quatro cartas de ás em um baralho com 52 cartas, portanto: $P(A) = \frac{4}{52}$;
- Evento B : a carta é de espadas. Assim, temos 13 cartas de espadas no baralho de 52 cartas, logo: $P(B) = \frac{13}{52}$.

Passo 2

Determinar a probabilidade da interseção dos eventos A e B :

- A interseção dos eventos ($A \cap B$) é composta por cartas que são às e de espadas. Portanto, considerando as informações no enunciado, temos um ás de espadas no baralho de 52 cartas. Assim: $P(A \cap B) = \frac{1}{52}$.

Passo 3

Verificar se $P(A \cap B) = P(A).P(B)$.

- $P(A).P(B) = \frac{4}{52} \cdot \frac{13}{52} = \frac{1}{52} = P(A \cap B)$. Com esse resultado, podemos afirmar que A e B são eventos independentes.

De acordo com Bussab e Morettin (2010, p. 116), “uma das relações mais importantes envolvendo probabilidades condicionais é dada pelo Teorema de Bayes”, utilizado para revisar as probabilidades quando se obtêm novas informações.

Segundo Anderson, Sweeney e Williams (2011, p. 155), frequentemente a análise é iniciada com estimativas da probabilidade inicial (*a priori*), para eventos de interesse específico e, a partir de fontes como uma amostra ou teste de produto, por exemplo, obtemos informações adicionais sobre os eventos. Considerando essas novas informações, os valores da probabilidade inicial são atualizados calculando as probabilidades revisadas, chamadas de probabilidades *a posteriori*. O teorema de Bayes é aplicado para a atualização das probabilidades a partir das novas informações.

Formalmente, considerando B_i eventos mutuamente excludentes e coletivamente exaustivos, para todo $i = 1, 2, \dots, n$, temos:

$$P(A) = \frac{P(B_i).P(A|B_i)}{\sum_{i=1}^k P(B_i).P(A|B_i)}$$

Vejamos como o Teorema de Bayes é aplicado no exemplo apresentado por Bussab e Morettin (2010, p. 116):

“ Temos cinco urnas, cada uma com seis bolas. Duas dessas urnas (tipo C1) têm 3 bolas brancas, duas outras (tipo C2) têm 2 bolas brancas, e a última urna (tipo C3) tem 6 bolas brancas. Escolhemos uma urna ao acaso e dela retiramos uma bola. Qual a probabilidade de a urna escolhida ser do tipo C3, sabendo-se que a bola sorteada é branca? ”

As informações dadas na questão podem ser formalmente apresentadas da seguinte forma:

$$P(C1) = \frac{2}{5}; P(C2) = \frac{2}{5}; P(C3) = \frac{1}{5}$$

$$P(B|C1) = \frac{1}{2}; P(B|C2) = \frac{1}{3}; P(B|C3) = 1$$

Para responder à questão temos que calcular:

$$P(C3|B) = \frac{P(C3 \cap B)}{P(B)} = \frac{P(C3) \cdot P(B|C3)}{P(B)}$$

Como não temos $P(B)$, vamos calculá-la da seguinte forma:

$$P(B) = P(C1 \cap B) + P(C2 \cap B) + P(C3 \cap B)$$

Portanto:

$$P(B) = P(C1) \cdot P(B|C1) + P(C2) \cdot P(B|C2) + P(C3) \cdot P(B|C3)$$

$$= \left(\frac{2}{5} \cdot \frac{1}{2}\right) + \left(\frac{2}{5} \cdot \frac{1}{3}\right) + \left(\frac{1}{5} \cdot 1\right) = \frac{8}{15}$$

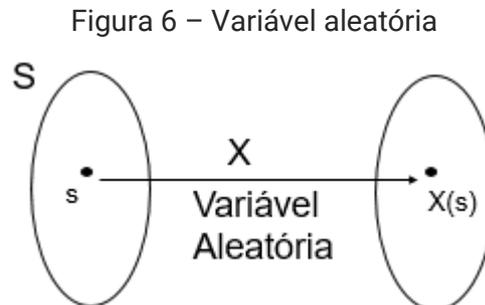
Assim, temos:

$$P(C3|B) = \frac{P(C3) \cdot P(B|C3)}{P(B)} = \frac{\frac{1}{5} \cdot 1}{\frac{8}{15}} = \frac{3}{8}$$

4.2 Variáveis aleatórias

Seja S um espaço amostral associado a um experimento aleatório. **Variável aleatória** (v.a.) é definida por Larson e Farber (2010), como uma função ou regra que associa a cada elemento $s \in S$ um número real $X(s)$, isto é, atribui um valor numérico

para cada resultado s no espaço amostral S de um experimento aleatório, conforme representado na Figura 6.



Fonte: Elaborado pela autora.

Vejamos um exemplo de variável aleatória: considere um experimento aleatório de lançamento de duas moedas, sendo C igual a coroa e Ca igual a cara. Seu espaço amostral é $S = \{CC, CCa, CaC, CaCa\}$ com a probabilidade de cada evento ocorrer igual a $\frac{1}{4}$, resultado da multiplicação da probabilidade de sair cara ou coroa em uma moeda ($\frac{1}{2}$) pela probabilidade de sair cara ou coroa na outra moeda ($\frac{1}{2}$). Dado que a variável aleatória X é o número de caras obtidas nas duas moedas, temos:

- $x_1 = 0 \rightarrow$ Correspondente ao evento (CC)
- $x_2 = 1 \rightarrow$ Correspondente aos eventos (CCa) e (CaC)
- $x_3 = 2 \rightarrow$ Correspondente ao evento $(CaCa)$

Para cada valor de X tem-se uma probabilidade associada, denominada **função de probabilidade** e denotada por $P(X)$. O conjunto de pares $[x_i, P(x_i)]$ é chamado de **distribuição de probabilidade** de X e deve seguir as regras de probabilidade:

- $0 \leq P(x_i) \leq 1$ para todo i ;
- $\sum_{i=1}^n P(x_i) = 1$.

Diante do exposto, no exemplo do lançamento de duas moedas, teremos a seguinte distribuição de probabilidade:

Tabela 3 – Distribuição de probabilidade da v.a. X

X	$P(X)$
0	$\frac{1}{4}$ – Correspondente ao evento (CC)
1	$\frac{1}{4} + \frac{1}{4} = \frac{1}{2}$ – Correspondente à ocorrência dos eventos (CCa) ou (CaC)
2	$\frac{1}{4}$ – Correspondente ao evento $(CaCa)$

Fonte: Elaborada pela autora.

Os autores Larson e Farber (2010) salientam que a palavra **aleatória** indica que x é determinado pelo acaso e que existem dois tipos de variáveis aleatórias: discreta e contínua.

a) As variáveis aleatórias discretas apresentam valores que podem ser enumerados e resultam de alguma forma de contagem, como, por exemplo, o número de ligações realizadas em um consultório médico em um dia.

b) As variáveis aleatórias contínuas não são enumeráveis, podem assumir qualquer valor dentro de um intervalo de variação, inclusive valores numéricos decimais, e resultam de uma mensuração, como, por exemplo, o número de horas gastas em ligações em um consultório médico em um dia.

Vejamos um exemplo apresentado por Bussab e Morettin (2010, p. 164), para uma melhor compreensão da diferença entre variáveis aleatórias discretas e contínuas.

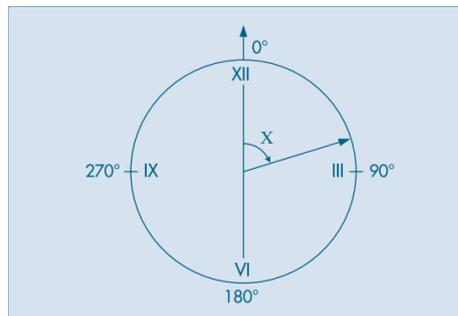
a) Variável aleatória discreta: o ponteiro dos segundos de um relógio mecânico pode parar a qualquer instante. A variável aleatória X indica o ângulo que o ponteiro forma com um eixo imaginário que passa pelo centro do mostrador e pelo número XII, conforme a Figura 7. Uma vez que o ponteiro dá 1 “salto” por segundo e que são 60 saltos a cada volta completa do ponteiro no mostrador do relógio, teremos 6° por salto ($360^\circ/60$). A distribuição de probabilidade de X será:

Tabela 4 – Distribuição de probabilidade da v.a. ângulo do ponteiro dos segundos de um relógio mecânico

x	0°	6°	12°	18°	...	348°	354°
$P(x)$	$1/60$	$1/60$	$1/60$	$1/60$...	$1/60$	$1/60$

Fonte: Bussab e Morettin (2010, p. 163).

Figura 7 – Ilustração de uma v.a. X discreta



Fonte: Bussab e Morettin (2010, p. 164).

b) Variável aleatória contínua: considere o mesmo problema do item anterior com um relógio elétrico, em que o ponteiro dos segundos move-se continuamente, sendo X o ângulo formado pelo ponteiro dos segundos com o eixo imaginário que passa pelo centro do mostrador e pelo número XII. Nesse caso, X pode assumir qualquer valor do intervalo $[0, 360) = \{x \in IR: 0 \leq x < 360\}$. Como existem infinitos pontos nos quais o ponteiro pode parar, cada um com a mesma probabilidade, então cada ponto tem a probabilidade de ocorrer igual a 0, ou seja, $P(X = x_i) = 0$. Contudo, podemos encontrar a probabilidade de que X esteja compreendido entre dois valores quaisquer. Assim:

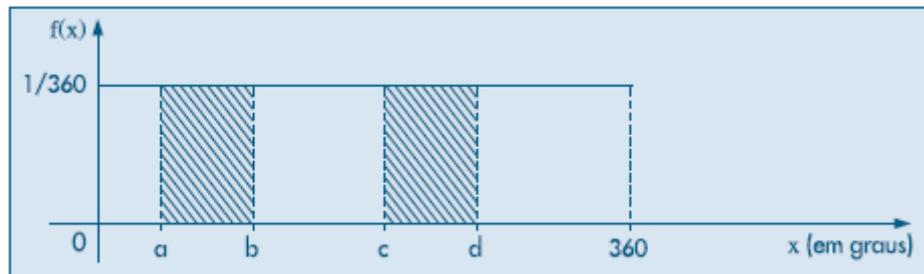
$$P(0^\circ \leq X \leq 90^\circ) = \frac{1}{4}$$

e

$$P(120^\circ \leq X \leq 150^\circ) = \frac{1}{12}$$

Nesse caso, temos uma distribuição uniforme contínua e, dados dois números a e b , a probabilidade de $X \in [a, b)$ é $P(a \leq X < b) = \frac{b-a}{360^\circ}$. Podemos ainda construir o histograma dessa v.a. X , conforme a Figura 8.

Figura 8 – Histograma da v.a. contínua ângulo do ponteiro do relógio elétrico



Fonte: Bussab e Morettin (2010, p. 165).

A figura acima corresponde à seguinte função $f(x)$, denominada **função de densidade de probabilidade** (f.d.p.):

$$f(x) = \begin{cases} 0, & \text{se } x < 0^\circ \\ \frac{1}{360}, & \text{se } 0^\circ \leq x < 360^\circ \\ 0, & \text{se } x \geq 360^\circ \end{cases}$$

Temos então que a distribuição de probabilidade para uma variável aleatória contínua X é representada por uma curva e a probabilidade de que X assumira um valor no intervalo de a e b (sendo $a < b$) é dada pela área sob a curva limitada por a e b . Como vimos no Capítulo 2, essa área é expressa matematicamente por meio da integral definida da função $f(x)$ no intervalo de a até b . Logo,

$$P(a \leq X < b) = \int_a^b f(x)dx = \int_a^b \frac{1}{360} dx = \frac{b-a}{360}.$$

No caso das v.a. contínuas, é importante observar que:

1. $P(X = x_0) = \int_{x_0}^{x_0} f(x)dx = 0$, e isso implica em $P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b)$;
2. A função de densidade de probabilidade de uma v.a. definida para todos os valores de x , tal que $a \leq x \leq b$, satisfaz às seguintes condições:
 - a. $f(x) \geq 0 \forall x$ entre a e b ;
 - b. $\int_a^b f(x)dx = 1$.
3. A f.d.p. não representa a probabilidade de coisa alguma. Ela somente produzirá uma probabilidade quando for integrada entre dois limites. Assim, como vimos, a probabilidade será a área sob a curva entre esses limites.

