

# TEORIA E PRÁTICA EM ESTATÍSTICA PARA CURSOS DE GRADUAÇÃO

FERNANDA CRISTINA WIEBUSCH SINDELAR  
SAMUEL MARTIM DE CONTO  
LUCILDO AHLERT



Fernanda Cristina Wiebusch Sindelar  
Samuel Martim de Conto  
Lucildo Ahlert

# Teoria e prática em estatística para cursos de graduação

1ª Edição



Lajeado, 2014



### **Centro Universitário UNIVATES**

Reitor: Ney José Lazzari

Pró-Reitor de Pesquisa, Extensão e Pós-Graduação: Carlos Cândido da Silva Cyrne

Pró-Reitora de Ensino: Luciana Carvalho Fernandes

Pró-Reitora de Ensino Adjunta: Daiani Clesnei da Rosa

Pró-Reitora de Desenvolvimento Institucional: Júlia Elisabete Barden

Pró-Reitor Administrativo: Oto Roberto Moerschbaeher



### **Editora Univates**

**Coordenação e Revisão Final:** Ivete Maria Hammes

**Editoração e Capa:** Glauber Röhrig e Marlon Alceu Cristófoli

### **Conselho Editorial da Editora Univates**

#### **Titulares**

Adriane Pozzobon

Augusto Alves

Beatris Francisca Chemin

Fernanda Cristina Wiebusch Sindelar

#### **Suplentes**

Simone Morelo Dal Bosco

Ieda Maria Giongo

Rogério Schuck

Ari Künzel

Avelino Tallini, 171 - Bairro Universitário - Lajeado - RS - Brasil

Fone: (51) 3714-7024 / Fone/Fax: (51) 3714-7000

E-mail: [editora@univates.br](mailto:editora@univates.br) / <http://www.univates.br/editora>

---

S615 Sindelar, Fernanda Cristina Wiebusch

Teoria e prática em estatística para cursos de graduação / Fernanda Cristina Wiebusch Sindelar, Samuel Martim de Conto, Lucildo Ahlert - Lajeado : Editora da Univates, 2014.

200 p.

ISBN 978-85-8167-089-8

1. Estatística 2. Probabilidade I. Título

CDU: 519.2

---

Catálogo na publicação - Biblioteca da Univates

**As opiniões e os conceitos emitidos, bem como a exatidão, adequação e procedência das citações e referências, são de exclusiva responsabilidade dos autores.**

## APRESENTAÇÃO

A presente obra contempla os tópicos mais importantes da estatística abordados em cursos de Graduação da área de Gestão. Parte de uma introdução geral à compreensão das finalidades e dos instrumentos de estatística. Visa a permitir ao aluno iniciante nos estudos de estatística adquirir conhecimentos sobre os métodos utilizados na transformação de dados em informações, para que possa fazer comparações e análises sobre fenômenos coletivos.

O seu conteúdo foi organizado tendo como base os materiais utilizados em sala de aula e a vasta experiência dos professores envolvidos, primando com exemplos práticos que se aproximam da realidade dos alunos. Assim, o texto, escrito de forma simples, apresentando-se de forma acessível, destina-se, como livro didático, para os alunos dos cursos superiores em que são ministradas disciplinas de Estatística.

Os diversos capítulos do livro são estruturados contendo conceitos, fórmulas de cálculo, exemplos de aplicabilidade, exercícios resolvidos e a resolver, e resumo em forma de tópicos para fixação e para fechamento do capítulo.

Nesse sentido, as demonstrações teóricas limitam-se ao essencial para não prejudicar a compreensão dos processos e conceitos subjacentes aos vários temas abordados. Os vários exemplos resolvidos consubstanciam e enriquecem o desenvolvimento técnico de cada capítulo, facilitando a absorção e a aplicação dos métodos e técnicas expostos.

A obra é dividida em nove capítulos. Apresenta como ponto inicial uma introdução geral à compreensão da estatística, a começar com um relato dos principais momentos que resultaram na formação da realidade atual da estatística. Segue com a apresentação de conceitos como: estatística, população, amostra, dados estatísticos com as suas diferentes formas e escalas de medidas, regras de arredondamento e caráter de variabilidade contínua e discreta, além de outros aspectos fundamentais.

No capítulo seguinte, buscando capacitar o aluno na coleta, agrupamento e apresentação de dados, o livro trata da organização de tabelas focalizando as tabulações de dados agrupados com as regras de apresentação da Associação Brasileira de Normas Técnicas – ABNT, bem como os elementos essenciais das distribuições e os tipos principais de frequências, fazendo referência, também, aos diferentes tipos de séries estatísticas.

No Capítulo 3 são apresentadas as modalidades gráficas de apresentação de dados para distribuições de frequência, gráficos cartesianos básicos e gráficos alternativos. Em terceiro lugar, apresenta as principais ferramentas utilizadas na caracterização de uma população estatística.

O Capítulo 4 é dedicado à análise das medidas de posição, como a média, a mediana e a moda.

O Capítulo 5 foi reservado para o estudo das medidas de dispersão ou de variabilidade, como, por exemplo, a variância, o desvio-padrão e o coeficiente de variação.

Ainda fazendo parte da estatística descritiva, é apresentado o Capítulo 6 com análise das medidas de assimetria e curtose que permitem caracterizar distribuições de frequência.

Nos próximos três capítulos são apresentados os elementos básicos das teorias de probabilidade para dados discretos e contínuos, usados na inferência estatística.

Assim, o Capítulo 7 trata da teoria elementar da probabilidade com os seus principais conceitos e regras aplicáveis a variáveis discretas. Já no Capítulo 8 são apresentadas, com as suas características de uso, as duas distribuições básicas para variáveis discretas, a distribuição Binomial para uso em populações finitas e a distribuição de Poisson em populações infinitas relacionadas com tempo, volume e área.

Completando o estudo básico da probabilidade, no Capítulo 9 é discutido e desenvolvido o conteúdo relacionado à distribuição Normal para dados contínuos.

Ao final do texto encontram-se as tabelas estatísticas para utilização nos tópicos apresentados.

**Os autores**

## COMO ESTUDAR ESTATÍSTICA

**Estudando estatística** - você está prestes a iniciar seu estudo de estatística. À medida que avançar na disciplina, você descobrirá como usar estatística em sua vida e em sua atividade profissional. Ao estudar estatística, é necessário ter uma mente aberta e entender que a matéria que você aprende a cada dia se soma àquilo que aprendeu anteriormente. Não existem atalhos – você deve manter-se atualizado com os seus estudos diariamente.

**Criando um plano** - trace seu próprio plano de estudo. Defina algumas horas do seu horário livre para estudar e fazer exercícios referentes a cada encontro semanal. Depois de seu primeiro teste, saberá se seus esforços foram suficientes. Se não obtiver a nota que queria, você deve aumentar seu tempo de estudo, melhorar sua eficiência no estudo ou ambos.

**Preparando-se para a aula** - antes de cada aula, reveja as anotações da aula anterior e leia a parte do livro que será abordada. Leia os exemplos e trabalhe os exercícios sugeridos.

**Frequentando as aulas** - frequente todas as aulas. Chegue no horário certo, com o seu livro, o seu caderno de anotações e a calculadora. Se você perder uma aula, procure um colega, pegue as anotações dele e lhe peça orientações sobre o conteúdo trabalhado. Tente aprender a matéria que foi abordada na aula em que faltou, antes de participar da aula seguinte.

**Participando da aula** - leia o livro, reveja as suas anotações de uma aula anterior e faça os exercícios. Anote quaisquer dúvidas a respeito da matéria. Exponha-as ao seu professor durante a aula. Isso o ajudará, e também aos demais alunos da classe, a entender melhor a matéria.

**Fazendo anotações** - durante a aula, certifique-se de anotar definições, exemplos, conceitos e regras. Concentre-se nas dicas do professor para entender a matéria. Então, assim que for possível, depois da aula, reveja suas anotações e acrescente quaisquer explicações que ajudem a tornar suas anotações mais compreensíveis para você.

**Fazendo exercícios** - aprender estatística é como aprender a tocar piano ou a jogar basquete. Não se consegue desenvolver habilidades simplesmente observando alguém fazendo; você mesmo deve desenvolver atividades. A melhor hora para fazer seus exercícios é imediatamente depois da aula, quando os conceitos ainda estão recentes em sua mente. Fazer os exercícios nesse momento aumenta suas chances de guardar a informação na memória por muito tempo.

**Encontrando um colega de estudo** - quando “emperrar” num problema, talvez você ache válido trabalhar com um colega. Mesmo que ache que dará mais ajuda do que receberá, descobrirá que ensinar outras pessoas é uma excelente maneira de aprender.

**Mantendo-se em dia com a matéria** - não fique para trás nesta disciplina. Se tiver problemas, procure ajuda imediatamente – de seu professor, de um orientador em estatística, de seu colega de estudo ou de materiais de estudos auxiliares. Lembre-se: se tiver problemas com um assunto do seu livro, há uma grande probabilidade de ter problemas com capítulos posteriores, a não ser que tome providências para melhorar seu entendimento.

**“Emperrando”** - todo estudante de estatística já passou por esta experiência: você trabalha num problema e não consegue resolvê-lo. Quando isso acontecer, peça ajuda ou faça uma pausa para aclarar suas ideias. Você pode até mesmo querer deixar o problema para mais tarde ou reler o capítulo do livro. Evite sentir-se frustrado ou gastar tempo demais num único problema.

**Fazendo um teste** – não se recomenda estudar até o último minuto antes do início da prova. Isso tende a tornar as pessoas ansiosas. A melhor cura para a ansiedade quando se faz provas é preparar-se bem antecipadamente. Assim que a prova iniciar, leia cuidadosamente as orientações e trabalhe num ritmo razoável. Sugere-se ler a prova por inteiro primeiro e depois iniciar a solução dos problemas na ordem dos mais fáceis aos mais difíceis. Não se apresse! Pessoas que se apressam tendem a cometer erros por falta de atenção. Se terminar cedo, dedique alguns minutos para aclarar suas ideias e depois reexamine cuidadosamente seu trabalho. Na dúvida, mantenha sempre a primeira resposta. Somente faça a alteração quando tiver convicção de que a resposta anterior realmente não está correta.

**Aprendendo com os erros** – depois que sua prova lhe for devolvida, reveja quaisquer erros que possa ter cometido. Isso o ajudará a evitar a repetição de alguns erros sistemáticos e conceituais. Não rejeite nenhum erro simplesmente como um “erro tolo”. Tire proveito de quaisquer erros procurando maneiras de melhorar suas habilidades para participar de testes ou provas (LARSON; FARBER, 2006).

## SUMÁRIO

<b>1 INTRODUÇÃO À COMPREENSÃO DA ESTATÍSTICA .....</b>	<b>13</b>
<b>1.1 Conhecimentos gerais.....</b>	<b>13</b>
<b>1.1.1 A estatística como ciência e como instrumento .....</b>	<b>13</b>
<b>1.1.2 História .....</b>	<b>13</b>
<b>1.1.3 Uso prático .....</b>	<b>15</b>
<b>1.2 Conceitos .....</b>	<b>16</b>
<b>1.2.1 Conceito geral .....</b>	<b>16</b>
<b>1.2.2 Conceito especificado .....</b>	<b>16</b>
<b>1.2.3 Coleta de dados .....</b>	<b>16</b>
<b>1.2.4 Classificação .....</b>	<b>16</b>
<b>1.2.5 Apresentação e organização de dados.....</b>	<b>16</b>
<b>1.2.6 Análise .....</b>	<b>17</b>
<b>1.2.7 Interpretação .....</b>	<b>17</b>
<b>1.2.8 Estatística descritiva .....</b>	<b>17</b>
<b>1.2.9 Estatística Indutiva .....</b>	<b>17</b>
<b>1.3 Fases que antecedem a Estatística.....</b>	<b>18</b>
<b>1.3.1 Definição do problema .....</b>	<b>18</b>
<b>1.3.2 Planejamento .....</b>	<b>18</b>
<b>1.4 Fenômenos estatísticos .....</b>	<b>19</b>
<b>1.5 Dados estatísticos .....</b>	<b>19</b>
<b>1.5.1 Primários.....</b>	<b>19</b>
<b>1.5.2 Secundários.....</b>	<b>19</b>
<b>1.6 Algumas classes específicas de variáveis .....</b>	<b>20</b>
<b>1.6.1 Características demográficas e socioeconômicas .....</b>	<b>20</b>
<b>1.6.2 Estilo de vida .....</b>	<b>20</b>
<b>1.6.3 Atitudes .....</b>	<b>21</b>
<b>1.6.4 Conscientização e conhecimento .....</b>	<b>21</b>
<b>1.6.5 Motivação.....</b>	<b>21</b>
<b>1.6.6 Comportamento .....</b>	<b>22</b>
<b>1.6.7 Intenções .....</b>	<b>22</b>

<b>1.7 Classificação das variáveis.....</b>	<b>22</b>
<b>1.7.1 Variável qualitativa.....</b>	<b>22</b>
<b>1.7.2 Variável quantitativa .....</b>	<b>22</b>
<b>1.8 Escalas de medidas de variáveis qualitativas e quantitativas .....</b>	<b>23</b>
<b>1.8.1 Escala nominal .....</b>	<b>23</b>
<b>1.8.2 Escala ordinal.....</b>	<b>23</b>
<b>1.8.3 Escala de intervalo (ou intervalar).....</b>	<b>23</b>
<b>1.8.4 Escalas de razão .....</b>	<b>24</b>
<b>1.9 Escalas de medidas de transformação de variável qualitativa .....</b>	<b>24</b>
<b>1.9.1 Escalas nominais .....</b>	<b>24</b>
<b>1.9.2 Escalas de avaliação .....</b>	<b>24</b>
<b>1.9.3 Escalas de ordenação .....</b>	<b>25</b>
<b>1.9.4 Escalas comparativas de avaliação .....</b>	<b>25</b>
<b>1.9.5 Escalas de comparação pareada.....</b>	<b>26</b>
<b>1.9.6 Escalas de diferencial semântico.....</b>	<b>26</b>
<b>1.9.7 Escalas de Stapel.....</b>	<b>26</b>
<b>1.9.8 Escalas indiretas.....</b>	<b>26</b>
<b>1.10 Arredondamento de dados .....</b>	<b>27</b>
<b>1.10.1 Arredondamento por falta .....</b>	<b>27</b>
<b>1.10.2 Arredondamento por excesso .....</b>	<b>27</b>
<b>1.10.3 Arredondamento de dígitos seguidos de cinco.....</b>	<b>27</b>
<b>1.11 População e amostra .....</b>	<b>28</b>
<b>1.11.1 População ou Universo Estatístico .....</b>	<b>28</b>
<b>1.11.2 Amostra.....</b>	<b>28</b>
<b>1.12 Levantamento de dados .....</b>	<b>28</b>
<b>1.12.1 Dados brutos .....</b>	<b>28</b>
<b>1.12.2 Rol .....</b>	<b>28</b>
<b>1.12.3 Amplitude amostral.....</b>	<b>29</b>
<b>1.13 Resumo para fixação .....</b>	<b>30</b>
<b>1.14 Exercícios propostos.....</b>	<b>31</b>

<b>2 TABELAS.....</b>	<b>32</b>
<b>2.1 Tabelas de frequência .....</b>	<b>32</b>
<b>2.1.1 Elementos de uma tabela de frequências.....</b>	<b>32</b>
<b>2.1.2 Tipos de intervalos de classe .....</b>	<b>34</b>
<b>2.1.3 Determinação do número de intervalos de classe .....</b>	<b>36</b>
<b>2.1.4 Normas para apresentação de tabelas .....</b>	<b>37</b>
<b>2.1.5 Roteiro para a elaboração de uma tabela de frequências.....</b>	<b>39</b>
<b>2.1.6 Tipos de frequência para análise de tabelas .....</b>	<b>40</b>
<b>2.2 Tabelas de séries estatísticas.....</b>	<b>41</b>
<b>2.2.1 Séries históricas, cronológicas, temporais ou marchas.....</b>	<b>42</b>
<b>2.2.2 Séries geográficas, espaciais, territoriais ou de localização .....</b>	<b>42</b>
<b>2.2.3 Séries específicas ou categóricas.....</b>	<b>42</b>
<b>2.2.4 Tabelas de séries complexas.....</b>	<b>43</b>
<b>2.3 Resumo para fixação .....</b>	<b>45</b>
<b>2.4 Exercícios propostos.....</b>	<b>46</b>
<b>3 GRÁFICOS.....</b>	<b>48</b>
<b>3.1 Representação gráfica de uma distribuição de frequência .....</b>	<b>48</b>
<b>3.1.1 Histograma .....</b>	<b>48</b>
<b>3.1.2 Polígono de frequência .....</b>	<b>49</b>
<b>3.1.3 Curva de frequência.....</b>	<b>49</b>
<b>3.2 Gráficos cartesianos básicos .....</b>	<b>50</b>
<b>3.2.1 Gráfico de colunas .....</b>	<b>50</b>
<b>3.2.2 Diagrama de Pareto .....</b>	<b>51</b>
<b>3.2.3 Gráfico de linhas .....</b>	<b>52</b>
<b>3.2.4 Diagrama de dispersão.....</b>	<b>53</b>
<b>3.3 Normas de apresentação de ilustrações .....</b>	<b>53</b>
<b>3.4 Gráficos de séries complexas .....</b>	<b>56</b>
<b>3.5 Outros gráficos .....</b>	<b>57</b>
<b>3.5.1 Gráficos de setores ou de “pizza” .....</b>	<b>57</b>
<b>3.5.2 Gráfico de barras .....</b>	<b>57</b>
<b>3.5.3 Gráficos de colunas transformadas:.....</b>	<b>58</b>
<b>3.5.4 Gráficos pictóricos .....</b>	<b>59</b>