Vejam os outro exemplo, também apresentado por Bussab e Morettin (2010), em que temos $f(x) = 2x$, para $0 \leq x \leq 1$ e 0 para outros valores de x . Primeiramente, podemos verificar que $f(x)$ é uma função de densidade de probabilidade, pois atende às seguintes propriedades:

a) $f(x) \geq 0 \forall x$ entre 0 e 1 \rightarrow **Ok**

b) $\int_0^1 2x dx = \frac{2x^2}{2} \Big|_0^1 = 1 \rightarrow$ **Ok**

Assim, temos que a $P\left(0 \leq x \leq \frac{1}{2}\right) = \int_0^{1/2} 2x dx = x^2 \Big|_0^{1/2} = 1/4$.

4.2.1 Valor médio ou valor esperado (esperança matemática) de uma variável aleatória

O valor esperado, $E(X)$, de uma variável aleatória discreta X é uma média ponderada dos diferentes valores de X com pesos dados pelas respectivas probabilidades. Portanto, é definido como:

$$E(X) = \sum_{i=1}^n x_i \cdot p(x_i)$$

onde, x_1, x_2, \dots, x_n são os valores possíveis de X e $p(x_i) = P(X = x_i)$.

No experimento de lançamento de um dado, considerando X o número de pontos obtidos, temos como distribuição de probabilidade os lados do dado $x_1 = 1, x_2 = 2, \dots, x_6 = 6$ e suas respectivas probabilidades iguais a $1/6$. O valor esperado de X será:

$$E(X) = \left(1 \cdot \frac{1}{6}\right) + \left(2 \cdot \frac{1}{6}\right) + \left(3 \cdot \frac{1}{6}\right) + \left(4 \cdot \frac{1}{6}\right) + \left(5 \cdot \frac{1}{6}\right) + \left(6 \cdot \frac{1}{6}\right) = 3,5$$

Destaca-se que $E(X)$ não é o resultado que podemos esperar quando X for observada uma única vez. Afinal, a $E(X)$ é a média aritmética dos resultados quando X for observada um grande número de vezes.

Para uma variável aleatória contínua, o valor esperado de X é definido como:

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx$$

onde, $f(x)$ é a função de densidade de probabilidade.

Voltando ao exemplo de v.a. contínua anterior, no qual $f(x) = 2x$, para $0 \leq x \leq 1$ e 0 para outros valores de x , seu valor médio ou valor esperado será:

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx = \int_0^1 2x^2 dx = 2 \frac{x^3}{3} \Big|_0^1 = \frac{2}{3}$$

4.2.2 Variância de uma variável aleatória

A variância de uma variável aleatória X é dada por:

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

A variância de uma variável aleatória discreta é:

$$V(X) = \sum_{i=1}^n [x_i - E(X)]^2 p(x_i)$$

E a variância de uma variável aleatória contínua é:

$$V(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx$$

Contudo, pelo teorema encontrado a partir do desenvolvimento da fórmula $V(X) = E[X - E(X)]^2$, podemos calcular a variância de X , em ambos os casos, como:

$$V(X) = E(X^2) - E^2(X)$$

Você sabia?

O desvio padrão de X é definido como a raiz quadrada da variância.

Vejamos agora um exemplo de cálculo da variância de uma variável aleatória discreta. Considere uma variável aleatória com a distribuição de probabilidade abaixo. Vamos calcular o valor esperado e a variância de X .

$$x_1 = 1; P(x_1) = \frac{1}{4}; x_2 = 2; P(x_2) = \frac{1}{2}; x_3 = 3; P(x_3) = \frac{1}{4}.$$

$$E(X) = \left(1 \cdot \frac{1}{4}\right) + \left(2 \cdot \frac{1}{2}\right) + \left(3 \cdot \frac{1}{4}\right) = 2$$

$$V(X) = \sum_{i=1}^3 (x_i - 2)^2 \cdot p(x_i) = \left(1 \cdot \frac{1}{4}\right) + \left(0 \cdot \frac{1}{2}\right) + \left(1 \cdot \frac{1}{4}\right) = \frac{1}{2}$$

Calculando a variância através da aplicação do teorema, temos que calcular primeiro a esperança de X^2 e, depois, substituí-la na fórmula do teorema junto com o quadrado do valor esperado de X :

$$E(X^2) = \sum_{i=1}^3 x_i^2 \cdot p(x_i) = \left(1 \cdot \frac{1}{4}\right) + \left(4 \cdot \frac{1}{2}\right) + \left(9 \cdot \frac{1}{4}\right) = \frac{9}{2}$$

$$V(X) = E(X^2) - E^2(X) = \frac{9}{2} - 2^2 = \frac{1}{2}$$

Agora vamos calcular a variância da variável aleatória contínua $f(x) = 2x$, para $0 \leq x \leq 1$ (a mesma variável da seção 4.4.1 acima), onde encontramos que a $E(X) = 2/3$.

$$V(X) = \int_0^1 [x - E(X)]^2 f(x) dx = \int_0^1 \left(x - \frac{2}{3}\right)^2 \cdot 2x dx = \int_0^1 \left(2x^3 - \frac{8x^2}{3} + \frac{8x}{9}\right) dx =$$

$$= \left(\frac{2x^4}{4} - \frac{8x^3}{9} + \frac{8x^2}{18}\right) \Big|_0^1 = \frac{2}{4} - \frac{8}{9} + \frac{8}{18} = \frac{1}{18}$$

Podemos também aplicar o teorema e, para isso, devemos primeiro calcular $E(X^2)$, conforme vemos a seguir:

$$E(X^2) = \int_0^1 x^2 \cdot 2x dx = \frac{x^4}{2} \Big|_0^1 = \frac{1}{2}$$

$$V(X) = E(X^2) - E^2(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

Observe que, em geral, o uso do teorema simplifica os cálculos necessários para se encontrar a variância, especialmente nos casos de variável aleatória contínua.

Revisando

- O espaço amostral de um experimento aleatório descreve todos os seus possíveis resultados.
- Uma probabilidade empírica é baseada em frequências relativas.
- O complemento de um evento são todos os resultados do experimento, exceto os do evento.
- A interseção de dois eventos corresponde apenas aos resultados que estão em ambos, é denotada por $P(A \cap B)$ e pode ser calculada por: $P(A \cap B) = P(A|B) \cdot P(B)$.
- A união de dois eventos é dada por todos os eventos que estão contidos em um deles ou em ambos e sua fórmula é igual a: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- A probabilidade condicional de um evento é a probabilidade de que ele ocorra dado que outro evento já tenha ocorrido e é calculada assim: $P(A|B) = \frac{P(A \cap B)}{P(B)}$.
- Uma variável aleatória atribui um valor numérico a cada evento no espaço amostral de um experimento aleatório.
- Uma variável aleatória discreta tem um número contável de valores diferentes.
- Em uma distribuição de probabilidade discreta, as probabilidades devem estar entre 0 e 1 e juntas devem somar 1.
- Valor esperado (média) de uma v.a. discreta: $E(X) = \sum_{i=1}^n x_i \cdot p(x_i)$
- Variância de uma v.a. discreta: $V(X) = \sum_{i=1}^n [x_i - E(X)]^2 p(x_i)$
- A função de densidade de probabilidade de uma variável aleatória contínua é uma curva e as probabilidades são as áreas sob a curva, que devem ser calculadas por meio da integral definida.
- Valor esperado (média) de uma v.a. contínua: $E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx$

- Variância de uma v.a. contínua: $V(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x) dx$
- Teorema para o cálculo da variância de variáveis aleatórias: $V(X) = E(X^2) - E^2(X)$

Saiba mais

Livro: **Estatística aplicada à administração e economia**, de David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm e James J. Cochran.

ANDERSON, David. R.; SWEENEY, Dennis J.; WILLIAMS, Thomas A. **Estatística aplicada à administração e economia**. 2. ed. São Paulo: Cengage Learning, 2011.

Artigo: **Statistics for management and economics**, de Gerald Keller.

KELLER, Gerald. **Statistics for management and economics**. 7. ed. Mason, OH, USA: Thomson South-Western, 2005.

Livro: **Notas de Aula – Probabilidade: teoria e exercícios**, de Élcio Lebensztayn e Cristian Coletti.

LEBENSZTAYN, Élcio; COLETTI, Cristian F. **Notas de Aula – Probabilidade: teoria e exercícios**. São Paulo: Programa de Pós-Graduação em Estatística do Departamento de Estatística da Universidade de São Paulo, [s.d.] Disponível em:

<http://hostel.ufabc.edu.br/~cristian.coletti/arquivos/Livro.pdf>. Acesso em: 07 dez. 2020.

Referências

ANDERSON, David. R.; SWEENEY, Dennis J.; WILLIAMS, Thomas A. **Estatística aplicada a administração e economia**. 2. ed. São Paulo: Cengage Learning, 2011.

BUSSAB, Wilton de O.; MORETTIN, Pedro A. **Estatística básica**. 6. ed. São Paulo: Saraiva, 2010.

DOANE, David P.; SEWARD, Lori E. **Estatística aplicada à administração e economia**. 4. ed. Porto Alegre: Mc Graw Hill, 2014.

LARSON, Ron; FARBER, Betsy. **Estatística aplicada**. 4. ed. São Paulo: Pearson, 2010.

5

Modelos de distribuição de probabilidade e inferência estatística

Métodos Quantitativos Aplicados
às Ciências Contábeis 2

Lidia Maria Alves Rodella Amorim

Objetivos da aprendizagem

- 1 Determinar as probabilidades de Binomial, Hipergeométrica e de Poisson;
- 2 Encontrar probabilidades com as distribuições normal e t de Student;
- 3 Realizar estimação por intervalo e testes de hipóteses.

Introdução

Algumas variáveis aleatórias adaptam-se muito bem a uma série de problemas práticos. Assim, segundo Bussab e Morettin (2010), um estudo dessas variáveis é de grande valor para a elaboração de modelos probabilísticos para situações reais e, a consequente estimação de seus parâmetros. Usamos os modelos probabilísticos para descrever as características essenciais de um processo estocástico, para orientar na tomada de decisões ou fazer previsões. (DOANE; SEWARD, 2014). Nas seções 5.1 e 5.2, veremos como alguns modelos probabilísticos, para variáveis aleatórias discretas e contínuas são desenvolvidos e como são comumente utilizados.

Na seção 5.3 você aprenderá como são usados certos métodos estatísticos, como a média amostral e a proporção amostral, para estimar a média e a proporção da população, bem como construir intervalos de confiança e realizar testes de hipóteses, compreendendo assim a inferência estatística, muito importante para o processo de tomada de decisões e testes de suposições e teorias.

5.1 Modelos de distribuição de probabilidade para variáveis aleatórias discretas

Iniciaremos pelos modelos de distribuição de probabilidade para variáveis aleatórias discretas, salientando que uma distribuição de probabilidade discreta lista cada valor possível que a variável aleatória pode assumir, junto com sua probabilidade.

5.1.1 Distribuição de Bernoulli

Os experimentos de Bernoulli apresentam apenas dois possíveis resultados: sucesso (se acontecer o evento de interesse) ou fracasso (o evento não se realiza). A probabilidade de sucesso é p e a probabilidade de fracasso é $(1 - p)$.

Atenção!

Em todos os modelos de distribuição de probabilidade, as condições de que a probabilidade é um valor entre 0 e 1 e a soma de todas as probabilidades é igual a 1 devem ser atendidas.

Alguns exemplos de experimentos de Bernoulli são apresentados por Bussab e Morettin (2010):

1. Uma moeda é lançada: ou o resultado é cara ou não (coroa);
2. Um dado é jogado: ou cai face 5 ou não (caindo 1, 2, 3, 4 ou 6);
3. Uma peça é selecionada ao acaso de uma caixa com 500 peças: essa peça tem defeito ou não;
4. Uma pessoa escolhida aleatoriamente dentre 1.000 é ou não do sexo masculino.

Teremos, portanto, a seguinte variável aleatória X de Bernoulli, representada também pela notação $X \sim Ber(p)$:

$$X = \{0, \text{fracasso com } P(0) = 1 - p, 1, \text{sucesso com } P(1) = p\}$$

Considerando o que vimos no Capítulo 4 sobre variáveis aleatórias discretas, temos:

$$E(X) = 0(1 - p) + 1(p) = p$$

$$V(X) = p - p^2 = p(1 - p)$$

Exemplo 1: Considere o experimento de lançamento de um dado e sucesso se cair o lado 5. Calcule as probabilidades, o valor esperado e a variância.

Solução:

$$P(X = 0) = \frac{5}{6} \text{ e } P(X = 1) = \frac{1}{6}$$

$$E(X) = \frac{1}{6} \text{ e } V(X) = \frac{1}{6} \left(\frac{5}{6} \right) = \frac{5}{36}$$

5.1.2 Distribuição Binomial

Repetições independentes de um ensaio de Bernoulli, com a mesma probabilidade de ocorrência de “sucesso”, dão origem ao modelo de distribuição binomial.

Assim, um experimento binomial é um experimento de probabilidade que apresenta as seguintes características:

1. Ocorre quando um ensaio de Bernoulli é repetido n vezes;
2. Para cada ensaio existe dois possíveis resultados: sucesso ou fracasso;
3. A probabilidade de sucesso é p e a de fracasso é $(1 - p)$.
4. Os ensaios são independentes “isto é, o resultado de um ensaio não tem influência nenhuma no resultado de qualquer outro ensaio.” (BUSSAB e MORETTIN, 2010, p.143) e a probabilidade de sucesso p permanece constante em cada ensaio;

5. A variável aleatória binomial X é o número de sucessos num experimento binomial, ou seja, o número de “sucessos” ocorridos em n repetições independentes do experimento de Bernoulli.

A função de probabilidade de uma variável aleatória que tem distribuição binomial é a seguinte:

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

onde: k = número de sucessos; n = número de ensaios; p = probabilidade de sucesso. Além disso, a notação utilizada é $X \sim b(n, p)$.

Para encontrar o valor esperado e a variância, basta multiplicar por n as respectivas fórmulas da distribuição de Bernoulli:

$$E(X) = np$$

$$V(X) = np(1 - p)$$

Exemplo 2: (LARSON; FARBER, 2010) Uma pesquisa indica que 41% das mulheres nos Estados Unidos têm a leitura como atividade de lazer preferida. São escolhidas, aleatoriamente, quatro mulheres norte-americanas e pergunta se elas têm a leitura como atividade de lazer preferida. Encontre a probabilidade de que: (A) Exatamente duas delas respondam sim; (B) No mínimo duas delas respondam sim; (C) Menos que duas delas respondam sim.

Solução:

- a) Dado que $n = 4$, $p = 0,41$, $(1 - p) = 0,59$ e $x = 2$, a probabilidade de que exatamente duas mulheres respondam sim é:

$$P(2) = \binom{4}{2} (0,41)^2 (0,59)^{4-2} = \frac{4!}{2! (4 - 2)!} (0,41)^2 (0,59)^2 = 6 \cdot (0,41)^2 (0,59)^2 \cong 0,351$$

b) A probabilidade de que no mínimo duas delas respondam sim será a soma de $P(2)$, $P(3)$ e $P(4)$:

$$P(3) = \binom{4}{3} (0,41)^3 (0,59)^{4-3} = 4 \cdot (0,41)^3 (0,59)^1 \cong 0,1626$$

$$P(4) = \binom{4}{4} (0,41)^4 (0,59)^{4-4} = 6 \cdot (0,41)^4 (0,59)^0 \cong 0,0282$$

$$P(X \geq 2) = P(2) + P(3) + P(4) = 0,351 + 0,1626 + 0,0282 = 0,5418$$

Para calcular a probabilidade de que menos que duas mulheres respondam sim iremos somar $P(0)$ e $P(1)$:

$$P(0) = \binom{4}{0} (0,41)^0 (0,59)^{4-0} = 1 \cdot (0,41)^0 (0,59)^4 \cong 0,1211$$

$$P(1) = \binom{4}{1} (0,41)^1 (0,59)^{4-1} = 4 \cdot (0,41)^1 (0,59)^3 \cong 0,3368$$

$$P(X < 2) = P(0) + P(1) = 0,1211 + 0,3368 = 0,4579$$

5.1.3 Distribuição Hipergeométrica

Quando consideramos extrações casuais feitas **sem reposição** de uma população dividida segundo dois atributos (Bussab e Morettin, 2010) teremos a distribuição hipergeométrica. Ela é semelhante à distribuição binomial, exceto pelo fato da amostragem ser sem reposição de uma população finita com N itens.

A sua função de probabilidade é o resultado do produto e divisão de algumas combinações:

$$P(X = k) = \frac{\binom{r}{k} \binom{N-r}{n-k}}{\binom{N}{n}}$$

onde: $\max(0, n - N + r) \leq k \leq \min(r, n)$, N = número de itens na população, r = número de sucessos na população, n = tamanho da amostra e X = número de sucessos na amostra.

O valor esperado e a variância de uma variável aleatória com distribuição hipergeométrica são dados por:

$$E(X) = n \cdot \frac{r}{N}$$

$$V(X) = n \cdot \frac{r}{N} \cdot \left(1 - \frac{r}{N}\right) \cdot \frac{N - n}{N - 1}$$

Exemplo 3: (www.portalection.com.br) Suponha que 3 moedas comemorativas foram colocadas por engano em um cofrinho no qual já haviam algumas moedas comuns, o qual ficou contendo um total de 12 moedas. Suponha que, devido à dificuldade de tirar as moedas do cofrinho sem quebrá-lo, vamos retirar ao acaso um total de 4 moedas, qual a probabilidade de retirarmos no mínimo 1 moeda comemorativa?

Solução: Dado que $N = 12$, $n = 4$ e $r = 3$, a probabilidade $P(X \geq 1)$ poderá ser calculada de duas formas, somando $P(1)$, $P(2)$ e $P(3)$ ou subtraindo de 1 a probabilidade de X ser menor que 1.

$$P(X \geq 1) = P(1) + P(2) + P(3) = 1 - P(X < 1) = 1 - P(0)$$

$$P(0) = \frac{\binom{3}{0} \binom{12-3}{4-0}}{\binom{12}{4}} = \frac{1 \cdot 126}{495} = \frac{126}{495}$$

Portanto, $P(X \geq 1) = 1 - \frac{126}{495} \cong 0,7454$.

5.1.4 Distribuição de Poisson

Como destacam Larson e Farber (2010), em um experimento binomial estamos interessados em encontrar a probabilidade de um número definido de sucessos em um dado número de tentativas. Na distribuição de Poisson, em vez disso, o que se busca é a probabilidade de que um número específico de ocorrência aconteça dentro de uma dada unidade de tempo ou espaço.

Para termos uma variável aleatória discreta que se comporte segundo uma distribuição de Poisson é preciso que a mesma satisfaça as seguintes condições, conforme Larson e Farber (2010):

- O experimento de Poisson consiste em calcular o número de vezes, k , que um evento ocorre em um dado intervalo escolhido aleatoriamente. O intervalo pode ser tempo, área ou volume.
- A probabilidade do evento acontecer permanece constante para cada intervalo.
- O número de ocorrências em um intervalo é independente do número de ocorrências em outro intervalo, isto é, os eventos são independentes.

São exemplos:

- a) Número de clientes chegando ao caixa eletrônico de um banco em dado minuto;
- b) Número de pacientes que chegam com asma em dada hora em um consultório;
- c) Número de relatórios de acidentes enviados a uma seguradora em uma semana.

A probabilidade de exatas k ocorrências em um intervalo é dada por:

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

onde λ é o número médio de ocorrências por intervalo de unidade e pode ser utilizada a notação $X \sim Pois(\lambda)$. Além disso, o valor esperado e variância de uma variável aleatória discreta com distribuição de Poisson são iguais a λ .

Exemplo 4: (DOANE; SEWARD, 2014) Segundo um estudo sobre Controle de Qualidade na empresa J.D. Associados realizado em 2006, os consumidores relataram em média 1,7 problemas por veículo com novos automóveis da VW 2006. Em um grupo de novos veículos da VW selecionados aleatoriamente, encontre a probabilidade de: (A) Nenhum problema; (B) Pelo menos um problema; (C) Mais do que três problemas.

Solução:

- a) Dado que $\lambda = 1,7$, temos:

$$P(0) = \frac{e^{-1,7} 1,7^0}{0!} = e^{-1,7} \cong 0,1826$$

- b) Considerando a quantidade quase infinita de possíveis problemas, vamos calcular a probabilidade de pelo menos um problema subtraindo de 1 a probabilidade de ter menos do que um problema:

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 1 - 0,1826 = 0,8174$$

- c) Seguindo a mesma ideia do item anterior, encontraremos a probabilidade de mais do que três problemas também subtraindo de 1 o seu evento complementar:

$$P(X > 3) = 1 - P(X \leq 3) = 1 - [P(0) + P(1) + P(2) + P(3)]$$

$$P(1) = \frac{e^{-1,7} 1,7^1}{1!} \cong 0,3105$$

$$P(2) = \frac{e^{-1,7} 1,7^2}{2!} \cong 0,264$$

$$P(3) = \frac{e^{-1,7} 1,7^3}{3!} \cong 0,15$$

$$P(X > 3) = 1 - [0,1826 + 0,3105 + 0,264 + 0,15] = 0,0929$$

5.2 Modelos de distribuição de probabilidade para variáveis aleatórias contínuas

Agora vamos ampliar nossa discussão sobre modelos de distribuição de probabilidade, incluindo modelos de variáveis aleatórias contínuas, que são aquelas que geralmente surgem a partir de uma mensuração e que podem assumir valores numéricos decimais.

5.2.1 Distribuição Normal

Como destacam Bussab e Morettin (2010), a distribuição normal é um modelo fundamental em probabilidade e inferência estatística. Distribuições normais podem ser usadas para modelar muitos grupos de mensuração de dados na natureza, indústria e negócios. (LARSON; FARBER, 2010).

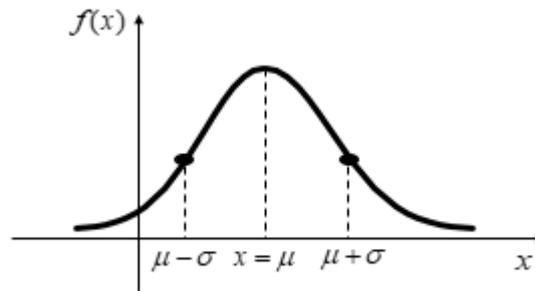
A variável aleatória X tem distribuição normal se sua função de densidade de probabilidade (f.d.p) for da forma:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2}, \quad -\infty < x < \infty$$

com parâmetros μ e σ^2 , tal que $-\infty < \mu < +\infty$ e $0 < \sigma^2 < \infty$.

A Figura 1 ilustra uma curva normal, para valores particulares de μ e σ^2 .

Figura 1 – f.d.p. de uma va.a normal com média μ e desvio padrão σ .



Fonte: Elaborado pela autora.