3.5.5 Cartogramas .....	60
3.6 Dicas de utilização .....	61
3.7 Exemplo de como elaborar um gráfico de pizza .....	61
3.8 Resumo para fixação .....	63
3.9 Exercícios propostos.....	64
<b>4 MEDIDAS DE TENDÊNCIA CENTRAL.....</b>	<b>66</b>
4.1 Média .....	66
4.1.1 Média aritmética de dados observados .....	67
4.1.2 Média aritmética para dados tabulados .....	68
4.1.3 Propriedades da média aritmética.....	71
4.2 Mediana.....	72
4.2.1 Mediana para dados brutos .....	72
4.2.2 Mediana para dados agrupados .....	73
4.2.3 Outras formas de separatrizes .....	74
4.3 Moda.....	76
4.3.1 Moda para dados brutos .....	76
4.3.2 Moda para dados agrupados .....	77
4.4 Relação entre a média, a mediana e a moda .....	79
4.5 Resumo para fixação .....	81
4.6 Exercícios propostos.....	82
<b>5 MEDIDAS DE DISPERSÃO OU DE VARIABILIDADE .....</b>	<b>84</b>
5.1 Desvio médio.....	84
5.2 Variância .....	86
5.3 Desvio-padrão .....	89
5.3.1 Utilização do Desvio-padrão em Estatística.....	91
5.3.2 Desigualdade de Chebyshev .....	93
5.4 Coeficiente de variabilidade ou de variação .....	93
5.4.1 Características .....	94
5.5 Amplitude interquartílica.....	95
5.6 Resumo para fixação .....	97
5.7 Exercícios propostos.....	98

<b>6 MEDIDAS DE ASSIMETRIA E CURTOSE.....</b>	<b>100</b>
6.1 Medidas de assimetria.....	100
6.1.1 Coeficiente de Assimetria .....	101
6.2 Medidas de curtose.....	106
6.3 Resumo para fixação .....	109
6.4 Exercícios propostos.....	110
<b>7 TEORIA ELEMENTAR DE PROBABILIDADE.....</b>	<b>112</b>
7.1 Introdução.....	112
7.2 Conceitos .....	112
7.2.1 Clássico ou conceito de Laplace.....	112
7.2.2 Conceito moderno .....	113
7.2.3 Conceito subjetivo .....	113
7.3 Conceitos básicos .....	113
7.3.1 Experimento aleatório .....	113
7.3.2 Espaço amostral .....	114
7.3.3 Evento .....	114
7.4 Definições axiomáticas de probabilidade.....	115
7.5 Axiomas .....	115
7.6 Teoremas .....	116
7.7 Avaliação de probabilidades.....	117
7.8 Exercícios resolvidos .....	118
7.9 Cálculos acessórios .....	122
7.9.1 Permutações.....	122
7.9.2 Combinações.....	122
7.9.3 Combinação especial .....	123
7.10 Teorema de <i>Bayes</i> .....	123
7.11 Distribuição de probabilidade.....	124
7.12 Resumo para fixação .....	125
7.13 Exercícios propostos.....	126

<b>8 DISTRIBUIÇÕES BÁSICAS DE PROBABILIDADE DE VARIÁVEIS DISCRETAS .....</b>	<b>127</b>
8.1 Distribuição de probabilidade.....	127
8.2 Distribuição binomial.....	127
8.2.1 Conceito .....	128
8.2.2 Método de Bernoulli.....	128
8.2.3 Fórmula de cálculo .....	129
8.2.4 Tabela de Probabilidade Binomial.....	130
8.2.5 Propriedades da Distribuição Binomial .....	134
8.3 Distribuição de <i>Poisson</i> .....	134
8.3.1 Características da Distribuição de <i>Poisson</i> .....	134
8.3.2 Fórmula de cálculo .....	135
8.3.3 Tabela da Probabilidade de Poisson.....	135
8.4 Resumo para fixação .....	138
8.5 Exercícios propostos.....	139
<b>9 DISTRIBUIÇÃO BÁSICA DE PROBABILIDADE DE VARIÁVEIS CONTÍNUAS .....</b>	<b>140</b>
9.1 Função densidade de probabilidade.....	140
9.2 Características da Distribuição Normal .....	141
9.3 Fórmula de cálculo da Função densidade.....	142
9.4 Determinação da probabilidade da Distribuição Normal padronizada .....	144
9.4.1 Determinação da probabilidade em função de $X_i$ , sendo a média e o desvio-padrão conhecidos .....	145
9.4.2 Determinação de $X_i$ , em função da probabilidade desejada, conhecendo a média e o desvio-padrão .....	157
9.5 Relação entre as Distribuições Binomial e Normal .....	161
9.5.1 Cálculo da aproximação do valor de $z$ .....	161
9.6 Resumo para fixação .....	168
9.7 Exercícios propostos.....	169

<b>REFERÊNCIAS.....</b>	<b>172</b>
<b>LISTA DE APÊNDICES .....</b>	<b>173</b>
<b>APÊNDICE I – Distribuição das Probabilidades Binomiais (1) .....</b>	<b>174</b>
<b>APÊNDICE I – Distribuição das Probabilidades Binomiais (2) .....</b>	<b>175</b>
<b>APÊNDICE I – Distribuição das Probabilidades Binomiais (3) .....</b>	<b>176</b>
<b>APÊNDICE I – Distribuição das Probabilidades Binomiais (4) .....</b>	<b>177</b>
<b>APÊNDICE I – Distribuição das Probabilidades Binomiais (5) .....</b>	<b>178</b>
<b>APÊNDICE I – Distribuição das Probabilidades Binomiais (6) .....</b>	<b>179</b>
<b>APÊNDICE I – Distribuição das Probabilidades Binomiais (7) .....</b>	<b>180</b>
<b>APÊNDICE I – Distribuição das Probabilidades Binomiais (8) .....</b>	<b>181</b>
<b>APÊNDICE I – Distribuição das Probabilidades Binomiais (9) .....</b>	<b>182</b>
<b>APÊNDICE II – Distribuição da probabilidades de Poisson (1).....</b>	<b>183</b>
<b>APÊNDICE II – Distribuição da probabilidades de Poisson (2).....</b>	<b>183</b>
<b>APÊNDICE II – Distribuição da probabilidades de Poisson (3).....</b>	<b>184</b>
<b>APÊNDICE II – Distribuição da probabilidades de Poisson (4).....</b>	<b>185</b>
<b>APÊNDICE II – Distribuição da probabilidades de Poisson (5).....</b>	<b>186</b>
<b>APÊNDICE III – Distribuição Normal Padronizada - z .....</b>	<b>187</b>
<b>APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 1 .....</b>	<b>188</b>
<b>APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 2 .....</b>	<b>189</b>
<b>APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 3 .....</b>	<b>191</b>
<b>APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 4 .....</b>	<b>195</b>
<b>APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 5 .....</b>	<b>195</b>
<b>APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 6 .....</b>	<b>196</b>
<b>APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 7 .....</b>	<b>196</b>
<b>APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 8 .....</b>	<b>197</b>
<b>APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 9 .....</b>	<b>198</b>

# 1 INTRODUÇÃO À COMPREENSÃO DA ESTATÍSTICA

## 1.1 Conhecimentos gerais

### 1.1.1 A estatística como ciência e como instrumento

Como do cotidiano das pessoas faz parte tomar decisões, que na maioria das vezes devem ser ágeis, há a necessidade de dispor de informações que embasam decisões mais acertadas. Nesse contexto, a estatística, que extrai informações dos dados para obter melhor compreensão das situações, representa uma técnica de grande alcance e de larga utilização.

Com o desenvolvimento e o aperfeiçoamento de técnicas e métodos, a estatística transformou-se em uma ciência que dispõe de processos apropriados para recolher, organizar, classificar, apresentar e interpretar conjuntos de dados, tendo, assim, como objetivo orientar a coleta, o resumo, a análise e a interpretação de dados.

Para as demais ciências que necessitam de uma base de dados quantitativos para as suas conclusões científicas, a estatística é considerada, no entanto, como um ramo da Matemática Aplicada dedicada à análise e tomada de decisão com dados observados por meio do método científico. Nessas ciências, a estatística serve como um instrumento que permite o tratamento quantitativo das observações para resumir e transformar em informações e comparações os dados de um fenômeno em análise.

### 1.1.2 História

A palavra estatística tem sua origem associada à palavra latina *STATUS*, que quer dizer Estado ou, também, situação de uma coisa. Apesar de existirem indícios sobre a realização de censos na Babilônia, China e Egito desde 3000 anos A.C. e das referências no 4º Livro do Velho Testamento relativas ao levantamento dos homens aptos para guerrear na Palestina, ordenado pelo imperador romano César Augusto, ao redor do ano zero da Era Cristã, apenas no século XVII a Estatística passou a ser considerada uma disciplina autônoma.

A designação atual se originou do termo *STATENKUNDE* (Ciência do Estado) da escola alemã. É eminentemente qualitativa, desenvolvida por meio das ideias de A. L. Von Schölzer, que partiu dos estudos do acadêmico alemão Gottfried Achenwal (1719 - 1772) e de Hermann Conrig (1606 - 1681).

Na Inglaterra, no século XVII, John Graunt (1620 - 1674) e William Petty (1623 - 1687) preocuparam-se com o estudo numérico dos fenômenos sociais e políticos, na busca de leis que pudessem explicá-los. Nesses estudos eram analisados os nascimentos e as mortes, por meio das Tábuas de Mortalidade. Essas Tábuas deram origem às atuais que são utilizadas pelas companhias de seguros. Eles formaram a escola dos aritméticos políticos e podem ser considerados o berço da Demografia. O alemão Süssmilch (1707 -1767) foi um dos adeptos e, com ele, aparece pela primeira vez a Estatística como meio indutivo de investigação.

Na última metade do século XIX, foram obtidos importantes resultados que contribuíram para o desenvolvimento da Inferência Estatística. O impulso decisivo deve-se, porém, a Karl Pearson (1857 - 1936), William Gosset (1876 - 1937) e Ronald Fisher (1890 - 1962).

Karl Pearson dedicou-se ao estudo da evolução de Darwin, em que aplicou métodos estatísticos aos problemas biológicos relacionados com a evolução e a hereditariedade. Contribuiu extremamente para o desenvolvimento da teoria da Análise de Regressão e do Coeficiente de Correlação, bem como do teste de hipótese do Qui-quadrado. Ele foi o responsável pelo reconhecimento da Estatística como disciplina autônoma.

William Gosset, que trabalhava como químico em uma cervejaria, devido à necessidade de manipular dados provenientes de pequenas amostras, extraídas para avaliar e melhorar a qualidade da cerveja, derivou o teste  $t$  de *Student*, baseado na distribuição de probabilidades  $t$ . Gosset. Publicou os resultados na revista *Biometrika* em 1908, sob o pseudônimo de *Student*, pois a cervejaria não desejava revelar aos concorrentes os métodos estatísticos que estava utilizando no controle de qualidade da cerveja.

Ronald Fisher contribuiu tanto para o desenvolvimento da Genética quanto para a Estatística. Ele apresentou os princípios de planejamento e de experimentos, introduzindo os conceitos de aleatorização e da Análise da Variância, procedimentos que são muito utilizados ultimamente. Foi também o responsável por estabelecer o que a maioria aceita como a estrutura da moderna Estatística Analítica, por meio do conceito da verossimilhança.

A teoria das Probabilidades é considerada outra área de investigação extremamente importante para o desenvolvimento da Estatística. Costuma-se atribuir a origem do Cálculo de Probabilidade a questões relacionadas aos jogos de azar que o célebre cavaleiro Meré (1697 - 1684) encaminhou a Blaise Pascal (1623 - 1662). No entanto, alguns autores atribuem a sua origem à Itália, em especial, a Luca Pacioli (1445 - 1517), Girolamo Cardano (1501 - 1576), Nicolo Tartaglia (1500 - 1557) e Galileu Galilei (1564 - 1642).

Christian Huygens (1629 - 1695) publicou o primeiro livro sobre o Cálculo de Probabilidades, no qual introduziu o conceito de Esperança Matemática.

Jacques Bernoulli (1654 - 1705) dedicou-se ao estudo do Cálculo de Probabilidade e à Lei dos Grandes Números de Bernoulli, que foi considerado o primeiro teorema do limite. Graças as suas contribuições, o Cálculo de Probabilidade adquiriu o *Status* de Ciência.

Ao matemático Thomas Bayes (1702 - 1761) deve-se o conceito de probabilidade inversa, relacionado com situações em que se parte do particular para o geral. Bayes formula, por meio do teorema e do postulado que leva o seu nome, a primeira tentativa de matematização da inferência estatística.

Os astrônomos Pierre-Simon Laplace (1749 - 1827), Johann Gauss (1777 - 1855) e Lambert Quetelet (1796 - 1874) foram fundamentais para o desenvolvimento do Cálculo de Probabilidades. Gauss apresentou uma teoria sobre a análise de observações aplicável a qualquer ramo da ciência, alargando o campo de aplicação do Cálculo de Probabilidades. Quetelet iniciou a aplicação aos fenômenos sociais.

Antoine Augustin Cournot (1801 - 1877) observou a importância da Teoria das Probabilidades na análise estatística. Foi o pioneiro no tratamento matemático dos fenômenos econômicos.

Um dos pontos mais altos da Teoria das Probabilidades foi atingido na segunda metade do século XIX, com os trabalhos da escola russa, em que o seu maior expoente foi o russo Kolmogorov (1903 - 1987).

Com o início da competitividade concorrencial, Montgomery introduziu a metodologia de qualidade, destacando os aspectos que envolvem o controle de qualidade estatístico. A partir disso, os métodos estatísticos e suas aplicações em melhoria de qualidade tiveram uma longa história. Em 1924, Walter A. Shewhart (*Bell Telephone Laboratories*) desenvolveu o conceito estatístico de carta de controle. Isso é considerado o começo formal do controle de qualidade estatístico.

Com a introdução dos princípios de administração científica por Frederick W. Taylor, foram padronizados os métodos de trabalho, surgindo o conceito de trabalho-padrão - um tempo padrão para executar uma tarefa, ou um número especificado de unidades que deveriam ser produzidas por período, que exige métodos estatísticos para análise dos resultados.

A partir da Segunda Guerra Mundial, o uso de conceitos de controle de qualidade estatístico na produção industrial foi expandido e obteve maior aceitação. A experiência do tempo de guerra

tornou aparente que as técnicas estatísticas eram necessárias para controle e melhoria da qualidade do produto.

Desde 1980, houve um crescimento profundo no uso de métodos estatísticos para a melhoria da qualidade. O conceito de Administração da Qualidade Total é uma estrutura de administração na qual são implementados métodos estatísticos.

### 1.1.3 Uso prático

Na atualidade, a alta competitividade pela busca de tecnologias e de mercados tem provocado uma constante corrida pela informação. Com o advento da melhoria da qualidade e da produtividade, as empresas estão empenhadas em compreender e otimizar os processos produtivos, os quais estão constantemente gerando dados. A transformação desses dados em informações é eficientemente realizada mediante o uso de métodos estatísticos, permitindo tomadas de decisão mais seguras.

Isso ocorre pois mesmo que as máquinas usadas no processo de produção tenham sido desenvolvidas com base em alta tecnologia, elas poderão se desregular devido ao seu uso, levando os seus operadores a cometer erros. Assim, há necessidade de um controle constante para monitorar o resultado da produção de forma que problemas sejam detectados rapidamente, para que o processo de produção seja ajustado ou corrigido.

A Estatística está presente em todas as ciências que se envolvem com coleta e análise de dados quantitativos e sua consequente transformação em informação, para postular, refutar ou validar hipóteses científicas sobre um fenômeno observável. Portanto, qualquer ciência experimental não pode prescindir das técnicas proporcionadas pela Estatística. Essa é uma tendência crescente e irreversível, de maneira que a Estatística possui um papel vital na tomada de decisões e na produção de conhecimentos ou de novas tecnologias, podendo, assim, ser considerada a “tecnologia da ciência”.