Segundo Bussab e Morettin (2010), podemos destacar as seguintes propriedades da distribuição normal, também possíveis de serem visualizadas graficamente (Figura 1):

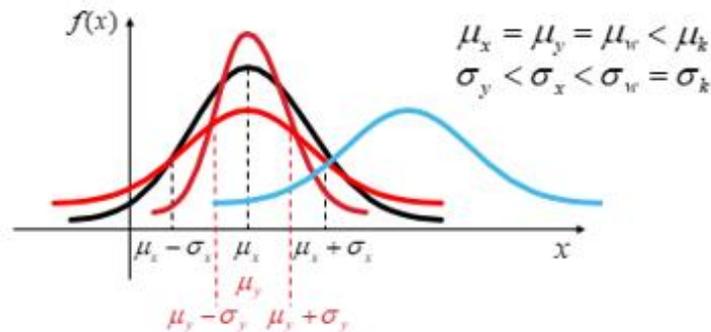
- A função $f(x)$ tende a zero quando x tende a mais ou menos infinito (∞).
- Em $x = \mu$ temos o ponto máximo da função.
- A função $f(x)$ é simétrica em relação à $x = \mu$.
- Temos que $x = \mu \pm \sigma$ são pontos de inflexão.

Os dois parâmetros μ e σ^2 , que caracterizam a distribuição normal, representam o valor médio (ou valor esperado) e a variância da variável aleatória X , respectivamente. Além disso, ela é frequentemente denotada por $N(\mu; \sigma)$, em que σ é o desvio padrão (ou seja, a raiz da variância de X).

O que acontece com a curva da normal quando alteramos os valores dos parâmetros μ e σ ? Observe na Figura 2 que alterações na média (μ) deslocam a curva para a direita ou para a esquerda, já modificações no desvio padrão (σ) resultam em curvas mais alongadas ou mais achatadas. Ou seja, quanto menor o desvio padrão, mais próximos os valores estarão do valor médio e quanto maior o desvio padrão, mais dispersos estarão os valores (mais achatada a curva).

Figura 2 – Curvas da normal com alterações em μ e σ .

$$X \sim N(\mu_x, \sigma_x^2) \quad Y \sim N(\mu_y, \sigma_y^2) \quad W \sim N(\mu_w, \sigma_w^2) \quad K \sim N(\mu_k, \sigma_k^2)$$



Fonte: Elaborado pela autora.

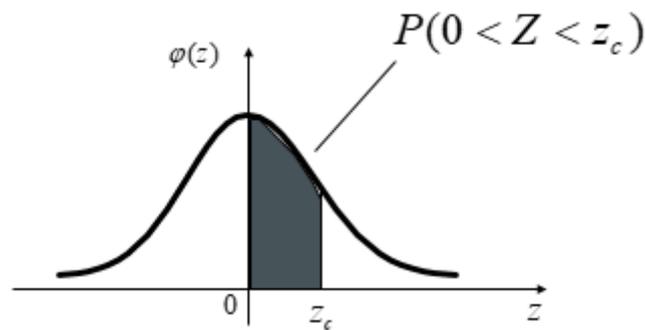
Considerando que existe uma distribuição normal diferente para cada valor de μ e σ , Doane e Seward (2014) destacam que frequentemente transformamos a variável em estudo subtraindo a média e dividindo pelo desvio padrão para produzir uma variável padronizada Z , conforme a seguir, que tem uma distribuição normal padrão com média 0 e desvio padrão 1, denotada por $N(0, 1)$.

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

Assim é possível calcular a probabilidade de uma variável aleatória X com distribuição normal, pertencer a um intervalo de valores através do uso da Tabela de Distribuição Normal Padrão (Anexo I), isto é, sem precisar utilizar a sua f.d.p. nada simples apresentada no início desta seção. Essa tabela fornece as probabilidades sob uma curva normal padrão, que nada mais são do que as correspondentes áreas sob a curva (BUSSAB; MORETTIN, 2010).

A tabela que vamos utilizar (Anexo I) fornece a área sob a curva normal padrão entre $Z = 0$ e qualquer valor positivo, conforme Figura 3.

Figura 3 – Probabilidade fornecida pela tabela da distribuição normal padrão (Anexo I).



Fonte: Elaborada pela autora.

Tendo em vista a simetria em torno de $Z = 0$, é possível obter a área (probabilidade) entre quaisquer valores de Z (positivos ou negativos). Vejamos, então, como obter probabilidades a partir da tabela apresentada no Anexo I.

Atenção!

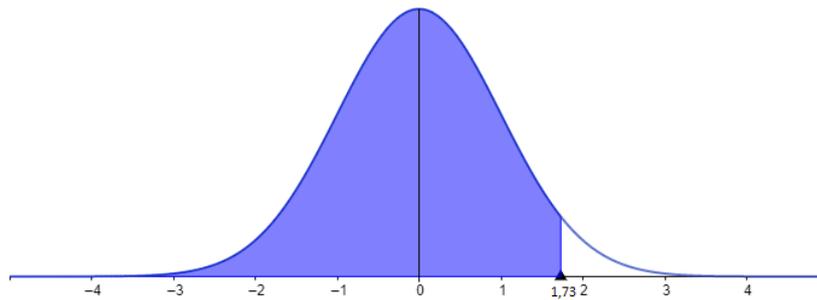
Sempre que for resolver exercícios que precise utilizar tabelas de distribuição de probabilidade, sugiro fazer o gráfico da área da questão, assim você conseguirá visualizar melhor como utilizar a tabela e fazer adaptações (quando necessário), para o seu uso adequado (exemplos 6 a 12).

Exemplo 5/Solução: $P(0 \leq Z \leq 1) = 0,34134$, conforme linha 1,0 e coluna 1 da Tabela da Distribuição Normal Padrão (Anexo I).

Exemplo 6/Solução: $P(Z \leq 1,73) = P(Z \leq 0) + P(0 \leq Z \leq 1,73) = 0,5 + 0,4582 = 0,9582$.

Graficamente, a área calculada é a seguinte (Figura 4):

Figura 4 – Ilustração do cálculo de $P(Z \leq 1,73)$.



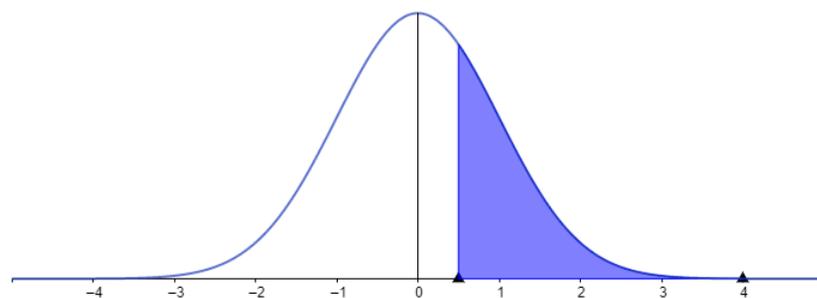
Fonte: Elaborada pela autora.

Exemplo 7/Solução: $P(-1 \leq Z \leq 0) = P(0 \leq Z \leq 1) = 0,34134$

Exemplo 8/Solução: $P(Z > -0,76) = P(-0,76 < Z < 0) + P(Z > 0) =$
 $= P(0 < Z < 0,76) + 0,5 = 0,27637 + 0,5 = 0,77637$

Exemplo 9/Solução: $P(Z > 0,5) = P(Z > 0) - P(0 < Z < 0,5) = 0,5 - 0,19146 =$
 $0,30854$. Observe a área calculada na Figura 5.

Figura 5 - Ilustração do cálculo de $P(Z > 0,5)$.

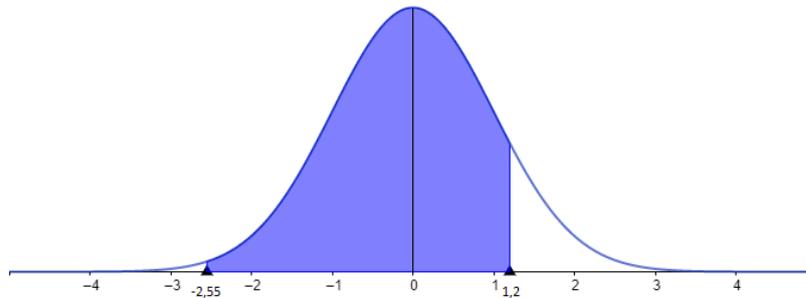


Fonte: Elaborada pela autora.

Exemplo 10/Solução: $P(Z \leq -1,53) = P(Z \geq 1,53) = 0,5 - P(0 \leq Z \leq 1,53) =$
 $= 0,5 - 0,43699 = 0,06301$

Exemplo 11/Solução: $P(-2,55 \leq Z \leq 1,2) = P(-2,55 \leq Z \leq 0) + P(0 < Z \leq 1,2) = P(0 \leq Z \leq 2,55) + 0,38493 = 0,49461 + 0,38493 = 0,87954$. Veja a área na Figura 6.

Figura 6 - Ilustração do cálculo de $P(-2,55 \leq Z \leq 1,2)$.



Fonte: Elaborada pela autora.

Exemplo 12/Solução: $P(0,48 < Z < 2,12) = P(0 < Z < 2,12) - P(0 < Z < 0,48) = 0,48300 - 0,18439 = 0,29861$

Exemplo 13: As alturas dos alunos de determinada escola são normalmente distribuídas com média 1,60m e desvio padrão 0,30m. Encontre a probabilidade de um aluno medir: (A) entre 1,50m e 1,80m; (B) Mais de 1,75m; (C) Menos de 1,48m.

Solução:

$$\begin{aligned} \text{a) } P(1,50 < X < 1,80) &= P\left(\frac{1,50-1,60}{0,3} < Z < \frac{1,80-1,60}{0,3}\right) = P(-0,33 < Z < 0,67) = \\ &= P(0 < Z < 0,33) + P(0 < Z < 0,67) = 0,12930 + 0,24857 = 0,37787 \end{aligned}$$

$$\begin{aligned} \text{b) } P(X > 1,75) &= P\left(Z > \frac{1,75-1,60}{0,3}\right) = P(Z > 0,5) = 0,5 - P(0 < Z < 0,5) = 0,5 - \\ &0,19146 = 0,30854 \end{aligned}$$

$$\begin{aligned} \text{c) } P(X < 1,48) &= P\left(Z < \frac{1,48-1,60}{0,3}\right) = P(Z < -0,4) = P(Z > 0,4) = 0,5 - P(0 < Z < 0,4) = \\ &= 0,5 - 0,15542 = 0,34458 \end{aligned}$$

5.2.2 Distribuição t de Student

Como destacam Bussab e Morettin (2010), a distribuição t de Student é importante no estudo da inferência sobre médias populacionais, assunto que veremos na seção 5.3.

Dizemos que a variável aleatória t tem distribuição t de Student se sua f.d.p. for do tipo:

$$f(t) = \frac{\Gamma[(v+1)/2]}{\Gamma(v/2)\sqrt{\pi v}} [1 + t^2/v]^{-(v+1)/2}, \quad -\infty < t < \infty$$

onde v são os graus de liberdade e a indicaremos por $t(v)$. Pode-se provar que (Bussab e Morettin, 2010):

$$E(t) = 0 \text{ e } V(t) = \frac{v}{v-2}, \text{ para } v > 2$$

Os autores Larson e Farber (2010) apresentam as propriedades da distribuição t de Student:

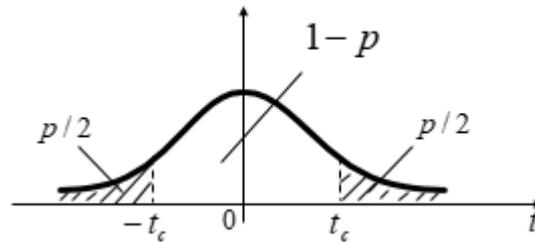
- Tem forma de sino e é simétrica sobre a média.
- É uma família de curvas, cada uma determinada por um parâmetro chamado de grau de liberdade (v). Os graus de liberdade são o número de escolhas livres deixadas depois que uma amostra estatística tal como \bar{x} é calculada, ou seja, o tamanho da amostra menos 1.
- A área total sob a curva t é 1 ou 100%.
- A média, a mediana e a moda da distribuição t são iguais a zero.
- Conforme os graus de liberdade aumentam, a distribuição t se aproxima da distribuição normal.

Como essa distribuição é muito utilizada na prática, existem tabelas fornecendo probabilidades relacionadas a ela. Como destacam Bussab e Morettin (2010), a Tabela apresentada no Anexo II fornece os valores de t_c tais que

$$P(-t_c < t(v) < t_c) = 1 - p$$

para alguns valores de p e de v , conforme Figura 7:

Figura 7 – A distribuição t de Student.



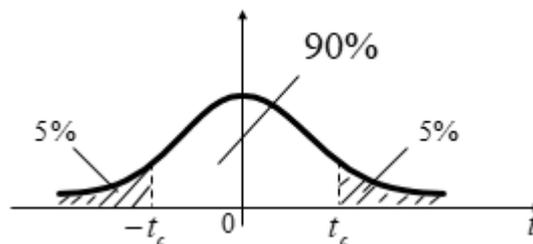
Fonte: Elaborada pela autora.

Observe que na Figura 7 temos um gráfico bicaudal e, portanto, a probabilidade p que será utilizada na Tabela do Anexo II é a soma das áreas das duas caudas.

Exemplo 14: Seja t uma variável aleatória com distribuição t de Student e 25 graus de liberdade. Calcule os valores de t_c tais que $P(-t_c < t < t_c) = 90\%$.

Solução: Vejamos, primeiramente, graficamente na Figura 8 o que a questão pede:

Figura 8 – Ilustração de $P(-t_c < t < t_c) = 90\%$.



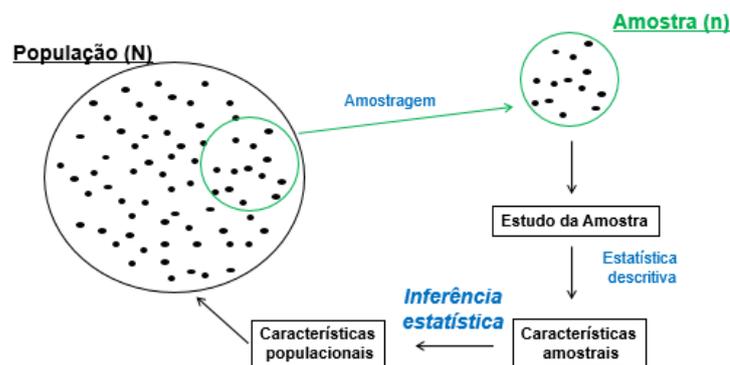
Fonte: Elaborada pela autora.

Temos então que $p = 10\%$ e $v = 25$, e consultando esses valores na tabela da distribuição t de Student (Anexo II), coluna e linha, respectivamente, encontramos $t_c = 1,7081$.

5.3 Inferência Estatística

A inferência estatística se refere ao processo de obter informações sobre dada característica de uma população a partir de resultados observados na amostra, conforme representado na Figura 9.

Figura 9 – Processo de inferência estatística.



Fonte: Elaborada pela autora.

Como destacam Anderson, Sweeney e Williams (2011), um dos propósitos fundamentais da inferência estatística é desenvolver estimativas e testar hipóteses a respeito dos parâmetros populacionais usando a informação contida em uma amostra.

Para uma melhor compreensão do que vamos estudar nesta seção, em que trataremos de distribuições amostrais, estimação e testes de hipóteses, é importante que você tenha em mente alguns conceitos básicos:

- **População** é o conjunto de todos os elementos de interesse de um dado estudo (BUSSAB; MORETTIN, 2010).
- **Amostra** é qualquer subconjunto da população (BUSSAB; MORETTIN, 2010).
- **Estatística** é uma característica da amostra e **parâmetro** é uma medida usada para descrever uma característica da população, como apresentado por Bussab e Morettin (2010).

Além disso, vamos considerar o uso do método de amostragem aleatória simples, um dos mais comuns para selecionar uma amostra de uma população, destacando que sua definição e processo de seleção dependem se a população é finita ou infinita, conforme Anderson, Sweeney e Williams (2011):

“ Uma amostra aleatória simples de tamanho n de uma população finita de tamanho N é uma amostra selecionada de tal maneira que cada amostra possível de tamanho n tenha a mesma probabilidade de ser escolhida. ” (ANDERSON; SWEENEY; WILLIAMS, 2011, p. 240).

Enquanto isso, no caso de uma população infinita, uma amostra aleatória simples é selecionada de forma que sejam atendidas as seguintes condições: “1. Cada elemento selecionado vem dessa população. 2. Cada elemento é selecionado de maneira independente.” (ANDERSON; SWEENEY; WILLIAMS, 2011, p. 241).

5.3.1 Estimação por ponto

A estimação é o processo que consiste no uso de dados da amostra (dados amostrais) para estimar valores de parâmetros populacionais desconhecidos, tais como média, desvio padrão, proporções etc. No processo de estimação temos o **estimador**, que é a estatística utilizada com o objetivo de estimar um parâmetro, e a **estimativa**, que é o valor numérico de um estimador, como definem Doane e Seward (2014).

A estimação pode ser feita de duas formas, definidas por Anderson, Sweeney e Williams (2011) como:

- Usando um **estimador pontual**, que estima o valor de um parâmetro não conhecido usando um simples valor.
- Usando um **estimador intervalar** (*por intervalo*), que estima o valor de um parâmetro não conhecido usando um intervalo. Esse estimador é o mais utilizado, principalmente porque o procedimento de estimação pontual não permite julgar qual a magnitude do erro que estamos cometendo.

Na estimação pontual, para estimar o valor do parâmetro de uma população calculamos uma característica correspondente da amostra, denominada **estatística amostral**. As principais estão na Tabela 1.

Tabela 1 – Parâmetros e estimadores

Denominação	População/Parâmetros	Estimativa/Estatística amostral
Média	$\mu = \frac{\sum_{i=1}^N x_i}{N}$	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$
Desvio padrão	$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$	$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$
Proporção	$p = \frac{X}{N}$ onde X é o número de elementos da população com a característica de interesse.	$\bar{p} \text{ ou } \hat{p} = \frac{x}{n}$ onde x é o número de elementos da amostra com a característica de interesse.

Fonte: Elaborada pela autora.

Exemplo 15: (ANDERSON; SWEENEY; WILLIAMS, 2011, p.245) Uma amostra aleatória simples dos dados de cinco meses de vendas forneceu a seguinte informação:

Mês	1	2	3	4	5
Unidades vendidas	94	100	85	94	92

- Elabore uma estimação por ponto do número médio de unidades da população vendidas por mês.
- Elabore a estimação por ponto do desvio padrão da população.

Solução:

$$\text{a) } \bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{94+100+85+94+92}{5} = \frac{465}{5} = 93$$

$$\text{b) } s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(94-93)^2 + (100-93)^2 + (85-93)^2 + (94-93)^2 + (92-93)^2}{4}} = \sqrt{29} \cong 5,38$$

5.3.2 Amostras e distribuições amostrais

Se considerarmos que o processo de escolher uma amostra aleatória simples é um experimento, a média amostral \bar{x} (uma estatística amostral) é uma descrição numérica do resultado do experimento e, portanto, é uma variável aleatória, explicam Anderson, Sweeney e Williams (2011).

Assim, como variável aleatória, a média amostral \bar{x} (ou qualquer outra estatística amostral) tem um valor esperado (ou valor médio), um desvio padrão (ou variância) e uma distribuição de probabilidade. Tendo em vista que os diversos valores possíveis de \bar{x} resultam de diferentes amostras aleatórias simples, a distribuição de probabilidade de \bar{x} é chamada de **distribuição amostral** de \bar{x} , afirmam os autores citados acima.

5.3.2.1 Distribuição amostral de \bar{x}

A distribuição amostral de \bar{x} é a distribuição de probabilidade de todos os valores possíveis da média amostral \bar{x} . Conhecer essa distribuição amostral e suas propriedades nos possibilitará fazer afirmações a respeito de quão próxima a média da amostra (\bar{x}) está da média da população (μ).

Vejamos a distribuição amostral de \bar{x} , de acordo com Anderson, Sweeney e Williams (2011):

- Valor esperado de \bar{x} : Quando se trata de uma amostragem aleatória simples, temos que $E(\bar{x}) = \mu$, ou seja, o valor esperado de \bar{x} é igual à média da população da qual a amostra é retirada.

- Desvio padrão de \bar{x} : Quando se trata de uma amostragem aleatória simples, o desvio padrão de \bar{x} depende de a população ser finita ou infinita:

$$s_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \rightarrow \text{Sempre que a população for infinita ou finita e } \frac{n}{N} \leq 0,05.$$

$$s_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \cdot \frac{\sigma}{\sqrt{n}} \rightarrow \text{Sempre que a população for finita e } \frac{n}{N} > 0,05.$$

- Forma da distribuição amostral de \bar{x} : consideramos duas situações para defini-la:
 - a) A população tem uma distribuição normal: nesse caso, a distribuição amostral de \bar{x} está normalmente distribuída para qualquer tamanho de amostra.
 - b) A população não tem uma distribuição normal: para $n > 30$ deve-se aplicar o Teorema do Limite Central, que diz:

“Ao selecionar amostras aleatórias simples de tamanho n de uma população, podemos aproximar a distribuição amostral da média da amostra \bar{x} por meio de uma distribuição normal à medida que o tamanho da amostra se torna maior.”

(ANDERSON; SWEENEY; WILLIAMS, 2011, p. 251).

Você sabia?

Quando o valor esperado de um estimador por ponto for igual ao parâmetro populacional, dizemos que o estimador por ponto é sem viés (ou não-viesado). Portanto, \bar{x} é um estimador sem viés de μ .

Exemplo 16: (ANDERSON, SWEENEY; WILLIAMS, 2011, p.257) O custo médio anual dos seguros de automóvel é US\$ 687. Use esse valor como média populacional e suponha que o desvio padrão da população seja $\sigma = \text{US\$ } 230$. Considere uma amostra de 45 apólices de seguro de automóveis. (A) Apresente a distribuição amostral de \bar{x} , em que \bar{x} é a média amostral do custo anual dos seguros de automóveis. (B) Calcule a probabilidade de a média amostral estar dentro dos US\$ 100 da média populacional.

Solução:

a) $E(\bar{x}) = \mu = 687$

$$s_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{230}{\sqrt{45}} = 34,28$$

\bar{x} tem distribuição normal, considerando que no enunciado não fala sobre a distribuição da população, mas justifica-se tal distribuição pelo Teorema do Limite Central e que $n > 30$.

b)

$$\begin{aligned} P(687 - 100 \leq \bar{x} \leq 687 + 100) &= P(587 \leq \bar{x} \leq 787) = P\left(\frac{587 - 687}{34,28} \leq \bar{x} \leq \frac{787 - 687}{34,28}\right) = \\ &= P(-2,91 \leq Z \leq 2,91) = 2 \cdot P(0 \leq Z \leq 2,91) = 2 \cdot 0,49819 = 0,99638 \end{aligned}$$

5.3.2.2 Distribuição amostral de \bar{p}

A distribuição amostral de \bar{p} é a distribuição de probabilidade de todos os valores possíveis da proporção amostral \bar{p} .

A proporção amostral \bar{p} é o estimador por ponto da proporção p da população:

$$\bar{p} = \frac{x}{n}$$

onde x é o número de elementos contidos na amostra que possuem a característica de interesse.