A pesquisa científica é um processo interativo de acumulação de conhecimentos e envolve a formulação de hipóteses, modelos e teorias, a observação de fenômenos e a verificação e rejeição de hipóteses sobre eles. A Estatística procura tornar esse processo o mais eficiente possível, por meio de suas técnicas de coleta de dados (amostragem e planejamento de experimentos); apresentação de dados (análise exploratória e descrição: tabelas e gráficos); modelagem (probabilidade e processos estocásticos); análise indutiva (inferência: testes e estimação) e verificação (ajustamento, previsão e controle).

Assim, o método estatístico pode ser usado pelo industrial ao controlar a qualidade de um produto; pelo comerciante ao pesquisar as condições do mercado consumidor; pelo governo ao indagar aspectos da vida econômica, social, cultural, administrativa e política; pelo político ao avaliar as condições potenciais de uma campanha eleitoral; pelo pesquisador para saber se as suas hipóteses são confirmadas.

Dessa forma, a Estatística tem um campo bastante diversificado, podendo ser aplicada em diversas áreas, entre as quais: Economia, Administração de Empresas, Agronomia, Engenharia, Física, Biologia, Medicina, Agricultura, Psicologia, Sociologia, Meteorologia, Informática, Contabilidade, Comunicação, Turismo etc.



## 1.2 Conceitos

### 1.2.1 Conceito geral

Estatística pode ser conceituada como um conjunto de métodos e processos quantitativos que serve para estudar e medir fenômenos coletivos. “É uma coleção de métodos para planejar experimentos, obter dados e organizá-los, resumi-los, analisá-los, interpretá-los e deles extrair conclusões” (TRIOLA, 1999, p. 2).

### 1.2.2 Conceito especificado

“Estatística representa uma atividade humana especializada ou um corpo de técnicas ou ainda uma metodologia científica desenvolvida para a coleta, a classificação, a apresentação, a análise e a interpretação de dados quantitativos obtidos (de fatos reais) e a utilização desses dados para a tomada de decisão” (TOLEDO; OVALLE, 1992, p. 14).

### 1.2.3 Coleta de dados

A coleta de dados é uma atividade que exige tempo de dedicação e custos. Assim, a identificação das informações necessárias à investigação é sempre uma dificuldade séria. Dessa forma, deve-se sempre analisar a melhor alternativa para fazer a coleta dos dados necessários. Entre as opções de decisão, podem-se coletar os dados diretamente, quando a busca de dados ocorre *in loco* na fonte, pelo uso de questionários, ou de forma indireta, quando a busca de dados ocorre por meio de fontes secundárias, como os disponibilizados em bancos de dados, publicações em anuários e revistas, entre outros.

### 1.2.4 Classificação

Após a coleta, os dados deverão ser reunidos segundo critérios previamente estabelecidos. Ex.: raça, idade, faixa de renda, grau de instrução etc. Esses critérios são oriundos de variáveis chamadas qualitativas ou classificatórias?

### 1.2.5 Apresentação e organização de dados

É uma forma prática do estudo. É a transformação dos dados em informações que podem ser em forma de tabelas ou gráficos, conforme exemplos abaixo, de modo a facilitar a sua análise e interpretação.

Exemplo de Tabela:

Tabela 1.1 - Produto Interno Bruto do Vale do Taquari 2007/2011

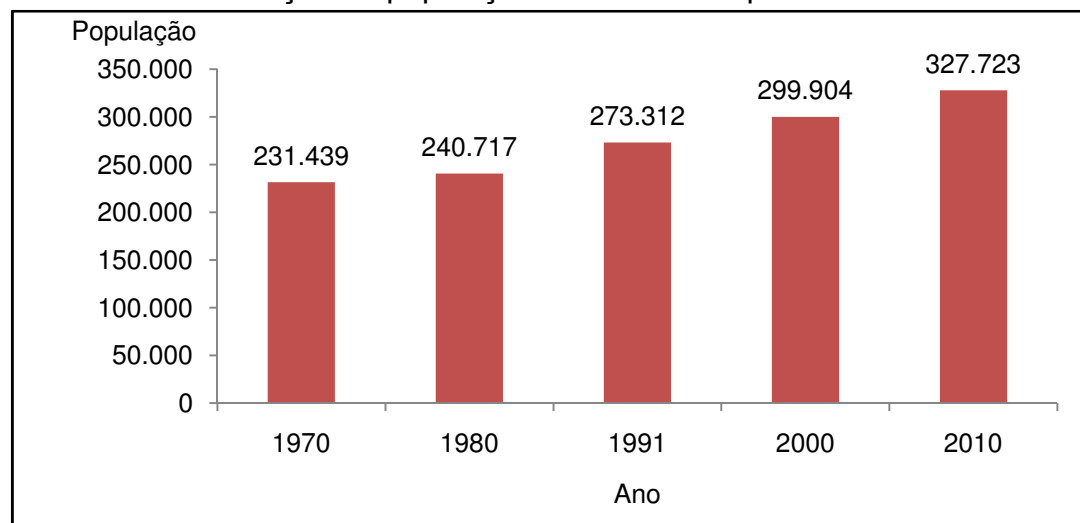
Ano	Vale do Taquari (R\$ mil)	Participação no RS (%)
2007	5.386.397	3,05
2008	5.759.721	2,89
2009	6.422.379	2,98
2010	7.739.330	3,07
2011	8.510.190	3,23

Fonte: Fundação de Economia e Estatística (FEE, 2013).



Exemplo de Gráfico:

Gráfico 1.1 – Evolução da população do Vale do Taquari – 1970/2010



Fonte: IBGE, Censos Demográficos, 1970/2010.

### 1.2.6 Análise

É realizada por meio do cálculo de medidas que permitem a descrição do fenômeno valendo-se de um número limitado de características. As medidas mais usuais são as de Tendência Central, como a média, a moda e a mediana, e de Dispersão, como o desvio-padrão, a variância e o coeficiente de variabilidade.

### 1.2.7 Interpretação

Leva a conclusões relativas a propriedades e características gerais do conjunto estudado. Refere-se à interpretação de cada um dos valores obtidos por meio do cálculo das várias medidas estatísticas disponíveis. As interpretações envolvem o entendimento das medidas calculadas e a sua generalização para a população, no caso em que os dados interpretados são oriundos de uma parcela representativa desse universo, conhecida como amostra.

### 1.2.8 Estatística descritiva

A estatística descritiva tem por objetivo a observação de fenômenos da mesma natureza para descrevê-los resumidamente. Assim, a estatística descritiva é o ramo que reúne a coleta de dados numéricos, a sua organização e classificação, a sua apresentação e a definição de características que permitem a sua análise e interpretação, que podem ser de variáveis quantitativas ou qualitativas.

Os métodos e técnicas da estatística descritiva são especialmente importantes na organização e apresentação dos resultados de pesquisas quantitativas descritivas. Os procedimentos, técnicas e métodos de estatística descritiva mais utilizados serão apresentados nos seguintes capítulos: Tabelas, Gráficos, Medidas de Tendência Central, Medidas de Dispersão e de Variabilidade e Medidas de Assimetria e Curtose.

### 1.2.9 Estatística Indutiva

A estatística indutiva ou inferência estatística refere-se a um processo de generalização a partir de amostras. Consiste em obter e generalizar conclusões, ou seja, inferir propriedades para o todo (população) com base na parte (amostra), no particular.

O processo de generalização está associado a uma margem de incerteza, que se deve ao fato de a conclusão, que se pretende obter para o conjunto de todos os indivíduos analisados quanto a determinadas características comuns, basear-se em uma parcela do total de observações. A medida de incerteza é tratada mediante técnicas e métodos que se fundamentam na Teoria da Probabilidade.

Tudo o que acontece na estatística provém de dados reais e o que possa ser observado. Todas as conclusões estatísticas se baseiam na inferência estatística que apresenta a “solidez científica”, que é o nível de certeza e incerteza dado em percentagem.

### **1.3 Fases que antecedem a Estatística**

#### **1.3.1 Definição do problema**

Saber exatamente aquilo que se pretende pesquisar, ou seja, uma pergunta que busca respostas para uma questão não solvida e que é objeto de discussão em qualquer domínio do conhecimento conceitua-se como um problema de natureza científica. Assim, um problema deve ser passível de verificação empírica.

Os problemas científicos envolvem variáveis que podem ser tidas como testáveis, não se referindo a valores subjetivos, pois esses conduzem a julgamentos morais e, conseqüentemente, a considerações subjetivas, invalidando os propósitos da investigação científica, que tem a objetividade como uma das mais importantes características.

Para que os problemas sejam passíveis de solução, é preciso que se tenha o domínio da tecnologia que permita coletar os dados necessários a sua resolução.

Nesse sentido, os meios disponíveis para a investigação forçam a delimitação do problema: em termos de área geográfica, aspectos socioeconômicos e aspectos intrínsecos do problema.

#### **1.3.2 Planejamento**

Definir que dados deverão ser obtidos são alguns aspectos a serem planejados. Como obtê-los, por meio de censo ou amostra? Como analisá-los e como proceder na inferência?

Inicialmente, deve ser elaborado um anteprojeto de pesquisa, com as definições dos instrumentos de coleta, o processo de observação (entrevista direta ou via postal), a determinação do número de entrevistadores, supervisores.

Após, concretiza-se o planejamento de pesquisa mediante a elaboração de um projeto, que é documento explicitador das ações a serem desenvolvidas ao longo de um processo de pesquisa. O projeto deve, portanto, especificar os objetivos da pesquisa, apresentar a justificativa de sua realização, definir a modalidade de pesquisa e determinar procedimentos de coleta e análise de dados. Deve, ainda, esclarecer acerca do cronograma a ser seguido no desenvolvimento da pesquisa e proporcionar a indicação de recursos humanos, financeiros e materiais necessários para assegurar o êxito da pesquisa.

Um projeto de pesquisa apresenta o roteiro de ações para o pesquisador, além de ser um documento fundamental para quem contrata a pesquisa, pois esclarece acerca do que será pesquisado e apresenta a estimativa de custos, sendo, também, um documento requerido por entidades que financiam pesquisas.

A elaboração de um projeto depende de inúmeros fatores, sendo a natureza do problema o mais importante. Assim, existem projetos bastante simples, como a pesquisa de intenção de voto, e mais complexos, como uma pesquisa que visa a identificar fatores que influenciam determinada situação constatada. Rigorosamente, um projeto só pode ser definitivamente elaborado quando se tem o problema claramente formulado, os objetivos, assim como o plano de coleta e análise dos dados, bem determinados. A estrutura é definida pelo tipo de problema.

É necessário, no entanto, que o projeto esclareça como se processará a pesquisa, ou seja, o seu delineamento, e quais os recursos que devem ser alocados para atingir seus objetivos.

Os elementos habitualmente requeridos num projeto são os seguintes:

- formulação do problema;
- delineamento da pesquisa ou metodologia a ser utilizada;
- cronograma da execução da pesquisa;
- definição dos recursos humanos, materiais e financeiros a serem alocados.

Antes de iniciar a pesquisa propriamente dita, deve ser definido o tipo de estatística a ser calculado (média, mediana, total, proporção, variabilidade etc.). Como passo seguinte, deve ser analisada a caracterização da população a ser pesquisada quanto a sua distribuição espacial, sua estrutura, existência de cadastros atualizados e completos, informações resultantes de pesquisas anteriores e, finalmente, deve-se verificar a existência de mapas geográficos, censitários, sensoriamentos, aerofotogrametrias e plantas topográficas.

Após a realização da investigação, deve ser elaborado, como conclusão, um relatório final em que são apresentados os dados estatísticos relevantes resultantes da investigação.

Ex.: marcas preferidas, unidades consumidas por família, por ano etc.

## 1.4 Fenômenos estatísticos

A Estatística dedica-se ao estudo de fenômenos de massa, que são resultantes do concurso de um grande número de causas, total ou parcialmente desconhecidas, denominadas de fenômenos estatísticos ou variáveis estatísticas.

O fenômeno estatístico é representado por qualquer evento que se pretende analisar, em cujo estudo seja possível a aplicação da técnica estatística.

## 1.5 Dados estatísticos

Os dados estatísticos são a matéria-prima da estatística, portanto são elementos importantes que precisam ser confiáveis e verdadeiros para evitar riscos e erros de interpretação.

São elementos quaisquer que podem ser medidos e que representam a intensidade efetiva de um caráter variável em cada um dos objetos ou pessoas observados. Como exemplo, temos: tamanho dos sapatos, altura das pessoas, renda da população etc.

A unidade do dado estatístico é chamada de elemento, que pode ser um objeto, peça, indivíduo ou conjunto agregado, como um fardo ou uma caixa com produtos.

Exemplo:

Variável: Venda de produtos

Dados: valor das vendas (R\$); quantidade de unidades comercializadas; nº de clientes atendidos (pessoas)

Os dados estatísticos podem ser classificados em primários e secundários.

### 1.5.1 Primários

Os dados primários são obtidos a partir da realização de pesquisas diretas ou pela observação experimental. Exemplos: aplicação de questionários, pesquisas de campo, experimentação agrícola.

### 1.5.2 Secundários

Os dados secundários são elaborados a partir da existência de dados primários, já disponíveis em anuários, bancos de dados, livros, periódicos. Exemplos: Censo Demográfico, Censo Agropecuário etc.

Normalmente é vantajoso usar dados secundários, pois não oneram o orçamento da pesquisa e estão imediatamente disponíveis. Além disso, permite reduzir o tempo necessário para obter os resultados.

Em muitas situações, no entanto, faz sentido coletar os próprios dados, como, por exemplo, a situação óbvia em que o dado coletado não existe. Assim sendo, as vantagens de coletar os próprios dados podem ser as seguintes:

1. é possível coletar os dados de todas as variáveis, com definição exata das medidas necessárias à pesquisa. Em uma pesquisa sobre desemprego, pode-se perguntar ao empresário, por exemplo, quantas máquinas estão parcial ou totalmente ociosas. Um Censo Industrial não incluiria perguntas desse teor;
2. há variáveis que são difíceis de serem quantificadas ou espera-se relutância, reticência ou má fé nas respostas. Coletando os próprios dados, tendo um caso específico em vista, justificam-se todas as precauções cabíveis. Incluem-se escalas múltiplas, digamos, de *status* socioeconômico;
3. em muitas situações o processo de coleta coloca o pesquisador em contato direto e pessoal com o objeto da investigação, o que permite compreender melhor o que está se estudando. No exemplo do desemprego, o empresário pode sugerir, justificar seu comportamento, propor soluções, levantar dúvidas e mesmo questionar a própria validade dos dados em que o pesquisador é levado a acreditar.

Ademais, podemos observar que dados estatísticos são resultantes da variação dos fenômenos estatísticos, e por este motivo, também são conhecidos como variáveis estatísticas. Assim, variável é algo que varia, que muda, ou seja, tudo aquilo que pode assumir diferentes valores ou diferentes aspectos.

## 1.6 Algumas classes específicas de variáveis

As variáveis, cujos dados serão coletados, podem ser classificadas de diferentes formas, de acordo com os seguintes tipos: características demográficas, socioeconômicas e de estilo de vida; atitudes e opiniões; conscientização e conhecimento; motivações; comportamento e intenções.

### 1.6.1 Características demográficas e socioeconômicas

As variáveis com essas características são representadas principalmente por sexo, idade, nível educacional, estado civil, número de filhos, renda, ocupação, bens e produtos possuídos. Com esses dados, o pesquisador pode elaborar tabulações cruzadas que lhe permitirão identificar o relacionamento entre essas variáveis e as opiniões e os comportamentos dos entrevistados em relação ao que está sendo pesquisado. Poderá, também, fazer uma análise ao inverso, e procurar, por meio da tabulação cruzada, por exemplo, identificar e caracterizar o perfil do consumidor típico de determinado produto, conhecido como segmentação de mercado.

Para levantar dados socioeconômicos, uma forma é fazer levantamentos conjugados de itens de posse e da escolaridade dos componentes das famílias. O critério, denominado de Classificação Econômica Brasil, por meio de pontuações dos itens de posse e do grau de instrução, define em classes econômicas as famílias.

### 1.6.2 Estilo de vida

É definido como um modo diferenciado de viver de uma sociedade ou de seus segmentos. A descrição dos entrevistados, segundo seu estilo de vida, envolve levantamento de informações sobre seus valores, hábitos, suas atividades, interesses e opiniões. As atividades estão relacionadas com: trabalho, passatempos, eventos sociais, férias, diversões, membro de clubes,

atividades comunitárias, compras e esporte. Os interesses são ligados a: família, lar, emprego, aspectos comunitários, recreação, moda, alimentos, meios de comunicação e aquisições. As opiniões referem-se: a si mesmo, sobre itens sociais, sobre políticas, sobre negócios, sobre questões econômicas, sobre educação, sobre produtos, sobre futuro e sobre cultura.

A inter-relação dos conceitos de estilo de vida e o que as pessoas estão dispostas a adquirir em termos de produtos e serviços, se suas condições econômicas permitirem, define o conceito do perfil psicográfico. Por isso, as pesquisas quantitativas psicográficas procuram, simultaneamente, levantar e relacionar variáveis demográficas, socioeconômicas, atitudinais, comportamentais e psicológicas, e, a partir de seu tratamento conjunto, estabelecer segmentos de mercado psicográficos.