A distribuição amostral de \bar{p} , conforme Anderson, Sweeney e Williams (2011), é dada por:

- Valor esperado de \bar{p} : É a média de todos os valores possíveis de \bar{p} e é igual à proporção populacional de p , ou seja, $E(\bar{p}) = p$ e \bar{p} é um estimador não viesado de p .
- Desvio padrão de \bar{p} : Semelhante ao que ocorre em \bar{x} , quando se trata de uma amostragem aleatória simples, o desvio padrão de \bar{p} também depende de a população ser finita ou infinita:

$$s_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}} \rightarrow \text{Sempre que a população for infinita ou finita e } \frac{n}{N} \leq 0,05.$$

$$s_{\bar{p}} = \sqrt{\frac{N-n}{N-1}} \cdot \sqrt{\frac{p(1-p)}{n}} \rightarrow \text{Sempre que a população for finita e } \frac{n}{N} > 0,05.$$

- Forma da distribuição amostral de \bar{p} : Para uma amostra aleatória simples de uma população grande, o valor de \bar{p} é uma variável aleatória binomial que indica o número de elementos contidos na amostra que possuem a característica de interesse. No entanto, a distribuição amostral de \bar{p} pode ser aproximada por meio de uma distribuição normal sempre que $np \geq 5$ e $n(1-p) \geq 5$.

Exemplo 17: (ANDERSON; SWEENEY; WILLIAMS, 2011, p. 261) O presidente de uma empresa acredita que 30% das encomendas feitas à firma são provenientes de clientes que compram pela primeira vez. Uma amostra aleatória simples de 100 pedidos será usada para estimar a proporção de clientes que compram pela primeira vez. (A) Suponha que o presidente esteja correto e $p = 30$. Apresente a distribuição amostral de \bar{p} neste estudo. (B) Qual é a probabilidade de a proporção amostral \bar{p} estar entre 0,20 e 0,40?

Solução:

a) $E(\bar{p}) = p = 0,30$

$$s_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0,30(1-0,30)}{100}} = \sqrt{0,0021} \cong 0,0458$$

$n \cdot p = 100 \cdot 0,30 = 30 > 5$ e $n(1-p) = 100(1-0,30) = 70 > 5$, assim, temos que a distribuição amostra de \bar{p} pode ser aproximada por meio de uma distribuição normal.

b) $P(0,2 \leq \bar{p} \leq 0,4) = P\left(\frac{0,2-0,3}{0,0458} \leq \bar{p} \leq \frac{0,4-0,3}{0,0458}\right) = P(-2,18 \leq Z \leq 2,18) =$
 $= 2 \cdot P(0 \leq Z \leq 2,18) = 2 \cdot 0,48537 = 0,97074$

5.3.3 Estimação por intervalo

Uma vez que não se pode esperar que um estimador por ponto produza o valor exato do parâmetro populacional, uma estimação por intervalo frequentemente é calculada, com a finalidade de fornecer informações sobre quão próximo o estimador por ponto, produzido pela amostra, está do valor do parâmetro populacional, como afirmam Anderson, Sweeney e Williams (2011).

As distribuições amostrais de \bar{x} e \bar{p} são fundamentais no cálculo das estimações por intervalo, uma vez que para a construção de um intervalo de confiança para a média desconhecida μ ou para a proporção desconhecida p precisamos da forma da distribuição para calcular a margem de erro que compõe o intervalo:

intervalo de confiança=estimação por ponto+margem de erro

O coeficiente de confiança $(1 - \alpha)$ para esse intervalo é expresso por um percentual. Nas ciências sociais, os graus de confiança mais utilizados são 90%, 95% ou 99%, de acordo com aqueles autores. Lembre-se que quanto maior o coeficiente (ou grau) de confiança, maior é o intervalo de confiança, como ficará evidenciado nos exemplos adiante.

5.3.3.1 Média da população: σ conhecido

Para a estimação por intervalo da média de uma população, o desvio padrão σ da população ou o desvio padrão s da amostra são usados para o cálculo da margem de erro. Anderson, Sweeney e Williams (2011) nos lembram que a partir de grandes quantidades de dados históricos relevantes disponíveis é possível calcular o desvio padrão da população antes de se fazer a amostragem.

Considerando, então, o σ conhecido e a distribuição amostral de \bar{x} (seção 5.3.2.1), que nos fornece informações sobre as possíveis diferenças entre \bar{x} e μ , poderemos construir o intervalo de confiança para a média da população.

Se as amostras são retiradas de uma população normal (ou a amostra é suficientemente grande para que \bar{x} seja aproximadamente normal pelo TLC) e σ é conhecido, então a margem de erro é calculada utilizando a distribuição normal. Teremos então o seguinte intervalo de confiança para a média μ com o σ conhecido:

$$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

em que $(1 - \alpha)$ é o coeficiente de confiança, α é o nível de significância e $z_{\alpha/2}$ é o valor de z que produz uma área de $\alpha/2$ na cauda superior da distribuição normal padrão de probabilidade.

Portanto, o valor de $z_{\alpha/2}$ dependerá do coeficiente de confiança desejado, conforme tabela 2 a seguir:

Tabela 2 – Níveis de confiança comuns e valores z

Coeficiente de confiança	$1 - \alpha$	α	$\alpha/2$	$z_{\alpha/2}$
90%	0,90	0,10	0,05	$z_{0,05} = 1,645$
95%	0,95	0,05	0,025	$z_{0,025} = 1,96$
99%	0,99	0,01	0,005	$z_{0,005} = 2,576$

Fonte: Doane e Seward, 2014, p. 304.

Afinal, qual a interpretação para um coeficiente de confiança de 95%, por exemplo? Doane e Seward (2014) explicam da seguinte forma: se seleccionássemos 100 amostras aleatórias da mesma população e utilizássemos exatamente esse procedimento para construir 100 intervalos de confiança com coeficiente de confiança de 95%, aproximadamente 95 dos intervalos conteriam a verdadeira média μ , e cerca de 5 intervalos não.

Exemplo 18: (ANDERSON; SWEENEY; WILLIAMS, 2011, p. 277) Em um esforço para estimar a quantia média que cada cliente gasta por jantar em um grande restaurante de São Francisco, foram coletados dados de uma amostra de 49 clientes. Considere um desvio padrão de US\$5,00 para a população. (A) Qual é a margem de erro para um grau de confiança de 95%? (B) Dado que a média amostral é US\$24,80, construa o intervalo de confiança de 95% para a média populacional.

Solução:

$$\text{a) } E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = z_{0,05/2} \cdot \frac{5}{\sqrt{49}} = z_{0,025} \cdot \frac{5}{\sqrt{49}} = 1,96 \cdot \frac{5}{\sqrt{49}} = 1,4$$

$$\text{b) } IC = \bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} = \bar{x} \pm E = 24,8 \pm 1,4 = 23,4 \text{ a } 26,2$$

5.3.3.2 Média da população: σ desconhecido

Nas situações em que o desvio padrão da população é desconhecido, usamos a mesma amostra para estimar μ e σ . Segundo Anderson, Sweeney e Williams (2011), temos, então, que s , é usado para estimar σ e a estimação por intervalo da média populacional se baseia em uma distribuição de probabilidade t de Student com $n - 1$ graus de liberdade.

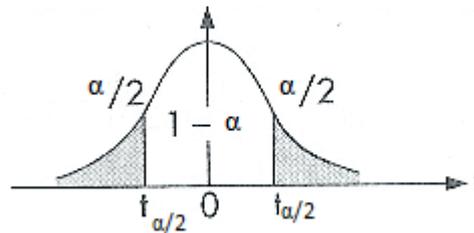
Além disso, tem-se que à medida que o n° de graus de liberdade aumenta, a diferença entre a distribuição t e a distribuição normal padrão torna-se cada vez menor.

O intervalo de confiança para a média μ com o desvio padrão populacional σ desconhecido é:

$$\bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

onde $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$ é o desvio padrão da amostra, $(1 - \alpha)$ é o coeficiente de confiança e α é o nível de significância, $t_{\alpha/2}$ é o valor de t que produz uma área de $\alpha/2$ na cauda superior da distribuição t , com $n - 1$ graus de liberdade, conforme Figura 10.

Figura 10 – Distribuição t de Student



Fonte: Adaptado de Bussab e Morettin, 2010, p. 513.

Exemplo 19: (ANDERSON; SWEENEY; WILLIAMS, 2011, p.285) A equipe de vendas de uma distribuidora apresenta semanalmente relatórios que relacionam os contatos feitos com clientes durante a semana. Uma amostra de 65 relatórios semanais exibiu uma média amostral de 19,5 contatos com clientes por semana. O desvio padrão da amostra foi 5,2. Construa o intervalo de confiança de 95% correspondente ao número médio da população de contatos semanais com clientes feitos pela equipe de vendas.

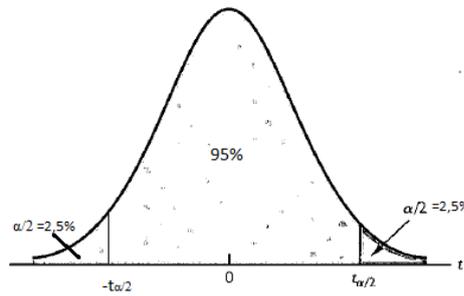
Solução:

$$1 - \alpha = 0,95 \rightarrow \frac{\alpha}{2} = \frac{0,05}{2} = 0,025$$

Observe na Figura 11 que para encontrarmos o $t_{\alpha/2}$ que precisamos para o cálculo do intervalo de confiança dessa questão, devemos entrar na tabela (Anexo II) com a soma das áreas das caudas, ou seja, 5% e $\nu = 60$ graus de liberdade (o valor de ν constante na tabela mais próximo de $n - 1 = 64$). Assim, teremos $t_{\alpha/2} = 2$ e, portanto:

$$IC = \bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} = 19,5 \pm 2 \cdot \frac{5,2}{\sqrt{65}} = 19,5 \pm 1,29 = 18,21 \text{ a } 20,79$$

Figura 11 – Distribuição t de Student com coeficiente de confiança de 95%.



Fonte: Adaptado de Anderson, Sweeney e Williams, 2011, p. 279.

5.3.3.3 Tamanho da amostra para intervalo de confiança de uma média populacional

Anderson, Sweeney e Williams (2011) nos mostram que é possível definir o tamanho da amostra de forma que ela seja grande o suficiente para produzir uma margem de erro desejada (pré-determinada), ao nível de confiança escolhido em decorrência da fórmula da margem de erro. Assim, vejamos como definir n no caso em que temos σ conhecido:

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \rightarrow n = \frac{(z_{\alpha/2})^2 \cdot \sigma^2}{E^2}$$

5.3.3.4 Proporção da população

O intervalo de confiança da proporção da população segue o mesmo princípio dos vistos anteriormente, em que resulta da soma e subtração da margem de erro à proporção amostral \bar{p} e depende da sua distribuição amostral (seção 5.3.2.2). Assim, teremos:

$$\bar{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

onde $(1 - \alpha)$ é o coeficiente de confiança, e $z_{\alpha/2}$ é o valor de z que produz uma área de $\alpha/2$ na cauda superior da distribuição normal padrão.

Exemplo 20: (ANDERSON; SWEENEY; WILLIAMS, 2011, p. 294) Dados sobre o perfil do público coletado no site da TVN, mostraram que 26% dos usuários eram mulheres. Considere que essa porcentagem tenha se baseado em uma amostra de 400 usuários. Qual é o intervalo de 95% de confiança relativo à proporção populacional de usuários do site da TVN que são mulheres?

Solução:

$$\begin{aligned} IC &= \bar{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\bar{p}(1 - \bar{p})}{n}} = 0,26 \pm 1,96 \sqrt{\frac{0,26(1 - 0,26)}{400}} = 0,26 \pm 0,043 \\ &= 0,217 \text{ a } 0,303 \end{aligned}$$

5.3.4 Testes de hipóteses

O teste de hipóteses usa dados de uma amostra para testar duas afirmações antagônicas a respeito de um parâmetro populacional, definidas por Anderson, Sweeney e Williams (2011) como:

- H_0 : hipótese nula (hipótese experimental a respeito de um parâmetro da população).
- H_a ou H_1 : hipótese alternativa (oposto de H_0).

Existem três formas possíveis para um teste de hipóteses, apresentadas abaixo.