### **1.6.3 Atitudes**

São importantes em estudos, principalmente de mercado, pois possuem estreito relacionamento com os comportamentos que as precedem. Raramente um consumidor irá adquirir um produto sobre o qual, ou sobre seu fabricante, não tenha uma atitude favorável.

O conhecimento das atitudes dos consumidores permite prever comportamentos para compra; prever comportamentos de pós-compra; prever aceitações/rejeições de produtos e marcas; tomar medidas para procurar mudar atitudes desfavoráveis à empresa e aos seus produtos; avaliar conceitos de novos produtos e avaliar propagandas e promoções de vendas.

A atitude não se manifesta, a não ser pelo comportamento apresentado, ou pelas declarações do próprio entrevistado, que permitem sua medição. A atitude é uma predisposição subliminar da pessoa na determinação de sua reação comportamental em relação a um produto, organização, pessoa, fato ou situação. A atitude possui três componentes: um cognitivo, resultante da vivência e da experiência do entrevistado com o rol de marcas e produtos existentes; um afetivo, resultante dos sentimentos adquiridos em relação ao rol de marcas e produtos existentes, e que está muito relacionado com a imagem que a empresa foi capaz de desenvolver em seu produto ou em sua marca; e um comportamental, relacionado com a rápida reação de comportamento diante de determinados estímulos de marcas e produtos.

### **1.6.4 Conscientização e conhecimento**

Essas variáveis referem-se ao que os entrevistados sabem, ou não, a respeito de determinado produto, serviço, marca, propaganda etc. Inclui-se, também, na mesma análise, a pesquisa do nível de consciência em relação ao produto, suas características, de onde está disponível para ser adquirido, seu preço, da empresa que o produz, de como deve ser usado e com que finalidade e da diferença em relação aos produtos concorrentes.

### **1.6.5 Motivação**

Expressam necessidades, desejos, vontades, impulsos ou qualquer outro estado que impele, conduz, ativa, move, direciona ou canaliza o comportamento de pessoas para atingir determinados objetivos.

As motivações tendem a ser mais estáveis nos indivíduos do que seu comportamento, e conhecer as motivações permite prever melhor o comportamento futuro do que a projeção de comportamentos passados e presentes. Conhecendo as motivações que estão por trás do comportamento das pessoas estar-se-á em condições de melhor entendê-las, de saber como influenciar seus comportamentos futuros ou, ao menos, ter uma ideia consistente de como poderão ser.

## 1.6.6 Comportamento

Estão relacionados ao que as pessoas fizeram ou ao que estão fazendo. O objetivo no estudo do mercado é entender melhor o consumidor e assim poder melhor antecipar ou prever seu comportamento futuro. Os comportamentos de maior interesse em pesquisas de mercado dizem respeito à compra e ao uso de produtos e serviços.

## 1.6.7 Intenções

Referem-se a previsões e antecipações dos comportamentos futuros declarados pelos próprios entrevistados. As intenções podem ser definidas em termos de escalas, como: intenção definitiva de comprar, provável intenção de comprar, indecisão quanto a comprar, provável intenção de não comprar, intenção definitiva de não comprar e sem opinião.

## 1.7 Classificação das variáveis

As variáveis estatísticas também podem ser divididas em duas categorias: variável qualitativa e variável quantitativa.

### 1.7.1 Variável qualitativa

Uma variável é classificada como qualitativa quando seus valores são expressos por atributos, como, por exemplo: sexo, cor da pele, nível cultural, preferências por marcas, estilos de vida, atitudes, opiniões (grau de importância, grau de satisfação), conscientização, conhecimento, motivações, comportamentos e intenções, também é conhecida por variável categórica, classificatória ou nominal e é utilizada para segmentação de pesquisas.

Essas variáveis não podem ser operadas matematicamente. No entanto, se utilizarmos alguma escala de medida, as quais serão apresentadas a seguir, podemos transformá-las em quantitativas. Nesse caso, a escala de medida mais utilizada é a escala Likert, em que cada variável é associada a uma resposta em formato numérico, o que permitirá o cálculo de medidas como a média e o seu desvio-padrão. No entanto, vale lembrar que esses números não têm significado como tal, sendo apenas usados como símbolos. Por exemplo, na aplicação de um questionário, as respostas de sondagem sobre um assunto poderiam ser: não, não sei ou sim, sendo tais respostas associadas a valores: 1. Não; 2. Não sei, e 3. Sim.

### 1.7.2 Variável quantitativa

Por outro lado, as variáveis quantitativas são aquelas cujos valores são expressos em números, como, por exemplo: salários de operários, idade das pessoas, renda familiar, número de acidentes, vendas mensais de uma empresa etc, essa variável, também conhecida como variável cardinal, representa a base da estatística quantitativa ou estatística de variável, pois os resultados das observações serão expressos sempre por valores numéricos.

Essas variáveis podem ser transformadas, no entanto, em qualitativas quando os dados forem agrupados em grupos ou classes nominais ou ordinais. Por exemplo: a renda é uma variável quantitativa, no entanto, pode ser transformada em uma variável qualitativa quando os dados foram classificados por classes de renda: classe A, classe B, classe C, classe D ou classe E. Outro exemplo ocorre quando agrupamos as pessoas por classe de idade: jovem, adulto ou idoso; assim, deixamos de utilizar a idade em números, e passamos a utilizar o atributo por característica de idade.

As variáveis quantitativas são expressas de duas formas: variáveis discretas e variáveis contínuas.

Variáveis discretas podem assumir apenas um valor inteiro, incluído o zero, de maneira que, se elas assumirem outros valores não previstos, elas se destroem e perdem a essência de



valor. Normalmente, a variável discreta resulta de contagem, razão pela qual seus valores são expressos por meio de números inteiros não negativos. Como exemplo de variável discreta citam-se: no de filhos, no de sapatos, no de alunos, no de acidentes, quantidade de livros, número de vendas diárias de uma empresa, número de movimentos de conta corrente em um banco, número de peças com defeito, número de erros de lançamento etc.

Variáveis contínuas podem assumir qualquer valor intermediário entre dois limites de valores inteiros reais. A variável contínua resulta normalmente de mensurações e a escala numérica dos seus possíveis valores corresponde ao conjunto dos números reais. Como exemplos de variável contínua tem-se: idade, peso, comprimento, altura, valor das vendas diárias de uma empresa, consumo mensal de energia elétrica, valor do movimento mensal das contas correntes, rentabilidade diária das ações na bolsa de valores etc.

Para evitar erros, deve-se ter o cuidado com a diferenciação entre os dois tipos de variáveis quantitativas na sua organização em tabelas e cálculos de medidas.

## **1.8 Escalas de medidas de variáveis qualitativas e quantitativas**

As escalas de medida são formas de medição de um conjunto de dados levantados de variáveis estatísticas. Como existem diversos tipos de variáveis, existem também diversas formas de medi-las. Assim, há quatro tipos básicos de escalas de medidas: nominais, ordinais, intervalares e razão.

### **1.8.1 Escala nominal**

É aquela em que as informações servem apenas para nomear, identificar e (ou) categorizar dados sobre pessoas, objetos ou fatos. As escalas nominais são muito utilizadas em pesquisas de mercado para nomear, identificar e classificar variáveis, como: marcas, cores, modelos, sexo, tipo de loja, regiões, uso/não uso, gosta/não gosta, ocupação etc.

### **1.8.2 Escala ordinal**

É aquela em que os dados servem para, além de nomear, identificar e (ou) categorizar, ordenar, segundo um processo de comparação, as pessoas, objetos ou fatos, em relação à determinada característica. Essa escala permite definir a marca preferida pelos consumidores sem, no entanto, estabelecer quanto é essa preferência. As escalas ordinais são usadas para obter o ordenamento das preferências, opiniões, atitudes e percepções de consumidores. Como exemplo, podemos citar: as notas usadas em alguns cursos sob a forma de conceitos (A, B, C, D, E); postos de trabalho (primeiro, segundo, terceiro); classificação de hotéis (cinco estrelas, quatro estrelas, três estrelas); escolaridade (fundamental, média, superior).

### **1.8.3 Escala de intervalo (ou intervalar)**

É aquela em que os dados quantitativos são organizados em intervalos de classe ou faixas de valores que informam a posição e quantificam a situação em termos de pessoas, objetos ou fatos em relação à determinada característica. As escalas de intervalo permitem comparar diferenças entre as medições, mas não permitem concluir quanto à magnitude absoluta das medições, pois não consideram um ponto de partida zero inerente ou natural. São exemplos de variáveis intervalares a temperatura e o tempo, pois essas variáveis não existem quando a informação for igual a zero absoluto (nulo), pois esse valor é arbitrado. No exemplo da temperatura, as escalas de centígrado e Fahrenheit medem um mesmo fenômeno, mas seus pontos zero indicam diferentes graus de temperatura. Assim, 0°C não é uma temperatura inexistente, mas sim representa uma temperatura fria.

Em pesquisas de mercado, as escalas de intervalo são muito utilizadas para medir atitudes, opiniões, conscientização e preferências, em que a preocupação é estabelecer medidas relativas e não absolutas. Nesse caso, variáveis qualitativas são transformadas em quantitativas, conforme já informado anteriormente.

#### 1.8.4 Escalas de razão

Têm as mesmas propriedades das escalas de intervalo, além da propriedade adicional de possuírem o zero absoluto. Em função disso, as medidas tomadas nessas escalas permitem concluir quanto a sua magnitude absoluta, além de informar a posição e quanto as pessoas, objetos ou fatos estão distantes entre si em relação à determinada característica. Essa escala se aplica em pesquisas de mercado para variáveis como: renda, número de consumidores, número de lojas, quantidade de produtos consumidos, número de vezes que o produto é comprado ao mês, volume de vendas, valores de venda, lucros, participação no mercado etc. Nas escalas de razão, é perfeitamente possível e correto comparar as medições absolutas das características e inferir conclusões sobre quanto uma medida é maior ou menor que outras.

### 1.9 Escalas de medidas de transformação de variável qualitativa

Além das medidas escalares básicas, ainda podem ser utilizadas outras escalas de medida em termos de crenças dos pesquisados, principalmente quando são considerados dados relativos a atitudes cuja medição é mais complexa, pois as atitudes existem somente na mente das pessoas, não sendo diretamente observáveis em pesquisas de mercado.

A utilização dessas escalas permite a transformação de uma variável qualitativa em uma variável quantitativa. Assim, foram adaptadas diversas escalas que, no entanto, não são muito diferentes entre si em termos de sua construção e da forma de aplicação.

As principais escalas são as seguintes: escalas nominais, escalas de avaliação, escalas de ordenação, escalas comparativas, escalas de comparação pareada, escalas de diferencial semântico, escalas de Stapel e escalas indiretas.

#### 1.9.1 Escalas nominais

São utilizadas para medir atitudes. Compreendem números que são associados às respostas com o objetivo de identificar as categorias de respostas para facilitar a digitação, o processamento e a análise de dados. Nessas escalas, as respostas das pessoas são classificadas em duas ou mais categorias. Exemplo:

Sua casa está equipada com lava-louça?

1. Não;      2. Não sei;      3. Sim.

#### 1.9.2 Escalas de avaliação

São utilizadas para medir variáveis que envolvem escalas ordinais, de intervalo e razão. Uma escala de avaliação típica procura efetuar a medição do componente afetivo da atitude. As escalas de avaliação são muito utilizadas para medir atitudes em pesquisas de mercado. Existem três formas de construir uma escala de avaliação: gráfica, verbal e itemizada.

A escala de avaliação gráfica compreende a apresentação das opções de respostas às pessoas, desde o extremo mais desfavorável, até o extremo mais favorável, de uma forma visual contínua ou por figuras que representem categorias ordenadas. Exemplo:

Com relação aos serviços prestados na recepção, qual a sua opinião? Assinale a resposta que coincide com a sua opinião:

1. ☹      2. 😊      3. 😊



A escala de avaliação verbal compreende a apresentação das opções de respostas às pessoas, desde o extremo mais favorável até o extremo mais desfavorável, pela identificação e ordenação das categorias por meio de expressões verbais. Exemplo:

Com relação aos serviços prestados na recepção, qual a sua opinião? Assinale a resposta que coincide com a sua opinião.

5. Ótimo;    4. Bom;    3. Regular;    2. Ruim;    1. Péssimo.

Outra escala de respostas que pode ser usada é a seguinte:

5. Muito favorável;    4. Pouco favorável;    3. Indiferente;    2. Pouco desfavorável;    1. Muito desfavorável.

Na construção de escalas verbais não há um número predeterminado de categorias, mas geralmente usa-se, no mínimo, três e, no máximo, sete. Normalmente, os pesquisadores usam escalas de cinco, as quais podem ainda ser reduzidas para três durante o processamento de dados. A escala de cinco categorias permite analisar melhor as nuances pró ou contra do que uma de três. As escalas de sete categorias são utilizadas apenas para medir situações em que se deseja captar pequenas nuances na medição de atitudes.

Ademais, a escala usada normalmente é de tamanho ímpar que tem um ponto neutro e está balanceada em termos de categorias prós e contras. Além das respostas normais, é conveniente oferecer a opção “não sei” ou “não tenho opinião” para não forçar a pessoa pesquisada e evitar, desta forma, respostas enviesadas.

Na escala de avaliação itemizada a pessoa pesquisada é solicitada a escolher uma categoria entre várias, de expressões verbais de opiniões sob determinado tópico que forma uma escala, desde a opinião mais desfavorável até a mais favorável. Esse tipo de escala verbal é o mais difícil de trabalhar, devido à subjetividade das expressões utilizadas. Exemplo:

Com relação ao produto “amaciante de roupa” que a senhora testou, qual das frases está mais de acordo com sua opinião:

1. Eu não acreditava no produto desde o começo e por isso não o usei;
2. Eu não acreditava no produto desde o começo e o seu uso confirmou minhas suspeitas;
3. Eu acreditei no produto no início, mas seu uso me decepcionou;
4. Eu não acreditava no produto, mas seu uso me surpreendeu;
5. Eu acreditei no produto no início, e fiquei muito satisfeita com seu uso;
6. Ainda não formei opinião a esse respeito.

### 1.9.3 Escalas de ordenação

A ordenação das respostas é uma das técnicas usadas em pesquisa de mercado em que é solicitado aos respondentes que ordenem os quesitos de acordo com suas preferências.

Apesar de ser muito usada, essa técnica pode apresentar problemas, pois a pessoa é forçada a ordenar determinado conjunto de objetos, mesmo que não goste de nenhum. Dessa forma, a ordenação pode expressar, ao invés “do que mais gosta para o que gosta menos”, “do que desgosta menos para o que desgosta mais”.

### 1.9.4 Escalas comparativas de avaliação

Têm como procedimento solicitar aos respondentes que expressem suas atitudes mediante comparação com um padrão de referência estabelecido. O padrão de referência pode ser a marca utilizada, a marca líder do mercado, a marca do concorrente, a versão corrente do produto.

As expressões normalmente utilizadas são:

1. Pior que A;    2. Igual a A;    3. Melhor que A;    4. Não sei.

### 1.9.5 Escalas de comparação pareada

Os respondentes são solicitados a comparar dois objetos (produtos, marcas, propagandas etc.) de cada vez, de um conjunto de vários objetos, em relação a suas opiniões sobre os objetos ou sobre vários de seus atributos. O procedimento compreende que os respondentes comparem todos os possíveis pares de objetos do conjunto apresentado. Considerando o número de pares a comparar, recomenda-se a utilização dessa escala para, no máximo, cinco objetos, sendo o ideal de três.

### 1.9.6 Escalas de diferencial semântico

Foram propostas por Osgood, Suci e Tannenbaum (1957) e consistem em que os respondentes avaliem determinado objeto num conjunto de escalas bipolares de sete pontos. Alguns adjetivos bipolares usados em pesquisas de mercado são os seguintes: puro/impuro, forte/fraco, saboroso/sem sabor, sabor diferente/sabor comum, aromático/sem aroma, alta qualidade/sem qualidade, barato/caro, bem torrado/mal torrado, embalagem bonita/embalagem feia, produto moderno/produto antigo. Exemplo:

Com relação à marca de café A, qual é a sua opinião sobre os seguintes atributos:

Puro \_\_\_\_\_ Impuro  
Forte \_\_\_\_\_ Fraco

### 1.9.7 Escalas de Stapel

As escalas de Stapel são uma modificação das escalas de diferencial semântico, sendo os pontos da escala representados por números, sendo a escala de 10 pontos. Exemplo:

Avalie o café da marca A, considerando os atributos de avaliação listados abaixo, segundo a sua opinião.

Atributos	Avaliação
Pureza	-5 -4 -3 -2 -1 +1 +2 +3 +4 +5
Sabor	-5 -4 -3 -2 -1 +1 +2 +3 +4 +5
Aroma	-5 -4 -3 -2 -1 +1 +2 +3 +4 +5

### 1.9.8 Escalas indiretas

Combinam um conjunto de respostas dos entrevistados em relação ao objeto em questão, para determinar qual sua posição na escala de atitude desenvolvida. As escalas indiretas mais importantes, por seu potencial de utilização em pesquisas de mercado, são: as escalas de intervalos aparentemente iguais de Thurstone e as escalas somatórias de Likert.

As escalas de intervalos aparentemente iguais de Thurstone constituem-se num conjunto de declarações, em que cada um possui um valor prédefinido na escala, que são apresentadas aos respondentes para que concordem ou delas discordem. Exemplo:

Assinale nas colunas à direita se você concorda ou discorda de cada uma das afirmações a seguir em relação ao café da marca A.

Afirmações	2. Concordo	1. Discordo
É um café puro		
É um café muito forte		
É muito saboroso		
Seu aroma é delicioso		

A escala Likert foi proposta por Rensis Likert em 1932, à semelhança das escalas de intervalos aparentemente iguais de Thurstone, para medir atitudes a partir de afirmações relacionadas com o objeto pesquisado. Diferentemente da escala de Thurstone, os respondentes são solicitados, não só a concordarem ou discordarem das afirmações, mas também a informarem qual seu grau de concordância/discordância. A cada célula de resposta é atribuído um número que reflete a direção da atitude do respondente em relação a cada afirmação. A pontuação total da atitude de cada respondente é dada pelo somatório das pontuações obtidas para cada afirmação.

Geralmente, o conjunto de termos utilizados para atribuir os graus de favorabilidade/desfavorabilidade é um dos seguintes:

5. Aprovo inteiramente; 4. Aprovo; 3. Indeciso (nem aprovo nem desaprovo); 2. Desaprovo; 1. Desaprovo inteiramente.

5. Concordo inteiramente; 4. Concordo; 3. Indeciso (nem aprovo nem desaprovo); 2. Discordo; 1. Discordo inteiramente

5. Concordo totalmente; 4. Concordo em parte; 3. Indeciso (nem aprovo nem desaprovo); 2. Discordo em parte; 1. Discordo totalmente.

5. Concordo muito; 4. Concordo pouco; 3. Indeciso (nem aprovo nem desaprovo); 2. Discordo pouco; 1. Discordo muito.

Aos vários graus de concordância/discordância são atribuídos números para indicar a direção da atitude do respondente. Geralmente, os números utilizados são 1 a 5, ou -2, -1, 0, 1 e 2. O conjunto de números utilizados não faz diferença em função das conclusões a que se pretende chegar.

Para facilitar a leitura dos resultados, ao usar os números de 1 a 5, é recomendável, uma vez que a média é 3, que as discordâncias sejam representadas pelos números menores e a concordância pelos números maiores. Dessa forma, quando o resultado for acima da média, ocorre a aprovação da expressão e, quando for abaixo da média, ocorre a reprovação.

### **1.10 Arredondamento de dados**

Segundo a Resolução 886/66, do IBGE, o arredondamento de um dado estatístico deve obedecer às seguintes regras: por falta, por excesso e por dígitos seguidos de cinco.

#### **1.10.1 Arredondamento por falta**

Esse arredondamento ocorre quando o primeiro dígito, aquele situado mais à esquerda entre os que irão ser eliminados, for igual ou menor que quatro, não deverá ser alterado o dígito remanescente. Ex.: 72,4 arredonda-se para 72.

#### **1.10.2 Arredondamento por excesso**

O arredondamento por excesso ocorre quando o primeiro dígito, após aquele que será arredondado, for maior ou igual a cinco seguido por dígitos maiores que zero, o dígito remanescente será acrescido de uma unidade. Ex.: 15,504 arredonda-se para 16; 16,56 arredonda-se para 16,6; 17,57 arredonda-se para 18.

#### **1.10.3 Arredondamento de dígitos seguidos de cinco**

Por outro lado, quando o dígito situado mais à esquerda dos que serão eliminados for um cinco seguido somente de zeros, o último dígito remanescente, se for par, não se alterará, e, se for ímpar, será aumentado de uma unidade. Ex.: 215,50 arredonda-se para 16; 216,50 arredonda-se para 216; 216,75 arredonda-se para 216,8; 216,705 arredonda-se para 216,70.

## 1.11 População e amostra

### 1.11.1 População ou Universo Estatístico

Ao conjunto de todos os elementos dá-se o nome de população, ou universo estatístico. A população congrega todas as observações que sejam relevantes para o estudo de uma ou mais características dos indivíduos, os quais podem ser concebidos tanto como seres animados ou inanimados. A população estatística é estudada em termos de observações de características nos indivíduos e não em termos de pessoas ou objetos em si. Assim, por exemplo, as alturas dos cidadãos do Brasil constituem uma população. Poderia haver também uma população correspondente aos pesos desses mesmos cidadãos. Essas características determinadas para toda uma população são denominadas de parâmetros.

### 1.11.2 Amostra

Quando se examina um subconjunto do conjunto maior (população) com características, senão idênticas, bastante assemelhadas ao universo, ele é denominado de amostra, sendo essas características denominadas de estatísticas, e seu valor pode variar de amostra para amostra.

Ao coletar os dados referentes às características de um grupo de objetos ou indivíduos, é muitas vezes impossível de observar toda a população ou universo pelas seguintes razões: custos elevados, tempo necessário e a própria característica do objeto que não permite a sua análise no seu conteúdo completo, como, por exemplo, o exame de sangue. Nesse caso coleta-se uma parte de dados que é chamada de amostra.

Outros exemplos, como medir a inteligência de toda a população brasileira, testar a qualidade de cada saco de café exportado pelo Brasil, aplicar um questionário de *status* socioeconômico a cada escolar do Estado do Rio Grande do Sul, não é possível e racional de ser feito.

Assim, recorre-se ao exame de uma amostra (uma fração da população) e coleta-se uma parte de dados. Se escolhermos a amostra por um processo aleatório (em que a probabilidade de escolha de cada indivíduo é a mesma), a amostra terá a maior probabilidade possível de reproduzir os parâmetros da população. E, ainda, as tabelas de probabilidade permitem avaliar a magnitude provável do erro cometido. Dessa forma, se uma amostra é representativa de uma população, conclusões importantes sobre essa população podem ser inferidas de sua análise.

O tamanho de uma amostra depende do grau de certeza que se quer ter na inferência, da disponibilidade de informantes e de dados, dos custos aprovados e do grau de dispersão dos dados estatísticos. Quanto menor é a homogeneidade da população maior deverá ser a amostra. Outros aspectos relacionadas à obtenção e análise de amostras serão abordadas em capítulo específico do livro.

## 1.12 Levantamento de dados

### 1.12.1 Dados brutos

São os dados que ainda não foram numericamente ordenados. São representados pelos dados levantados como, por exemplo, os pesos de 6 pessoas: 40, 60, 30, 80, 50, 60.

### 1.12.2 Rol

Representa o arranjo dos dados brutos em ordem de grandeza crescente ou decrescente. O exemplo do peso das seis pessoas apresentadas em ordem crescente representa um rol: 30, 40, 50, 60, 60, 80.

### 1.12.3 Amplitude amostral

É a magnitude de variação existente entre os dados levantados. Quando se está analisando dados de variáveis contínuas, pode-se calcular essa magnitude pela diferença entre o valor máximo e o valor mínimo da amostra ou do rol. Assim, o peso das seis pessoas do exemplo apresentado na página anterior, a amplitude amostral é 50:

$$\text{Amplitude amostral: } 80 - 30 = 50$$

Em se tratando de dados de variáveis discretas, o cálculo da simples diferença entre o dado maior e o menor causa um erro em uma unidade, uma vez que, nesse cálculo, os dois extremos estão incluídos na magnitude de variação. Nesse caso o último número que não faz mais parte da amostra está a distância de uma unidade do extremo inferior do rol, diferentemente da variável contínua, em que não há uma distância significativa, coincidindo o valor imediatamente anterior ao extremo inferior desse rol praticamente com esse valor.

Assim, considerando como sendo os valores utilizados no exemplo do peso das seis pessoas oriundos de uma variável discreta, como, por exemplo, número de acidentes, a amplitude amostral será de 51, adicionando-se uma unidade.

$$\text{Amplitude amostral: } 80 - 30 + 1 = 51.$$

### 1.13 Resumo para fixação

Para sintetizar os principais aspectos abordados no capítulo, os seguintes pontos são importantes para fixar como resumo:

- A Estatística é a ciência de coletar, organizar, apresentar, analisar e interpretar dados numéricos ou não (informações) com o objetivo de tomar melhores decisões.
- A estatística se divide em estatística descritiva, a qual preocupa-se com a coleta, organização e apresentação dos dados (amostrais), sem inferir sobre a população; e a estatística indutiva (ou inferência), que se preocupa com a análise e a interpretação dos dados amostrais, com o objetivo de fazer previsões a partir desses.
- Fenômeno estatístico é representado por qualquer evento que se pretende analisar, em cujo estudo seja possível a aplicação da técnica estatística.
- Dados primários são os dados obtidos a partir da realização de pesquisas diretas (questionários) ou pela observação experimental; e dados secundários são os dados elaborados a partir da existência de dados primários, já disponíveis em anuários, bancos de dados, censos, entre outros.
- As variáveis estatísticas se dividem em duas categorias: variável qualitativa, quando seus valores são expressos por atributos; e variável quantitativa, quando seus valores são expressos em números.
- As variáveis quantitativas podem ser classificadas como discretas, quando podem assumir apenas um número inteiro, incluído o zero, ou contínuas, quando podem assumir qualquer valor real, isto é, valor inteiro, valor decimal ou valor fracionário. Podemos notar que os valores decimais e fracionários são valores intermediários entre dois limites de valores inteiros reais.
- As variáveis também podem ser mensuradas por meio de quatro níveis escalares: nominais, quando os dados consistem em nomes, rótulos ou categorias que não podem ser ordenadas; ordinais, quando podem ser arranjados em alguma ordem, mas não é possível medir a diferença entre elas; intervalares, que permitem comparar diferenças entre as medições, mas não permitem concluir quanto à magnitude absoluta das medições, pois não consideram o zero natural; razão, igual a anterior, mas com a propriedade adicional de que há também um ponto inicial de zero natural.
- A escala Likert é a mais utilizada para transformar variáveis qualitativas em variáveis quantitativas.
- O arredondamento de números pode se dar por: a) arredondamento por falta, quando o primeiro dígito, aquele situado mais à esquerda entre os que irão ser eliminados, for igual ou menor que quatro; b) arredondamento por excesso, quando o primeiro dígito, após aquele que será arredondado, for maior ou igual a cinco seguido por dígitos maiores que zero; c) arredondamento de dígitos seguidos do cinco, quando o dígito situado mais à esquerda dos que serão eliminados for um cinco seguido somente de zeros, o último dígito remanescente, se for par, não se alterará, e se for ímpar será aumentado de uma unidade.
- População (ou universo) é o conjunto de unidades com uma ou mais características comuns.
- Amostra é um subconjunto do conjunto maior (população) com características senão idênticas, mas bastante assemelhadas ao universo.
- Dados brutos são os dados que ainda não foram numericamente ordenados.
- Rol representa o arranjo dos dados brutos em ordem de grandeza crescente ou decrescente.
- Amplitude amostral é a magnitude de variação existente entre os dados levantados.

## 1.14 Exercícios propostos

1.1. Classifique cada variável abaixo em qualitativa e quantitativa, e informe a mensuração mais adequada (nominal, ordinal, intervalar ou razão):

- a) As respostas de sondagem em um questionário: sim, não, e indeciso.
- b) As cores dos carros dirigidos por estudantes de uma faculdade.
- c) A produção leiteira de um município do Vale do Taquari.
- d) Faturamento diário em uma loja de confecções.
- e) Nacionalidade dos atletas de uma olimpíada.
- f) Tamanho das camisas em um mostruário.
- g) Temperatura no início de um dia.
- h) Número de reclamações diárias em um SAC.
- i) Classificação da categoria de hotéis (uma, duas, três, quatro ou cinco estrelas).
- j) Marca dos computadores utilizados pelos alunos de uma turma de fundamentos de estatística.

1.2. Diga quais das variáveis abaixo são discretas e quais são contínuas:

- a) Quantidade de livros retirados em uma biblioteca durante o mês de março.
- b) Peso dos recém-nascidos em uma maternidade.
- c) Distância existente entre os municípios de Lajeado e Encantado.
- d) Número de clientes atendidos em uma loja de telefonia em um determinado mês.
- e) Altura dos alunos da turma.
- f) Número de aparelhos eletrônicos com defeitos produzidos mensalmente.
- g) Faturamento diário da loja de vestuário feminino.
- h) Acidentes de trabalho ocorridos durante o último ano.
- i) Número de consultas anuais realizadas por pacientes que apresentam plano de saúde.
- j) Produção agrícola do Vale do Taquari em toneladas.

1.3. Arredonde os números abaixo para números inteiros e para números com um dígito após a vírgula.

- |           |           |
|-----------|-----------|
| a) 68,485 | f) 13,46  |
| b) 322,55 | g) 70,13  |
| c) 29,68  | h) 129,05 |
| d) 203,91 | i) 154,02 |
| e) 217,50 | j) 91,69  |

1.4. Em um levantamento de dados, observou-se que os salários pagos em determinada empresa da cidade no último mês foram os seguintes: R\$ 2.620,00; R\$ 3.735,00; R\$ 1.590,00; R\$ 1.040,00; R\$ 1.750,00; R\$ 910,00; R\$ 985,00; R\$ 800,00; R\$ 1.340,00; R\$ 2.100,00. A partir desses dados, pede-se:

- a) Coloque os dados em forma de rol.
- b) Calcule a amplitude amostral.

1.5. O número mensal de funcionários ausentes sem justificativa em uma determinada empresa da região no último ano foram os seguintes: 12, 14, 4, 1, 4, 5, 8, 3, 2, 0, 2, 9.

- a) Coloque os dados em forma de rol.
- b) Calcule a amplitude amostral.



## 2 TABELAS

Tabelas são estruturas básicas para organização e apresentação de dados, muito utilizadas em textos científicos, pela imprensa, pelas empresas e pelos órgãos governamentais. O principal objetivo da organização de dados é viabilizar a produção de informações rápidas, de forma sintética, das variáveis em estudo, permitindo uma leitura simples e uma interpretação precisa. Os dados podem ser organizados em tabelas de frequência ou em tabelas de séries. Cada qual tem os seus objetivos. A tabela de frequência permite uma análise estática das características intrínsecas da variável, enquanto a tabela de séries permite fazer análises dinâmicas e comparativas.

### 2.1 Tabelas de frequência

Uma vez feita a coleta, a crítica e a apuração dos dados, deve-se efetuar a sua organização, que geralmente é feita por meio de tabelas de frequência, que permitem uma análise melhor dos dados, principalmente quando esses forem em grande número.

Assim, as tabelas servem para ilustrar o conteúdo em desenvolvimento de forma agrupada ou condensada em faixas e descrever dados e informações relevantes para o estudo. No entanto, com o agrupamento, as observações individuais deixam de estar disponíveis, sendo, portanto, importante que se tenha cuidado para que o agrupamento não descaracterize os dados.

A tabela de frequência representa a distribuição das frequências de forma resumida, na qual os dados estão organizados em grupos de classes ou categorias estabelecidos numericamente de forma ordenada, para melhor expressar a situação para sua análise.

Na primeira coluna de uma tabela de frequência aparecem os valores da variável em faixas ou de forma individual, normalmente de forma crescente. Na segunda coluna são anotadas as contagens resultantes do número de vezes que se repete o respectivo valor da variável no rol conforme consta na primeira coluna. Assim, as tabelas de frequência são representações nas quais os valores se apresentam em correspondência com suas repetições, como, por exemplo, a Tabela 2.1, evitando-se assim que eles apareçam mais de uma vez na tabela, como ocorre com o rol.

Tabela 2.1 - Valor da compra dos clientes na promoção da empresa AXD em dezembro de 2013

Valor da compra (R\$)	Número de clientes
300  — 400	5
400  — 500	15
500  — 600	25
600  — 700	5

Fonte: Dados fictícios, apenas para fins ilustrativos.

As tabelas de frequências podem representar tanto valores agrupados com classes quanto valores individuais. Os valores individuais são conhecidos como pontos médios e representam a média de cada classe.

Na construção de tabelas de distribuição de frequência, deve-se atentar para os seguintes pontos: seleção adequada do número de grupos de classe; tipo de intervalo e amplitude apropriados para cada grupo de classe; e estabelecimento de limites (fronteiras) para cada grupo de classe a fim de evitar a sobreposição de dados.

#### 2.1.1 Elementos de uma tabela de frequências

A organização de uma tabela de frequência exige alguns conhecimentos básicos em relação aos seus elementos principais, que são os seguintes: marca de classe, frequência, limites de



classes, amplitude dos intervalos, amplitude total e ponto médio dos intervalos de classe. Na Figura 2.1 podem ser visualizados os elementos de uma tabela de distribuição de frequências.