Unicaudal inferior

$$H_0: \mu \geq \mu_0$$

$$H_a: \mu < \mu_0$$

Unicaudal superior

$$H_0: \mu \leq \mu_0$$

$$H_a: \mu > \mu_0$$

Bicaudal

$$H_0: \mu = \mu_0$$

$$H_a: \mu \neq \mu_0$$

Para auxiliar na formulação das hipóteses, os autores supracitados destacam que o termo de igualdade sempre aparece na hipótese nula e a hipótese alternativa frequentemente é aquilo que o teste tenta estabelecer.

As hipóteses nula e alternativa são afirmações excludentes a respeito da população e, afirmam Anderson, Sweeney e Williams (2011), uma vez que os testes de hipóteses baseiam-se em informações de amostras, temos que admitir a possibilidade de erros. A Tabela 3 apresenta os dois tipos de erro que podem ser incorridos no teste de hipóteses.

Tabela 3 – Erros e conclusões corretas no teste de hipóteses.

		Situação da população	
		H_0 verdadeira	H_a verdadeira
Conclusão	Aceitar H_0	Conclusão correta	Erro do Tipo II
	Rejeitar H_0	Erro do Tipo I	Conclusão correta

Fonte: Anderson, Sweeney e Williams, 2011, p. 313.

Portanto, erro do tipo I ocorre ao rejeitar H_0 quando H_0 é verdadeira e erro do tipo II ocorre quando aceitar H_0 quando ela é falsa, ou seja, H_a é verdadeira.

Temos que o nível de significância α , aquele mesmo do coeficiente de confiança $(1 - \alpha)$ definido pelo pesquisador no intervalo de confiança, é a probabilidade de se cometer um erro do tipo I quando H_0 é verdadeira enquanto igualdade.

Para realizar qualquer teste de hipóteses é necessário o cumprimento das seguintes etapas, conforme Anderson, Sweeney e Williams (2011):

1. Elaborar as hipóteses nula e alternativa.
2. Definir o nível de significância.
3. Coletar os dados da amostra de tamanho n e calcular o valor da estatística de teste.

Critério do valor p

4. Usar o valor da estatística de teste para calcular o valor p .
5. Rejeitar a hipótese nula se o *valor $p \leq \alpha$* .

Critério do valor crítico

6. Usar o α (nível de significância) para estabelecer o valor crítico e o valor de rejeição.
7. Usar o valor da estatística de teste e a regra de rejeição para determinar se deve ou não rejeitar H_0 .

Vamos aprender as etapas 3, 4 e 5 e sobre os possíveis critérios nas próximas seções.

5.3.4.1 Teste de hipótese sobre a média da população com σ conhecido

Iniciaremos estudando um teste unicaudal sobre a média da população no caso em que o desvio padrão populacional é conhecido. O teste poderia ter os seguintes formatos:

Teste da cauda inferior

$$\begin{cases} H_0: \mu \geq \mu_0 \\ H_a: \mu < \mu_0 \end{cases}$$

Teste da cauda superior

$$\begin{cases} H_0: \mu \leq \mu_0 \\ H_a: \mu > \mu_0 \end{cases}$$

A primeira coisa a ser definida é a distribuição amostral de \bar{x} . Assim, devemos selecionar uma amostra de tamanho n , calcular \bar{x} , $s_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ e assumindo que \bar{x} tem uma distribuição normal (conforme seção 5.3.2.1).

A seguir é preciso calcular a estatística de teste que, nesse caso, usamos a variável aleatória z , abaixo, como estatística de teste para determinar se \bar{x} se desvia do valor hipotético μ para justificar a rejeição da hipótese nula.

$$z = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Dessa forma, podemos utilizar a tabela da distribuição normal padrão para encontrar a probabilidade da cauda inferior ou superior referente a qualquer z .

Por fim, é feita a análise das hipóteses através de dois critérios, que produzem resultados coincidentes, que segundo Anderson, Sweeney e Williams (2011):

- Critério do valor p :
 - O valor p é uma probabilidade, calculada usando a estatística teste z , que mede o apoio (ou a falta de apoio) proporcionado pela amostra à hipótese nula e é a base para determinar se a hipótese nula deve ser rejeitada, dado o nível de significância (α).
 - O cálculo do valor p depende se o teste é de cauda inferior ou de cauda superior. Se de cauda *inferior*, ele será a área sob a curva da normal padrão à *esquerda* da estatística de teste. Já no de cauda *superior*, o valor p é a área sob a curva da normal padrão à *direita* da estatística de teste.
 - A regra de rejeição quando se usa o valor p é rejeitar H_0 se o *valor p* $\leq \alpha$.
- Critério do valor crítico (z_α):
 - É o valor da estatística de teste que corresponde a uma área α (nível de significância), localizada na cauda inferior (ou superior, a depender do tipo de teste) da distribuição amostral da estatística de teste. Ou seja, é o maior (ou menor, depende de que cauda) valor da estatística de teste que resultará na rejeição da hipótese nula.
 - Como regra de rejeição teremos que: no caso da cauda inferior, rejeita-se H_0 se $z \leq -z_\alpha$; quando for de cauda superior, rejeita-se H_0 se $z \geq z_\alpha$.

Exemplo 21: (ANDERSON; SWEENEY; WILLIAMS, 2011, p. 326) Considere o seguinte teste de hipóteses, tendo uma amostra de tamanho 40 que produziu uma média amostral 26,4. O desvio padrão da população é 6. (A) Calcule o valor da estatística de teste. (B) Encontre o valor p e indique qual a sua conclusão, para um $\alpha = 0,01$. (C) Pelo critério do valor crítico, apresente a regra de rejeição.

$$H_0: \mu \leq 25$$

$$H_a: \mu > 25$$

Solução:

- a) Dado que $n = 40$, $\bar{x} = 26,4$, $\sigma = 6$, assumindo distribuição normal e que se trata de um teste unicaudal de cauda superior, teremos a seguinte estatística de teste:

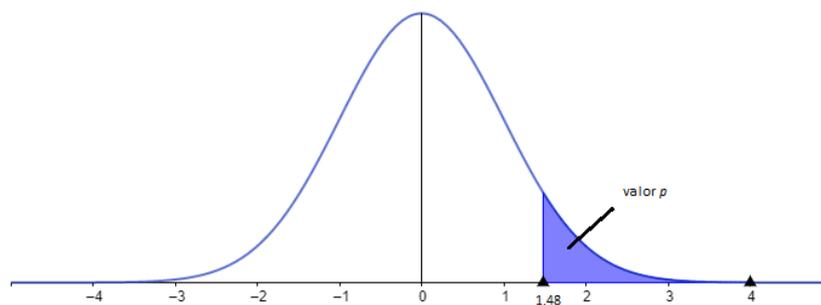
$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{26,4 - 25}{\frac{6}{\sqrt{40}}} \cong 1,48$$

- b) O valor p é a área destacada na Figura 12, portanto será igual a:

$$\text{valor } p = 0,5 - P(0 \leq Z \leq 1,48) = 0,5 - 0,43056 = 0,06944$$

$\text{valor } p (0,06944) > \alpha (0,01) \rightarrow$ Conclusão: não rejeitar H_0 .

Figura 12 – Ilustração do valor p do exemplo 21.



Fonte: Elaborada pela autora.

- c) $P(0 \leq Z \leq z_\alpha) = 0,5 - \alpha = 0,5 - 0,01 = 0,49$. Agora consultando na tabela da distribuição normal padrão (Anexo I) essa probabilidade calculada acima, teremos $z_\alpha \cong 2,32$. A regra de rejeição, portanto, será rejeitar H_0 se a estatística de teste z for maior do que 2,32.

Vejamos agora o caso de um teste **bicaudal** sobre a média da população em que o desvio padrão populacional é conhecido. Temos, portanto, as seguintes hipóteses:

$$H_0: \mu = \mu_0$$

$$H_a: \mu \neq \mu_0$$

Nessa situação, a estatística de teste e a regra de rejeição quando se usa o valor p permanecem as mesmas dos testes unicaudais acima, contudo temos uma alteração na regra de rejeição quando se usa o critério do valor crítico:

- Estatística de teste: $z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$.
- Regra de rejeição quando usa o critério do valor p : rejeita-se H_0 quando *valor* $p \leq \alpha$.
- Regra de rejeição quando usa o critério do valor crítico: rejeita-se H_0 se $z \leq -z_{\alpha/2}$ ou $z \geq z_{\alpha/2}$.

Para o cálculo do valor p no teste bicaudal, Anderson, Sweeney e Williams (2011) orientam:

1. Calcular o valor da estatística de teste.
2. Se z estiver na cauda superior ($z > 0$), encontre a área sob a curva normal padrão à direita de z . Se z estiver na cauda inferior ($z < 0$), encontre a área da curva normal padrão à esquerda de z .
3. Duplique a área da cauda (probabilidade) encontrada na etapa anterior para obter o valor p .

Exemplo 22: (ANDERSON; SWEENEY; WILLIAMS, 2011, p. 327) Considere o seguinte teste de hipóteses:

$$H_0: \mu = 22$$

$$H_a: \mu \neq 22$$

Uma amostra de tamanho 75 é usada e o desvio padrão da população é 10. Calcule o valor p e apresente sua conclusão para $\bar{x} = 25,1$. Use $\alpha = 0,01$.

Solução:

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{25,1 - 22}{\frac{10}{\sqrt{75}}} \cong 2,68$$

$$\text{valor } p = 2.P(Z \geq 2,68) = 2.[0,5 - P(0 \leq Z \leq 2,68)] = 2.(0,5 - 0,49632) = 0,00736$$

Assim, temos um o *valor* $p < \alpha$ e a conclusão é rejeitar H_0 .

5.3.4.2 Teste de hipótese sobre a média da população com σ desconhecido

E como realizar testes de hipóteses a respeito de uma média populacional considerando o caso em que o desvio padrão populacional σ é desconhecido? Nessa situação a amostra deve ser usada para desenvolver uma estimativa de μ e de σ , ou seja, a partir dos dados da amostra deve-se calcular \bar{x} e s .

Como a amostra é usada para duas estimativas, Anderson, Sweeney e Williams (2011), destacam que a distribuição amostral da estatística de teste tem uma maior variabilidade. Assim, ela tem uma distribuição t de Student com $n-1$ graus de liberdade e é a seguinte:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

onde $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$ é o desvio padrão da amostra.

A regra de rejeição quando usa o critério do valor p é rejeita-se H_0 quando *valor* $p \leq \alpha$. E a regra de rejeição quando usa o critério do valor crítico, é rejeita-se H_0 se t estiver na região crítica.

A Tabela 4 apresenta as hipóteses e análises por esses critérios de acordo com o tipo de teste, se unicaudal ou bicaudal.

Tabela 4 – Características dos testes de hipóteses sobre a média da população com σ desconhecido

Teste da cauda inferior	Teste da cauda superior	Teste bicaudal
$H_0: \mu \geq \mu_0$	$H_0: \mu \leq \mu_0$	$H_0: \mu = \mu_0$
$H_a: \mu < \mu_0$	$H_a: \mu > \mu_0$	$H_a: \mu \neq \mu_0$
Valor p igual à área (probabilidade) <i>abaixo</i> de t	Valor p igual à área (probabilidade) <i>acima</i> de t	Valor p igual à área (probabilidade) sob a curva em ambas as caudas, isto é, duas vezes a probabilidade de uma das caudas, em virtude da simetria da distribuição t .
Valor crítico: rejeita-se H_0 se $t \leq -t_\alpha$	Valor crítico: rejeita-se H_0 se $t \geq t_\alpha$	Valor crítico: rejeita-se H_0 se $t \leq -t_{\alpha/2}$ ou $t \geq t_{\alpha/2}$

Fonte: Elaborada pela autora conforme Anderson, Sweeney e Williams (2011).