**Marca de classe:** a primeira coluna é chamada marca de classe de frequência ou simplesmente classe, na qual são apresentados os valores individuais observados de uma variável. Muitas vezes, mesmo com risco de se sacrificar algum detalhe manifestado na ordenação de valores individuais, há vantagem em resumir os dados originais em uma distribuição de frequência, em que os valores não mais aparecerão individualmente, mas agrupados em intervalos de classes.

**Frequência ( $f_i$ ):** a segunda coluna é chamada de frequência e representa o número de vezes que o dado ou fato do intervalo de classe ocorre no rol de observações. O somatório de todas as frequências é igual ao número total de observações.

**Limites de classe:** são os números extremos de cada intervalo de classe. Sendo assim, há dois limites: o inferior e o superior. No exemplo do intervalo de classe 500 |— 600, o limite inferior é 500 e o superior 600.

Figura 2.1 – Visualização dos elementos de uma tabela de distribuição de frequência

Tabela 2 – Retiradas diárias no Banco Alfa no município Beta em Janeiro de 2014.			Marca de classe
Valor da Retirada (R\$)	$\bar{X}$	Número de contas bancárias	Limites de classe: 500  — 600
500  — 600	550	12	Frequência ( $f_i$ )
600  — 700	650	36	Amplitude intervalo classe: 600 - 500 = 100
700  — 800	750	63	Amplitude total ou Range: 1200 - 500 = 700
800  — 900	850	81	Ponto médio das classes: 550, 650, 750, 850, 950, 1050, 11500
900  — 1000	950	77	
1000  — 1100	1050	42	
1100  — 1200	1150	24	
Total		335	

Fonte: Banco Alfa

Fonte: Dados fictícios, apenas para fins ilustrativos.

**Amplitude de um intervalo de classe:** é a medida do intervalo que define a classe. Ela é obtida pela diferença entre os limites superior e inferior dessa classe. Assim, no exemplo no exemplo mostrado na Figura 2.1, para o intervalo de classe 500 |— 600, a amplitude é 600 - 500 = 100.

Para facilitar os cálculos com uma distribuição de frequências, é importante que a amplitude de todos os intervalos de classe seja a mesma.

**Amplitude total ou Range:** é a diferença entre o limite superior do último intervalo de classe (limite superior máximo) e o limite inferior do primeiro intervalo de classe (limite inferior mínimo). Por exemplo, na Figura 2.1 a amplitude total é igual a 700, ou seja, 1200 - 500 = 700

**Ponto médio das classes ( $X_i$ ):** é a média aritmética entre o limite superior e o limite inferior de cada classe. É o número que representa todos os valores observados dentro de um intervalo de classe dado. Ex.:  $(500 + 600)/2 = 1100/2 = 550$

### 2.1.2 Tipos de intervalos de classe

Os intervalos de classes são especialmente recomendados para variáveis quantitativas contínuas. Para variáveis discretas, os procedimentos variam de acordo com a amplitude amostral dos dados.

Variáveis contínuas: existem várias maneiras de apresentação dos intervalos de classe, recomendando-se usar como limites inferiores e superiores, sempre que possível, números inteiros, sendo a amplitude de cada intervalo dada em função do número total de intervalos a serem utilizados na tabela. Os intervalos mais comuns são os seguintes:

**Intervalo 0 —| 10:** compreende todos os valores entre 0 e 10, inclusive o 10 e exclusive o 0.

**Intervalo 0 |— 10:** compreende todos os valores entre 0 e 10, inclusive o 0 e exclusive o 10.

No exemplo da Tabela 2.2, o limite inferior pertence a cada intervalo, enquanto que o superior não faz parte da respectiva marca de classe. Dessa forma, pode-se observar que o consumo mensal mínimo de energia elétrica é de 30 Kwh e que o consumo máximo mensal fica abaixo de 60 Kwh.

Tabela 2.2 - Consumo mensal de energia elétrica na Propriedade “A” localizada em um município do Vale do Taquari no período de maio de 2008 a abril de 2011

Consumo mensal (Kwh)	Meses
30  — 35	5
35  — 40	12
40  — 45	8
45  — 50	6
50  — 55	3
55  — 60	2
Total	36

Fonte: Dados fictícios, apenas para fins ilustrativos.

Para variáveis discretas, os dados normalmente não são organizados em intervalos, mas são listados todos os dados levantados, do menor para o maior, computando as repetições (TABELA 2.3).

Tabela 2.3 - Número de vítimas fatais em rodovias estaduais nos municípios do Vale do Taquari em 2013

Número de vítimas fatais	Número de municípios
0	25
1	5
2	3
3	2
4	0
5	0
6	0
7	0
8	1

Fonte: FEE (2013).

Quando os dados de variáveis quantitativas discretas são em grande número, há, no entanto, dificuldades de organizar a tabela apresentando todas as marcas de classe da variável. Assim, com amplitude amostral dos dados acima de 15, recomenda-se organizar os dados discretos em intervalos apropriados de classe.

Considerando, contudo, que os intervalos de classe são especialmente recomendados para variáveis contínuas, as recomendações de intervalos de classe para variáveis discretas são diferentes e necessitam de adaptação.

Para tanto, deve-se fazer a aplicação da correção de continuidade, uma vez que os dados discretos não apresentam a continuidade, pois são intermitentes. Com a aplicação da correção de continuidade, os limites inferiores e superiores de cada dado são transformados em espaço de ocupação de forma a criar a forma contínua para os dados, como, por exemplo, o número inteiro 5 que é representado pelo intervalo de 4,5 a 5,5, e o número 6, de 5,5 a 6,5. Como o limite de cada dado discreto transformado em contínuo não é mais um inteiro, os limites inferiores e superiores dos intervalos também não serão mais inteiros. Dessa forma, evitam-se erros na centralização do intervalo de classe. Para que o ponto médio de cada intervalo seja representado por um número inteiro, que é a realidade de um dado discreto, recomenda-se usar sempre intervalos com amplitude ímpar, como, por exemplo, 1, 3, 5, 7 etc. Assim, sempre o ponto médio do intervalo será um número inteiro, coerente com o conceito da variável discreta, que não admite números não inteiros. Os intervalos mais comuns para esse tipo de dado são:

**Intervalo 0,5 — 1,5:** compreende todos os valores entre 0,5 e 1,5, exclusive os extremos. Nesse caso, o intervalo da variável discreta inclui apenas o número inteiro 1.

**Intervalo -0,5 — 2,5:** compreende os números inteiros 0, 1, 2.

**Intervalo 2,5 — 5,5:** compreende os números inteiros 3, 4, 5.

Na Tabela 2.3A é apresentada a distribuição do número de vítimas fatais em rodovias estaduais nos municípios do Vale do Taquari em 2013, em que há a organização de cinco intervalos de classe de variável discreta, com amplitude 3 e amplitude total 9, sendo o número mínimo de vítimas fatais por dia igual a 0 e o máximo a 8.

Tabela 2.3A - Número de vítimas fatais em rodovias estaduais nos municípios do Vale do Taquari em 2013

Número de vítimas fatais	Número de municípios
-0,5 — 2,5	33
2,5 — 5,5	2
5,5 — 8,5	1

Fonte: FEE (2013).

Nas apresentações usa-se a opção de colocar na primeira coluna, ao invés dos intervalos, os respectivos pontos médios, que são 1, 4 e 7 (TABELA 2.3B).

Tabela 2.3B - Número de vítimas fatais em rodovias estaduais nos municípios do Vale do Taquari em 2013

Número de vítimas fatais	Número de municípios
1	33
4	2
7	1

Fonte: FEE (2013).

Outra opção é apresentar os números discretos que fazem parte de cada intervalo, cuja apresentação na sequência dos intervalos é a seguinte: 1 a 3; 4 a 6; 7 a 9; 10 a 12; 13 a 15 (TABELA 2.4 B).

Tabela 2.3C - Número de vítimas fatais em rodovias estaduais nos municípios do Vale do Taquari em 2013

Número de vítimas fatais	Número de municípios
0  —  2	33
3  —  5	2
6  —  8	1

Fonte: FEE (2013).

Variáveis qualitativas: dados de variáveis qualitativas não são numéricos, mas quando são obtidos por meio de levantamentos de pesquisas quantitativas, com o uso de questionários estruturados com as categorias padronizadas, também são organizados por meio de tabelas. No lugar dos intervalos de classe são apresentadas as categorias. Recomenda-se ordenar as categorias de acordo com suas frequências, da maior para a menor. Categorias com frequências não significativas normalmente são agrupadas com outras. Quando as categorias se referem a variáveis escalares, como, por exemplo, o grau de instrução, a ordem de apresentação deve seguir a ordem da escala e não a sua frequência. As categorias oriundas da análise de conteúdo de variáveis qualitativas, obtidas por meio de entrevistas em profundidade, com o uso de listas de questões em pesquisas qualitativas, não devem ser tabuladas, e sim listadas.

A Tabela 2.4 apresenta um exemplo de tabulação da variável qualitativa padronizada dos veículos de comunicação por meio dos quais as pessoas buscam se informar.

Tabela 2.4 – Domicílios particulares permanentes por situação de domicílio no Brasil em 2010

Tipo de domicílio	Número de domicílios particulares permanentes	%
Casa	49.837.314	86,94
Casa de vila ou em condomínio	1.018.494	1,78
Apartamento	6.157.162	10,74
Habitação em casa de cômodos, cortiços ou cabeça de porco	296.754	0,52
Oca ou maloca	14.443	0,03
Total	57.314.167	100

Fonte: Dados fictícios, apenas para fins ilustrativos.

### 2.1.3 Determinação do número de intervalos de classe

Para determinar o número de classes, há diversas formas a adotar, podendo ser conforme a preferência do pesquisador, definição de algum critério para fazer comparações com pesquisas já realizadas ou então com base em critérios matemáticos. A rigor, não existe nenhuma exigência em termos de opção a adotar, mas, no entanto, não se recomenda usar mais de 15 intervalos, para não dificultar a análise da variável.

Em não havendo nenhuma recomendação especial, pode-se adotar o cálculo matemático do número de intervalos por meio da regra de **Sturges** que estabelece que o número de classes é igual a:

$$k = 1 + 3,3 \log_{10} n \quad (2.1)$$

onde:

k = número de classes

n = número total de observações

Exemplo:

Se o número de observações for 500:

$$n = 500$$

$$k = 1 + 3,3 \log. 500$$

$$\log 500 = 2,69897$$

$$k = 1 + (3,3 \times 2,69897) = 1 + 8,906601$$

k = 9,906601 e, arredondando,

k = 10, logo, recomenda-se usar 10 intervalos de classe.

Outra alternativa para definição do número de intervalos é usar a função:

$$k = \sqrt{n} \quad (2.2)$$

Assim, usando o mesmo exemplo de 500 observações, temos:

$$k = \sqrt{500} = 22,36, \text{ e arredondando, } k = 22$$

Como o número 22 ultrapassa o limite de intervalos recomendados, escolhe-se o resultado mais adequado que é o da primeira alternativa, ou seja, 10 intervalos.

#### 2.1.4 Normas para apresentação de tabelas

Segundo a NBR 14724/2011, tabelas são elementos demonstrativos de síntese que constituem unidade autônoma, ou seja, devem ter significação própria, em que o dado numérico se destaca como informação central. É importante lembrar que as tabelas são elementos essenciais nas pesquisas quantitativas e devem ter o texto descritivo dos seus dados o mais próximo possível delas. Elas apresentam informações tratadas estatística e numericamente (tabelas de séries estatísticas e tabelas de frequência), segundo as normas de apresentação tabular do Instituto Brasileiro de Geografia e Estatística (IBGE).

Conforme o IBGE (1993), sua característica principal é a apresentação de informações tratadas estatisticamente e, para que atenda as especificações, deve apresentar as partes fundamentais que são a seguir apresentadas (FIGURA 2.2).

**Título:** é representado por um conjunto de informações, as mais completas possíveis, localizado no topo da tabela, que respondem às seguintes perguntas: O quê?, Quando?, Onde?, ou seja, o fato, a época e o lugar a que se referem as informações apresentadas na tabela. O título deve ser claro e conciso. O título é precedido da palavra Tabela e seu número de ordem é apresentado em algarismos arábicos, sempre que um trabalho apresentar duas ou mais tabelas, para identificá-la, permitir a sua fácil localização e a referência no texto.

**Corpo:** é representado pelo conjunto de linhas, colunas e casas ou células que apresentam as informações sobre a variável em estudo, em relação ao corpo da tabela. Os dados devem ser dispostos de modo a possibilitar comparações e ressaltar as relações existentes, destacando o que se pretende demonstrar.

**Coluna indicadora:** é a parte da tabela representada pela primeira coluna que especifica o conteúdo das linhas.

**Linhas:** retas imaginárias que facilitam a leitura, no sentido horizontal, de dados que se inscrevem nos seus cruzamentos com as colunas. Utilizam-se “fios horizontais para separar os títulos das colunas no cabeçalho e fechá-las na parte inferior, evitando-se fios verticais para

separar colunas e fios horizontais para separar as linhas” (ABNT, NBR 14724:2001, p.6). A tabela não deve ser fechada lateralmente.

Casa ou célula: espaço destinado a um só número.

Cabeçalho: parte superior da tabela que especifica o conteúdo das colunas. Os títulos das colunas podem ser digitados verticalmente, quando necessário, para economizar espaços.

Rodapé: serve para colocar a legenda e todas as observações que venham a esclarecer a interpretação dos dados, também conhecidos por fonte, notas e chamadas.

Figura 2.2 – Visualização das partes fundamentais de uma tabela de distribuição de frequência

Tabela 1 – Distribuição da população do Vale do Taquari pro faixa etária em 2010

Faixa etária	Número de habitantes
0 a 20 anos	86.855
21 a 59 anos	190.254
Mais de 60 anos	50.588
Total	327.697

Fonte: IBGE, Censo Demográfico, 2010.

As partes fundamentais da tabela são:

- Título:** o quê? Aonde? Quando?
- Casa ou célula:** espaço destinado a um só número.
- Coluna Indicadora:** indica o conteúdo das colunas.
- Linhas:** as linhas da tabela.
- Rodapé:** serve para colocar a legenda e todas as observações que venham a esclarecer a interpretação dos dados.
- Fonte:** indica a origem dos dados ou a instituição responsável pelo fornecimento ou elaboração dos dados e informações nela contidos.

Fonte: Dados fictícios, apenas para fins ilustrativos.

Fonte: deve-se indicar o autor, ou autores, o título do livro ou revista, a coluna, a página, a data etc., de tal modo que, se alguém quiser comprovar a informação, poderá fazê-lo com facilidade. Também é usada para mostrar cortesia com o autor e a responsabilidade pelos dados. A fonte da tabela indica a origem dos dados ou a instituição responsável pelo fornecimento ou elaboração dos dados e informações nela contidos. A palavra “Fonte” deve ser colocada após o traço inferior da tabela, alinhando-se à margem esquerda da primeira coluna.

A fonte de dados pode ser de diversas origens. Quando os dados se originarem de mais de uma fonte, os nomes ou siglas são separados por vírgula. Em caso de a fonte tratar-se de pessoa física, responsável pelos dados levantados, utiliza-se como fonte o autor. Quando as tabelas são elaboradas com base em fontes que constituem documentos do próprio autor do trabalho (apresentação dos dados, por exemplo), a partir de pesquisa de campo (com o uso de questionários, entrevistas ou observação), podem ser utilizadas como fonte as seguintes expressões: “pesquisa de campo”, “formulários preenchidos”, “entrevistas realizadas”, “questionários aplicados”, “observação direta”, conforme o caso. Quando do uso de tabelas extraídas de outros documentos, é obrigatória a indicação da fonte.

Notas: são informações de natureza geral, destinadas a conceituar ou esclarecer o conteúdo da tabela, ou para indicar a metodologia adotada no levantamento ou redação dos dados.

Chamadas: informações de natureza específica sobre determinada parte da tabela destinada para conceituar ou esclarecer dados.

Observação: é usada para esclarecimentos, tais como o uso de percentuais que devem ser explicados sobre o que se referem e se são sobre o total, de diminuição ou aumento. Quando os números foram arredondados, explicar o sistema de arredondamento. Serve também para algum outro esclarecimento sobre as informações constantes no título ou no cabeçalho que serão apresentadas com asteriscos.

Nenhuma “casa” da tabela deve estar vazia. De acordo com a convenção internacional e Resolução 886 da Fundação IBGE, nas casas ou células, ou quando existem dúvidas quanto à exatidão dos dados, deve-se usar os seguintes critérios:

- **Um traço horizontal ( - )** quando, pela natureza do fenômeno, o dado não existir;
- **Três pontos (...)** quando não se dispuser dos dados;
- **Um ponto de interrogação (?)** quando se tem dúvida quanto à exatidão de determinado valor;
- **Zero ( 0 )** quando o valor é zero, não só quanto à natureza das coisas como quanto ao resultado do inquérito, ou quando o valor é muito pequeno para ser expresso pela unidade utilizada;
- **- ou +** quando os dados anteriores ao símbolo não forem comparáveis aos posteriores;
- **X** quando o dado for omitido para evitar a individualização da informação.