Exemplo 23: (ANDERSON; SWEENEY; WILLIAMS, 2011, p. 329) Aeroportos que receberem uma avaliação média populacional maior que 7 serão designados como aeroportos com atendimento de alto nível. Uma amostra de 60 viajantes de um aeroporto produziu uma avaliação média amostral de 7,25 e um desvio padrão da amostra igual a 1,052. Os dados indicam que o aeroporto deveria ser designado como aeroporto de alto nível? Utilizar um nível de significância de 0,05.

Solução: Os dados fornecidos pela questão são: $n = 60$, $\bar{x} = 7,25$, $\alpha = 0,05$, $s = 1,052$. Considerando que σ é desconhecido, temos que \bar{x} tem uma distribuição t com $n - 1 = 59$ graus de liberdade. A hipótese a ser testada é:

$$H_0: \mu \leq 7$$

$$H_a: \mu > 7$$

Vamos calcular a estatística teste:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{7,25 - 7}{1,052/\sqrt{60}} \cong 1,84$$

Consultando a tabela da distribuição t de Student (Anexo II), com 60 graus de liberdade (valor mais próximo de 59), teremos 1,84 com p entre 10% e 5%. Como esse p é a soma das duas caudas e o valor p para o teste de cauda superior é a área acima de 1,84, o valor p será um percentual entre a metade desses, ou seja, entre 5% e 2,5%. Como o valor p é menor ou igual a $\alpha = 0,05$, nós rejeitamos H_0 e, portanto, esse aeroporto pode ser classificado como de alto nível.

Exemplo 24: (ANDERSON; SWEENEY; WILLIAMS, 2011, p. 333) Considere o seguinte teste de hipóteses:

$$H_0: \mu = 18$$

$$H_a: \mu \neq 18$$

Uma amostra de tamanho 48 produziu uma média amostral $\bar{x} = 17$ e um desvio padrão amostral $s = 4,5$. (A) Calcule o valor da estatística de teste e apresente a conclusão utilizando o valor p , com $\alpha = 0,05$. (B) Qual é a regra de rejeição e a sua conclusão usando o valor crítico?

Solução:

a) Temos aqui um teste bicaudal com σ desconhecido. Então:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{17 - 18}{4,5/\sqrt{48}} \cong -1,5396$$

O teste de hipóteses que estamos realizando tem 47 graus de liberdade. Consultando a tabela da distribuição t (Anexo II), o grau de liberdade mais próximo é 50. Considerando ainda a simetria da distribuição t , identificamos na linha 50 da tabela que 1,5396 está entre 20% e 10%, portanto:

$$\text{valor } p = 2 \cdot \frac{p}{2} = p \text{ (da tabela da dist. } t) = \text{entre 20\% e 10\%}$$

Temos um valor p maior que $\alpha = 0,05$. Conclusão: não rejeitar H_0 .

b) Temos $gl = 47$, aproximando ao que temos na tabela da distribuição t usaremos $gl = 50$. Dado que $\alpha = 0,05$, ele será a probabilidade p representada na coluna da referida tabela, onde encontramos, então, o valor crítico será $t_\alpha = 2,009$.

Assim, a regra de rejeição pelo critério do valor crítico é rejeitar H_0 se $t \leq -2,009$ ou $t \geq 2,009$. Como nosso t é igual a $-1,5396$ e, portanto, maior do que $-2,009$, a conclusão é não rejeitar H_0 .

Revisando

- Distribuição binomial: $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$
- Distribuição hipergeométrica: $P(X = k) = \frac{\binom{r}{k} \binom{N-r}{n-k}}{\binom{N}{n}}$
- Distribuição de Poisson: $P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$, $k = 0, 1, 2, \dots$
- A distribuição normal, simétrica e em forma de sino, é denotada por $N(\mu; \sigma)$, e tem dois parâmetros, média μ e um desvio padrão σ .
- Variável aleatória normal padrão: $Z = \frac{X - \mu}{\sigma}$ para $-\infty < x < +\infty$
- Valor esperado de \bar{x} : $E(\bar{x}) = \mu$
- Desvio padrão (erro padrão) de \bar{x} :

$$s_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \rightarrow$$

Sempre que a população for infinita ou finita e $\frac{n}{N} \leq 0,05$.

$$s_{\bar{x}} = \sqrt{\frac{N - n}{N - 1} \cdot \frac{\sigma}{\sqrt{n}}} \rightarrow$$

Sempre que a população for finita e $\frac{n}{N} > 0,05$.

- Valor esperado de \bar{p} : $E(\bar{p}) = p$

- Desvio padrão (erro padrão) de \bar{p} :

$$s_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}} \rightarrow$$

Sempre que a população for infinita ou finita e $\frac{n}{N} \leq 0,05$.

$$s_{\bar{p}} = \sqrt{\frac{N-n}{N-1}} \cdot \sqrt{\frac{p(1-p)}{n}}$$

Sempre que a população for finita e $\frac{n}{N} > 0,05$.

- Intervalo de confiança para μ , σ conhecido: $\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$
- Intervalo de confiança para μ , σ desconhecido: $\bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$ com $gl = n - 1$
- Intervalo de confiança para p : $\bar{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$
- Estatística do teste para uma média amostral, σ conhecido: $z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$
- Estatística do teste para uma média amostral, σ desconhecido:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \text{ com } gl = n - 1$$

Saiba mais

Livro: **Probabilidade – Aplicações à Estatística**, de Paul L. Meyer.

MEYER, Paul L. **Probabilidade – Aplicações à Estatística**. 2ª ed., Rio de Janeiro: Livros Técnicos e Científicos Editora S. A., 1983.

Vídeo: **Distribuição Normal de Probabilidades EP 01**, de Marcos Murakami

DISTRIBUIÇÃO NORMAL de probabilidade - EP 1. Publicado pelo canal Prof. Marakami - Matemática Rapidola. Disponível em: <https://www.youtube.com/watch?v=Cclh1bN1U9Q>. Acesso em: 14 jan. 2021.

Livro: *Statistics for Management and Economics*, de Gerald Keller.

KELLER, Gerald. *Statistics for Management and Economics*. Mason, OH, USA: Editora Thomson South-Western, Seventh Edition, 2005.

Referências

ANDERSON, David. R.; SWEENEY, Dennis J.; WILLIAMS, Thomas A. **Estatística Aplicada à Administração e Economia**. 2ª ed. São Paulo: Cengage Learning, 2011.

BUSSAB, W.; MORETTIN, P. **Estatística Básica**. 6ª ed. São Paulo: Editora Saraiva, 2010.

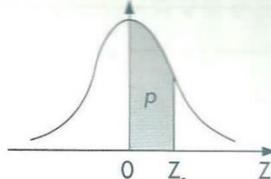
DISTRIBUIÇÃO HIPERGEOMÉTRICA. Portal Action, 2021. Disponível em: <http://www.portalaction.com.br/probabilidades/54-distribuicao-hipergeometrica>. Acesso em: 14 jan. 2021.

DOANE, David P.; SEWARD, Lori E. **Estatística Aplicada à Administração e Economia**. 4ª ed. Porto Alegre: Editora Mc Graw Hill, 2014.

LARSON, Ron; FARBER, Betsy. **Estatística Aplicada**. 4ª ed. São Paulo: Editora Pearson, 2010.

Anexo A – Distribuição Normal Padrão

Tabela III – Distribuição Normal Padrão
 $Z \sim N(0, 1)$
 Corpo da tabela dá a probabilidade p , tal que $p = P(0 < Z < Z_c)$

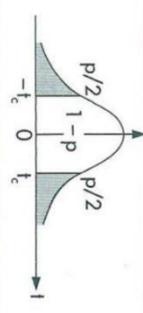


parte inteira e primeira decimal de Z_c	Segunda decimal de Z_c										parte inteira e primeira decimal de Z_c
	0	1	2	3	4	5	6	7	8	9	
	p = 0										
0,0	00000	00399	00798	01197	01595	01994	02392	02790	03188	03586	0,0
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535	0,1
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409	0,2
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173	0,3
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793	0,4
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240	0,5
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490	0,6
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524	0,7
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327	0,8
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891	0,9
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214	1,0
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298	1,1
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147	1,2
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774	1,3
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189	1,4
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408	1,5
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449	1,6
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327	1,7
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062	1,8
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670	1,9
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169	2,0
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574	2,1
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899	2,2
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158	2,3
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361	2,4
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520	2,5
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643	2,6
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736	2,7
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807	2,8
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861	2,9
3,0	49865	49869	49874	49878	49882	49886	49889	49893	49897	49900	3,0
3,1	49903	49906	49910	49913	49916	49918	49921	49924	49926	49929	3,1
3,2	49931	49934	49936	49938	49940	49942	49944	49946	49948	49950	3,2
3,3	49952	49953	49955	49957	49958	49960	49961	49962	49964	49965	3,3
3,4	49966	49968	49969	49970	49971	49972	49973	49974	49975	49976	3,4
3,5	49977	49978	49978	49979	49980	49981	49981	49982	49983	49983	3,5
3,6	49984	49985	49985	49986	49986	49987	49987	49988	49988	49989	3,6
3,7	49989	49990	49990	49990	49991	49991	49992	49992	49992	49992	3,7
3,8	49993	49993	49993	49994	49994	49994	49994	49995	49995	49995	3,8
3,9	49995	49995	49996	49996	49996	49996	49996	49996	49997	49997	3,9
4,0	49997	49997	49997	49997	49997	49997	49998	49998	49998	49998	4,0
4,5	49999	50000	50000	50000	50000	50000	50000	50000	50000	50000	4,5

Fonte: Bussab e Morettin, 2010, p. 511.

Anexo B – Distribuição t de Student

Tabela V – Distribuição t de Student
Corpo da tabela dá os valores t_c tais que $P(-t_c < t < t_c) = 1 - p$.
Para $v > 120$, usar a aproximação normal.



Graus de liberdade v	Corpo da tabela dá os valores t_c tais que $P(-t_c < t < t_c) = 1 - p$. Para $v > 120$, usar a aproximação normal.													Graus de liberdade v	
	p = 90%	80%	70%	60%	50%	40%	30%	20%	10%	5%	4%	2%	1%		0,2%
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	15,894	31,821	63,657	318,309	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	4,849	6,965	9,925	22,327	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	3,482	4,541	5,841	10,214	12,924
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	2,998	3,747	4,604	7,173	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	2,756	3,365	4,032	5,893	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	2,612	3,143	3,707	5,208	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,517	2,998	3,499	4,785	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,449	2,896	3,355	4,501	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,398	2,821	3,250	4,297	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,359	2,764	3,169	4,144	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,328	2,718	3,106	3,025	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,303	2,681	3,055	3,930	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,282	2,650	3,012	3,852	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,264	2,624	2,977	3,787	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,248	2,602	2,947	3,733	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,235	2,583	2,921	3,686	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,224	2,567	2,898	3,646	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,214	2,552	2,878	3,610	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,205	2,539	2,861	3,579	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,197	2,528	2,845	3,552	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,189	2,518	2,831	3,527	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,183	2,508	2,819	3,505	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,177	2,500	2,807	3,485	3,768
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,172	2,492	2,797	3,467	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,166	2,485	2,787	3,450	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,162	2,479	2,779	3,435	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,158	2,473	2,771	3,421	3,690
28	0,127	0,256	0,389	0,530	0,684	0,855	1,056	1,313	1,701	2,048	2,154	2,467	2,763	3,408	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,150	2,462	2,756	3,396	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,147	2,457	2,750	3,385	3,646
35	0,126	0,255	0,388	0,529	0,682	0,852	1,052	1,306	1,690	2,030	2,133	2,448	2,724	3,340	3,591
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,123	2,438	2,704	3,307	3,551
50	0,126	0,254	0,387	0,528	0,679	0,849	1,047	1,299	1,676	2,009	2,109	2,403	2,678	3,261	3,496
60	0,126	0,254	0,387	0,527	0,679	0,848	1,045	1,296	1,671	2,000	2,099	2,390	2,660	3,232	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,076	2,358	2,617	3,160	3,373
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,054	2,326	2,576	3,090	3,291

Fonte: Bussab e Morettin, 2010, p. 513.