Se os valores são expressos em numerais decimais, precisa-se acrescentar à parte decimal um número correspondente de zeros (0,0; 0,00; 0,000; ...).

A tabela, em relação à sua localização e apresentação gráfica, deve estar inserida o mais próximo possível do trecho do texto a que se refere e ocupar, preferencialmente, uma única página, evitando quebras de página.

A Tabela 2.5 representa um exemplo de apresentação de uma tabela.

Tabela 2.5 – Salários pagos aos operários da empresa XYZ no mês de março de 2012

Número de salários *	Número de operários
0  — 2	40
2  — 4	30
4  — 6	10
6  — 8	15
8  — 10	5
Total	100

Fonte: Dados fictícios, apenas para fins ilustrativos.

(\*) Nota: os valores são em número de salários mínimos.

Se a tabela não couber em uma folha, deve ser continuada na folha seguinte e, neste caso, não é delimitada por traço horizontal na parte inferior, sendo o título e o cabeçalho repetidos na folha seguinte. Quando a tabela continuar na página seguinte, cada página deve ter uma das seguintes indicações: ‘Continua’ para a primeira, ‘Conclusão’ para a última, e ‘Continuação’ para as demais (NBR 14724:2011).

A referência à tabela no texto se faz pela indicação: Tabela ou Tab, acompanhada do número de ordem na forma direta, ou entre parênteses no final da frase (TAB. 6, ou TABELA 6). Para as tabelas apresentadas em anexo, a referência no texto pode seguir o exemplo: TAB. 20, Anexo A. Não se deve usar o plural na abreviatura (TAB. 6 e 7).

A tabela deve estar alinhada, preferencialmente, às margens laterais do texto, ou, quando a sua largura for pequena, deve ser centralizada.

### 2.1.5 Roteiro para a elaboração de uma tabela de frequências

Para a construção de uma tabela de frequências, é conveniente adotar um roteiro que facilite e torne mais operacional o trabalho de quem irá montar a tabela. O roteiro consta dos seguintes passos:

- a) Criar lista de dados brutos que pode ou não ser transformada em rol;
- b) encontrar a amplitude amostral;



c) escolher o número de classes. Para tanto, usar a fórmula de **Sturges**, ou escolher, arbitrariamente, entre um mínimo de 5 e máximo de 20 classes, considerando que quanto maior o número de observações maior deverá ser o número de classes;

d) determinar a amplitude do intervalo de classe pela seguinte fórmula:

$$\text{Amplitude do intervalo de classe} = \frac{\text{Amplitude Amostral}}{\text{Número de Classes}} \quad (2.3)$$

Observação: o resultado deve ser arredondado arbitrariamente para um número mais adequado que facilite os cálculos conforme os tipos de intervalos para dados discretos ou contínuos;

a) determinar os limites inferiores e superiores dos intervalos de classe;

b) construir a tabela de frequência;

c) apresentar a tabela conforme as normas da ABNT.

### 2.1.6 Tipos de frequência para análise de tabelas

As frequências resultantes da organização de tabelas podem ser expressas de quatro formas diferentes: absoluta, relativa, absoluta acumulada e relativa acumulada.

Frequência simples ou absoluta ( $f_i$ ): é o número de vezes que determinado(s) valor(es) ocorre(m) em um extrato, grupo ou intervalo de classe, sendo a soma das frequências de todos os intervalos igual ao total de dados levantados:  $\sum f_i = n$ . A frequência absoluta é a mais utilizada na apresentação de tabelas.

Frequência relativa ( $fr$ ): é a proporção com que um determinado valor ocorre com respeito ao número total de observações, ou seja, é um percentual obtido pelo quociente entre a frequência absoluta de cada marca de classe e o total de elementos, conforme pode se visualizar na fórmula abaixo:

$$fr = \frac{f_i}{n} \quad (2.4)$$

O propósito das frequências relativas é permitir a análise comparativa, pois a sua expressão, quando multiplicada por 100, é um dado em percentual.

A Tabela 2.6 apresenta as diversas frequências relacionadas com os salários recebidos pelos funcionários da empresa “A” de Lajeado.

Tabela 2.6 – Pontos médios e frequências dos salários pagos pela empresa “A” localizada em Lajeado em fevereiro de 2012

Salários (R\$ por hora)	$X_i$	$f_i$	$F_i$	$fr$ (%)	$Fr$ (%)
10  — 20	15	5	5	12,5	12,5
20  — 30	25	8	13	20,0	32,5
30  — 40	35	11	24	27,5	60,0
40  — 50	45	9	33	22,5	82,5
50  — 60	55	4	37	10,0	92,5
60  — 70	65	3	40	7,5	100
Total		40	-----	100	-----

Fonte: Dados fictícios, apenas para fins ilustrativos.

Frequência absoluta acumulada ( $F_i$ ): representa a soma do número de elementos que existem desde o primeiro intervalo até um intervalo de classe determinado, ou seja, o total de frequências absolutas de todos os intervalos igual ou inferiores a de determinada classe, conforme abaixo:

$$F_i(k) = f_1 + f_2 + \dots + f_k, \text{ sendo } k \text{ o intervalo de classe determinado.}$$



Frequência relativa acumulada ( $F_r$ ): representa a soma das frequências relativas que existem desde o primeiro intervalo de classe até um determinado intervalo K desejado, conforme ilustrado na fórmula:  $F_r(k) = fr_1 + fr_2 + \dots + fr_k$ .

O seu valor também pode ser representado como uma divisão da frequência absoluta acumulada de um determinado intervalo de classe pela frequência total ou número de observações, segundo a fórmula a seguir:

$$Fr = \frac{Fi}{n} \quad (2.5)$$

Para não poluir as tabelas, usa-se como regra apresentar as frequências simples, absolutas e relativas, enquanto as acumuladas, absolutas e relativas são usadas para análise e interpretação dos dados (TABELA 2.7).

Tabela 2.7 - Salários recebidos pelos funcionários da empresa “A” de Lajeado em fevereiro de 2012

Salários (R\$ por hora)	Nº de pessoas	Percentual de pessoas
10  — 20	5	12,5
20  — 30	8	20,0
30  — 40	11	27,5
40  — 50	9	22,5
50  — 60	4	10,0
60  — 70	3	7,5
Total	40	100

Fonte: Dados fictícios, apenas para fins ilustrativos.

## 2.2 Tabelas de séries estatísticas

Além das tabelas de frequências que normalmente procuram organizar os dados de forma a reduzir o seu número, mantendo apenas aqueles que têm frequências de repetência, as tabelas de séries estatísticas procuram organizar os dados mantendo todos os que foram levantados. Enquanto as tabelas de frequência procuram organizar os dados para uma análise estática da variável em termos de características, as tabelas de séries procuram mostrar a dinamicidade dos dados, com possibilidades de comparação, em termos de tempo, local e espécie.

Dessa forma, uma série estatística é toda a tabela que apresenta a distribuição de um conjunto de dados estatísticos em função da época, do local ou da espécie. É importante salientar que os três fatores fazem parte das informações do título de cada tabela, ou seja, o quê? (espécie), onde? (local) e quando? (época). A tabela apresenta na primeira coluna a discriminação do fator de análise: época, local ou espécie e, na segunda coluna os respectivos dados. Pode-se usar uma terceira coluna com as informações das quantidades em relação ao total.

Observa-se, portanto, a existência de três elementos ou fatores que podem ser objeto de uma série estatística, que são: o tempo, o espaço e a espécie.

Conforme o elemento que faz variar a série, pode-se classificá-la, em três tipos diferentes: histórica, geográfica e específica.

### 2.2.1 Séries históricas, cronológicas, temporais ou marchas

Este tipo de tabela descreve os valores da variável de determinado local, alterando em termos de intervalos de tempo. As alternativas para relacionar os dados com o tempo são as mais diversas. Como regra geral pode-se citar as seguintes: por século, por ano, por semestre, por trimestre, por bimestre, por mês, por quinzena e por dia da semana.

Como exemplo de uma série histórica, a Tabela 2.8 apresenta dados discriminados por ano:

Tabela 2.8 - Produção anual de leite de vaca no Brasil em mil litros no período de 2004/2010

Ano	Litros de leite (mil)
2004	23.474.694
2005	24.620.859
2006	25.398.219
2007	26.137.266
2008	27.585.346
2009	29.085.495
2010	30.715.460

Fonte: IBGE, Produção da Pecuária Municipal, 2004/2010.

### 2.2.2 Séries geográficas, espaciais, territoriais ou de localização

Esta forma de apresentação de tabela descreve os valores da variável, em determinado instante, segundo regiões de onde se originaram os dados. Da mesma forma como a discriminação por tempo, existem também diversas alternativas para apresentar os dados por elementos geográficos. As formas mais comuns são por continente, por bloco econômico, por país, por região, por município, por bairro e por filial.

Na Tabela 2.9, a produção é apresentada pelas diferentes regiões, representando uma série geográfica:

Tabela 2.9 - Produção de ovos de galinha por região geográfica do Brasil - 2010

Região	Ovos de galinha (mil dúzias)
Norte	116.843
Nordeste	497.904
Sudeste	1.433.915
Sul	845.751
Centro-Oeste	352.307

Fonte: IBGE, Produção da Pecuária Municipal, 2010.

### 2.2.3 Séries específicas ou categóricas

As tabelas de séries específicas ou categóricas descrevem os valores da variável, em determinado tempo e local, discriminados segundo suas especificações ou categorias. No mundo real existem muitas formas de categorias e espécies a serem apresentadas. Dentro das empresas existem os tipos de matéria-prima, os diversos produtos, as diversas marcas, as rubricas de contas de despesas e receitas, além de outras. Fora das empresas existem outras formas, entre as quais a apresentada na Tabela 2.10, como exemplo, das espécies que fazem parte do rebanho brasileiro.

Tabela 2.10 – Composição do rebanho brasileiro por espécie - 2010

<b>Espécie</b>	<b>Quantidade (1.000 cabeças)</b>
Galos, frangas, frangos e pintos	1.028.151.477
Galinhas	210.761.060
Bovino	209.541.109
Suíno	38.956.758
Ovino	17.380.581
Codornas	12.992.269
Caprino	9.312.784
Equino	5.514.253
Muar	1.277.419
Bubalino	1.184.511
Asinino	1.001.587
Coelhos	226.359

Fonte: IBGE, Produção da Pecuária Municipal, 2010.

## 2.2.4 Tabelas de séries complexas

Além da apresentação simples, as tabelas de séries também podem ser organizadas de forma complexa. Esta forma permite ampliar as possibilidades de análise dinâmica das variáveis, sendo dessa forma muito utilizada no dia a dia, nas empresas e pelos órgãos de pesquisa. Essa sistemática apresenta um fator na vertical, como primeira coluna, e outro fator na horizontal, formando as diferentes colunas. Como forma prática, recomenda-se colocar o fator com mais alternativas na primeira coluna da tabela.

As formas mais comuns de organização de tabelas de séries complexas são as seguintes: com dados duplamente históricos, com dados históricos e geográficos, com dados históricos e específicos, e com dados geográficos e específicos. Os dados apresentados na primeira coluna identificam a classificação da série, em termos históricos, geográficos ou específicos.

A Tabela 2.11 representa um exemplo de uma série específica e histórica, pois apresenta na primeira coluna a especificação dos produtos e nas demais colunas a descrição dos anos em que ocorreram as importações.

Tabela 2.11 – Exportação brasileira por fator agregado - 2010 e 2011, em US\$ milhões FOB (*Free on Board*)

<b>Fator agregado</b>	<b>Anos</b>		<b>Variação em %</b>	<b>Participação em %</b>	
	<b>2010</b>	<b>2011</b>		<b>2010</b>	<b>2011</b>
Básicos	90.005	122.457	36,06	45,51	48,71
Semimanufaturados	28.207	36.026	27,72	14,26	14,33
Manufaturados	79.563	92.929	16,80	40,23	36,96

Fonte: MDIC/SECEX/DEPLA 2010/2011.

Já a Tabela 2.12 apresenta um exemplo de uma série geográfica e histórica, pois na primeira coluna constam os municípios de origem e nas demais colunas, abertas por anos, são apresentados os dados das exportações.

Tabela 2.12 - Principais municípios exportadores do Rio Grande do Sul nos anos de 2010 e 2011, em US\$ FOB

Município	Anos		Variação em % 2010/2011	Participação em % no total do RS	
	2010	2011		2010	2011
Rio Grande	1.690.234.123	2.529.139.144	49,63	11,52	13,40
Porto Alegre	1.434.159.666	1.972.619.310	37,55	9,77	10,45
Triunfo	1.506.633.706	1.953.730.128	29,68	10,27	10,35
Santa Cruz do Sul	955.160.397	1.283.057.131	34,33	6,51	6,80
Caxias do Sul	893.042.755	1.066.563.454	19,43	6,09	5,65
Canoas	956.335.114	1.045.305.195	9,30	6,52	5,54
Total do Estado RS	14.671.867.036	18.867.768.081	28,60	-	-

Fonte: MDIC/SECEX/DEPLA, 2010/2011.

Como exemplo de uma série duplamente histórica é apresentada a Tabela 2.13, que traz os dados da Balança Comercial brasileira, separados por exportação e importação, com os respectivos anos e os meses do ano. Como os meses são em número de 12 e os anos de importação e exportação em número de 4, colocam-se nas linhas os meses e nas colunas os anos.

Tabela 2.13 - Balança Comercial brasileira de janeiro a dezembro de 2010 e 2011, em US\$ milhões FOB

Meses	Exportação		Variação % 2010/2011	Importação		Variação % 2010/2011
	2010	2011		2010	2011	
Janeiro	11.305	15.214	34,58	11.486	14.816	28,99
Fevereiro	12.197	16.732	37,18	11.808	15.538	31,59
Março	15.727	19.286	22,63	15.055	17.734	17,79
Abril	15.161	20.173	33,06	13.879	18.312	31,94
Maio	17.703	23.209	31,10	14.252	19.684	38,11
Junho	17.094	23.689	38,58	14.827	19.259	29,89
Julho	17.673	22.252	25,91	16.329	19.114	17,06
Agosto	19.236	26.159	35,99	16.846	22.280	32,26
Setembro	18.833	23.285	23,64	17.762	20.213	13,80
Outubro	18.380	22.140	20,46	16.554	19.785	19,52
Novembro	17.687	21.773	23,10	17.396	21.195	21,84
Dezembro	20.918	22.127	5,78	15.574	18.312	17,58
Total	201.914	256.039	26,81	181.768	226.242	24,47

Fonte: MDIC, SISCOMEX, 2010/2011.

## 2.3 Resumo para fixação

Para sintetizar os principais aspectos abordados no capítulo, sugerem-se os seguintes pontos como resumo para fixação:

- As tabelas são formas de organização de dados para permitir a análise e interpretação de variáveis.
- Existem três tipos de tabelas: tabelas de frequência com intervalos para variáveis quantitativas, tabelas de frequências de atributos ou categorias indicadas para variáveis qualitativas padronizadas e tabelas de séries estatísticas.
- As tabelas de frequências e de categorias apresentam os dados na primeira coluna e na segunda, a respectiva frequência, enquanto que as tabelas de séries apresentam na primeira coluna o fator da série (tempo, local ou espécie) e na segunda coluna, os dados.
- As tabelas de frequência com intervalos de classe são esboçadas para variáveis quantitativas contínuas. Assim, para usá-las com variáveis quantitativas discretas, devem-se ajustar os limites inferiores e superiores dos intervalos com a correção de continuidade de Yates.
- Para facilitar na análise dos dados, recomenda-se usar a mesma amplitude para todos os intervalos de classe da tabela de frequência.
- Para tabelas de frequência com intervalos de classe de variáveis discretas, recomenda-se usar uma amplitude com valor ímpar, de forma que o ponto médio sempre seja um valor real inteiro.
- O número de intervalos de classe pode ser definido por regras, como a regra de Sturges ou a regra da raiz do número de dados, mas pode ser definido também de forma arbitrária, recomendando-se um mínimo de 4 e um máximo de 15.
- Nas tabelas de frequência normalmente são apresentadas as frequências simples absolutas e relativas em percentual. As demais frequências, absoluta e relativa acumuladas, são usadas na análise e incluídas no texto descritivo.
- Nas tabelas não são usadas linhas horizontais nem verticais na separação dos dados do corpo das tabelas. Usam-se apenas linhas horizontais para destacar os cabeçalhos e os totais e linhas horizontais externas à tabela: superior e inferior.
- As tabelas devem ser enumeradas conforme a ordem de apresentação no texto e o título sempre deve responder as questões: o quê?, onde?, e quando?.
- No rodapé sempre deve constar a fonte dos dados.
- Todas as tabelas apresentadas em um texto devem ser referenciadas com o seu respectivo número e devem ser precedidas e seguidas de texto interpretativo. A tabela é uma ilustração que deve ser comentada no texto, não sendo por si só uma informação textual.
- Deve-se evitar a quebra de tabelas. Nesse sentido, deve-se apresentar a tabela de forma a mantê-la unida, ajustando-se os parágrafos anteriores e posteriores.

## 2.4 Exercícios propostos

2.1. Considerando as alternativas abaixo, em relação à questão a seguir apresentada, assinale as respostas com V se verdadeiras e com F se forem falsas.

Uma distribuição de frequência apresenta características entre as quais se têm as seguintes:

- ( ) O ponto médio de uma classe é resultante da soma do seu limite inferior com a metade de sua amplitude.
- ( ) A frequência absoluta representa o número de dados que pertencem a um intervalo de classes.
- ( ) As amplitudes dos intervalos não devem ser iguais.
- ( ) O número de intervalos deve ser igual à amplitude deles.

2.2. Assinale a resposta correta da questão a seguir.

Segundo as normas da ABNT, uma tabela deve apresentar informações entre as quais se podem citar:

- a) Título com as informações: assunto, época e o local de onde provem os dados.
- b) Na primeira coluna, as frequências e na segunda os respectivos intervalos com os dados.
- c) Nem sempre é possível informar a fonte.
- d) No rodapé sempre deve constar o assunto e a fonte.
- e) A tabela deve apresentar linhas horizontais e verticais

2.3 O consumo mensal de energia elétrica, em kWh, em uma residência do Vale do Taquari, no período de outubro de 2013 a setembro de 2014 foi o seguinte: 210, 242, 338, 274, 217, 213, 267, 493, 406, 322, 268, 214. Organize estes dados em uma tabela de frequência (com frequências absolutas e relativas) e apresente-a segundo as normas da ABNT.

2.4 A empresa SAPATO NOVO, preocupada com o desperdício que ocorre na fábrica, passou a fazer uma avaliação diária da produção. No levantamento feito durante determinado mês de 2014, o número diário de sapatos produzidos com defeito foi: 10, 20, 5, 0, 40, 10, 15, 18, 22, 2, 8, 15, 1, 25, 33, 34, 12, 3, 2, 16, 14, 23, 13, 0, 11, 27, 17, 2, 12, 35. Organize estes dados em uma tabela de frequência (com frequências absolutas e relativas) e apresente-a segundo as normas da ABNT, usando a regra de Sturges, com o ajuste dos intervalos de classe recomendados para variáveis discretas.

2.5 O gasto mensal com alimentos consumidos na residência de uma família residente em Lajeado, ao longo de 2013 foi o seguinte: R\$ 1.240,00; R\$ 1.164,00; R\$ 896,00; R\$ 1.459,00; 1.043,00; R\$ 929,00; R\$ 860,00; R\$ 958,00; R\$ 1.399,00; R\$ 1.456,00; R\$ 1.274,00; R\$ 914,00. Organize estes dados em uma tabela de frequência (com frequências absolutas e relativas) e apresente-a segundo as normas da ABNT, com 5 intervalos.

2.6. Os salários pagos aos funcionários na empresa JKL Marketing, no mês de março de 2004, foram os seguintes: 950, 1150, 1950, 2180, 2660, 2400, 724, 724, 1820, 1610, 1355, 1710, 990, 3245, 2890, 1050.

- a) Organize os dados em uma tabela de frequência com frequências absolutas e relativas e apresente-a segundo as normas da ABNT, com amplitude de cada intervalo igual a 500.

2.7. Uma empresa exportadora de aço registrou os seguintes números de encomendas (em toneladas) dos seus clientes durante um semestre: 4; 3; 5; 4; 5; 6; 5; 6; 4; 3; 5; 5; 6; 6; 3; 3; 3; 5; 5; 5.

- a) Organize os dados em uma tabela de frequência com frequências absolutas e relativas e apresente-a segundo as normas da ABNT, usando a regra de Sturges.
- b) Determine também as frequências acumuladas: absoluta e relativa e os pontos médios de cada intervalo.

2.8. Os números de empréstimos para habitação, de 50 bancos comerciais, concedidos durante um mês são os seguintes: 2; 4; 2; 32; 9; 9; 2; 6; 3; 1; 14; 9; 16; 7; 8; 19; 6; 4; 4; 2; 4; 18; 0; 6; 13; 7; 2; 8; 0; 1; 14; 1; 2; 2; 18; 8; 24; 1; 8; 5; 8; 26; 3; 12; 23; 1; 8; 5;

- a) Organize os dados em uma tabela de frequência com frequências absolutas e relativas e apresente-a segundo as normas da ABNT, usando a regra de Sturges, com ajuste dos intervalos de classe recomendados para variáveis discretas.
- b) Determine também as frequências acumuladas: absoluta e relativa e os pontos médios de cada intervalo.



### 3 GRÁFICOS

Correspondem à representação dos dados sob diferentes formas, a fim de permitir uma visão rápida e global do fato estudado.

A estrutura de um gráfico baseia-se em uma função que expressa a relação existente entre a variável **dependente (Y)** variando em função da variável **independente (X)**:

$$Y = f(X)$$

Para representar a função, o mais usado é o gráfico cartesiano.

A variável dependente é colocada no eixo das ordenadas e a variável independente é colocada no eixo das abscissas.

Assim, a estrutura do gráfico apresenta as seguintes características:



#### 3.1 Representação gráfica de uma distribuição de frequência

Uma distribuição de frequência pode ser representada graficamente pelo **histograma**, pelo **polígono de frequência** e pela **curva de frequência**.

Construímos qualquer um dos gráficos mencionados utilizando o primeiro quadrante do sistema de eixos coordenados cartesianos e ortogonais. Na linha horizontal colocamos os valores da variável e na linha vertical as frequências. Por essa razão, o ponto médio é dado pelo símbolo  $X_i$ .

A apresentação gráfica de uma distribuição de frequências permite transformar dados em informação. Pode-se perceber, ao analisar um gráfico, como é a distribuição dos dados, se há uma uniformidade, se há concentração e as tendências.

Deve-se ter o cuidado na montagem para que seja respeitada a distribuição dos dados dentro de uma escala, a fim de evitar a manipulação na ampliação gráfica.

##### 3.1.1 Histograma

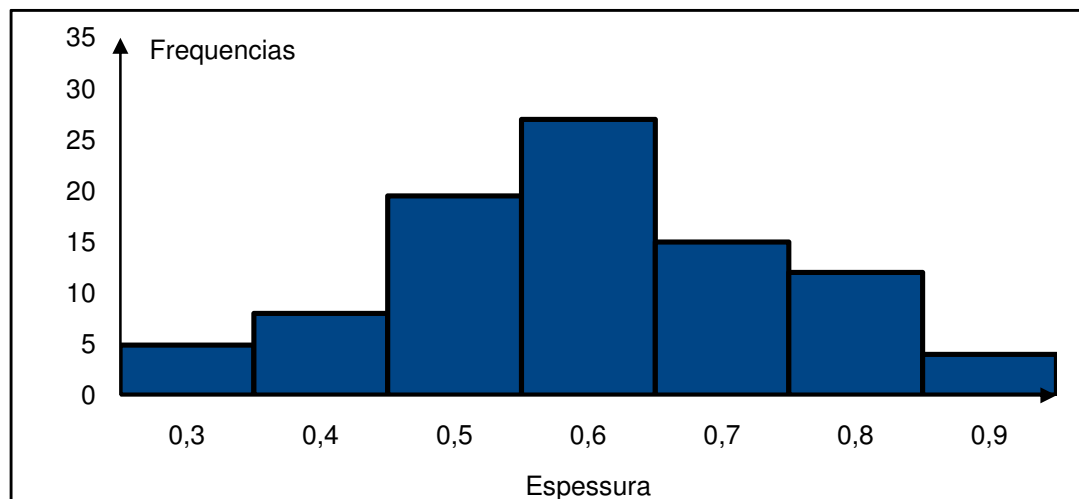
É um diagrama de colunas representadas graficamente por meio de retângulos adjacentes em que a base colocada no eixo das abscissas corresponde aos intervalos das classes e a altura é dada pela frequência absoluta ou relativa das classes.

Os intervalos devem ser de igual magnitude, podendo cada intervalo ser considerado com a dimensão 1, sendo a altura de cada coluna diretamente proporcional à frequência.

O centro do retângulo representa o ponto médio e as larguras são iguais às amplitudes dos intervalos de classe.

As áreas são proporcionais às frequências de cada intervalo de classe.

Gráfico 3.1 - Histograma da espessura de peças produzidas pela empresa Alfa no mês de maio de 2014



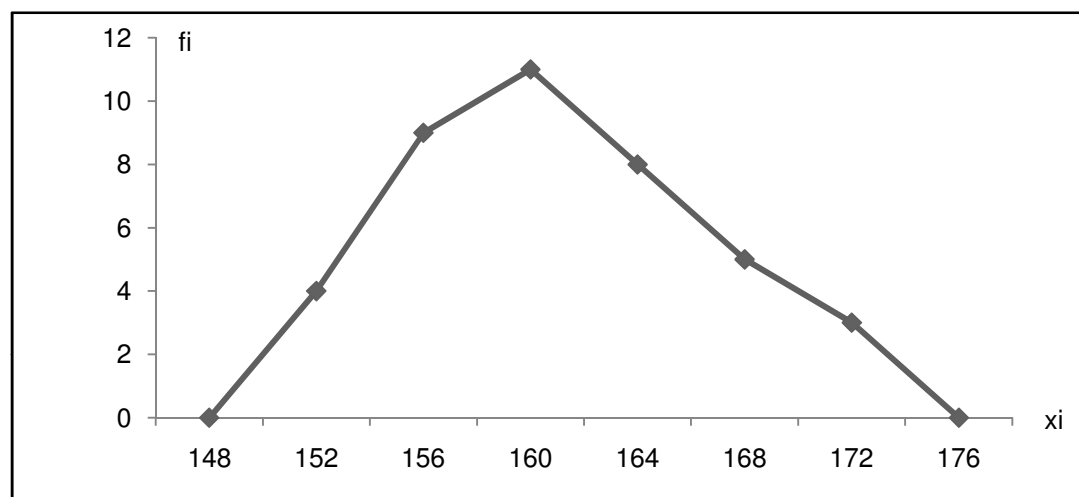
Fonte: Dados fictícios, apenas para fins ilustrativos.

### 3.1.2 Polígono de frequência

É a representação gráfica de uma distribuição de frequências por meio de um polígono, em que os pontos são obtidos por perpendiculares traçadas a partir dos pontos médios das classes, e de altura proporcional à frequência de cada uma das classes.

Para realmente obtermos um polígono (linha fechada), devemos completar a figura, ligando os extremos da linha obtida aos pontos médios, da classe anterior à primeira e da posterior à última, da distribuição. Os intervalos fictícios assim criados têm frequência zero.

Gráfico 3.2 - Polígono de Frequências da altura dos alunos do 6º de uma escola municipal de Lajeado em 2014



Fonte: Dados fictícios, apenas para fins ilustrativos.

### 3.1.3 Curva de frequência

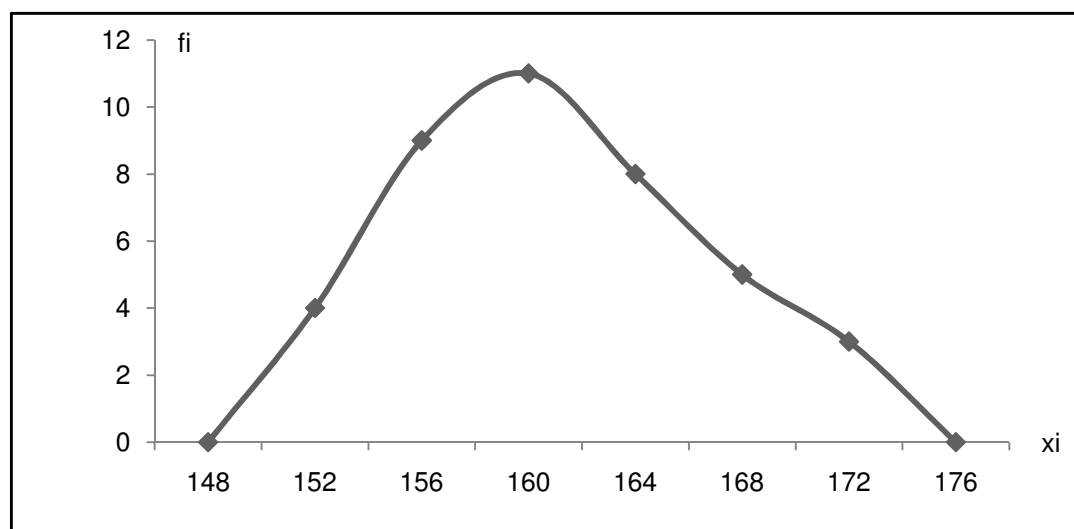
É a representação gráfica de uma distribuição por meio de uma curva regular que procura acompanhar o mesmo traçado e a mesma área subentendida pelo polígono de frequência. É o “polimento” da linha fechada do polígono.

Como, em geral, os dados coletados pertencem a uma amostra extraída de uma população, pode-se imaginar que, com amostras mais amplas, a amplitude dos intervalos se torne cada vez menor, o que permite concluir que a linha poligonal tende a se transformar numa curva.

Assim sendo, o polígono de frequência dá a imagem real do fenômeno estudado e a curva de frequência dá a imagem tendencial.

Esse procedimento dá uma ideia do que se encontraria se estivesse sendo usado um grande número de dados.

Gráfico 3.3 - Curva de Frequências da altura dos alunos do 6º de uma escola municipal de Lajeado em 2014



Fonte: Dados fictícios, apenas para fins ilustrativos.

As áreas de todos os gráficos de distribuição de frequências (**histograma, polígono de frequência e curva de frequência**) são iguais.

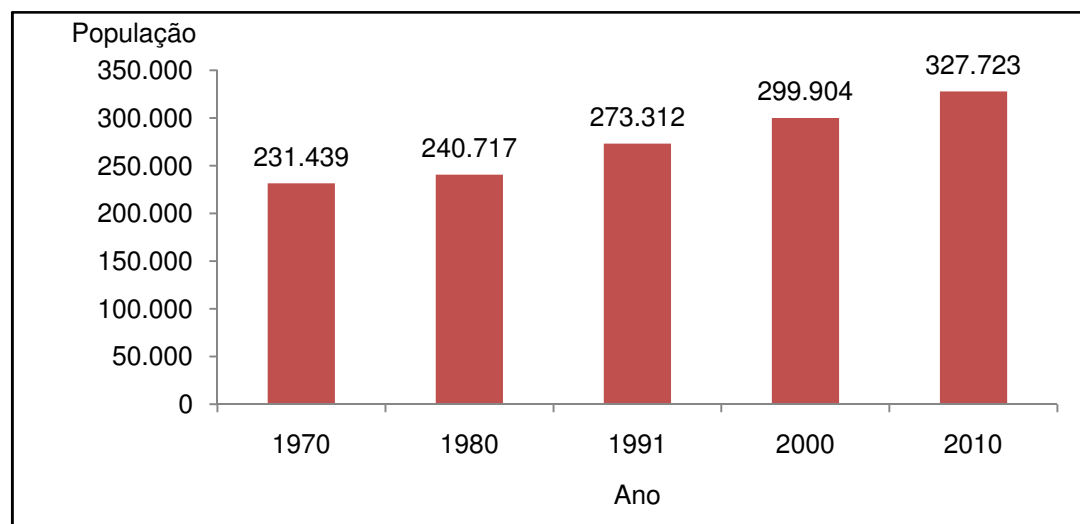
### 3.2 Gráficos cartesianos básicos

Além dos gráficos das distribuições de frequência, existem as mais variadas formas de apresentar graficamente dados de séries e dados de variáveis qualitativas. No entanto, considerando as características gerais dos gráficos, pode-se classificá-los em quatro tipos diferentes: gráfico de colunas, diagrama de Pareto, gráfico de linhas e diagrama de dispersão.

#### 3.2.1 Gráfico de colunas

O gráfico de colunas é representado por um conjunto de retângulos expostos em forma de colunas, que podem ser apresentados, justapostos ou não, ou ainda em três dimensões. Nas ordenadas são colocadas as frequências absolutas ou o percentual de repetições e na abcissa os fatores analisados, que podem ser de variáveis discretas, qualitativas ou até intervalos de classe, ou os seus pontos médios, de uma tabela de frequência. O gráfico de colunas é semelhante ao histograma, sem a exigência de as colunas serem justapostas. Deve-se ter um cuidado especial quanto à intersecção da ordenada com a abcissa, para que ela aconteça no valor zero, a fim de evitar interpretações errôneas na comparação das alturas das diversas colunas. O Gráfico 3.4 apresenta um exemplo de gráfico de colunas de uma série histórica.

Gráfico 3.4 – População da região do Vale do Taquari no período entre 1970/ 2010

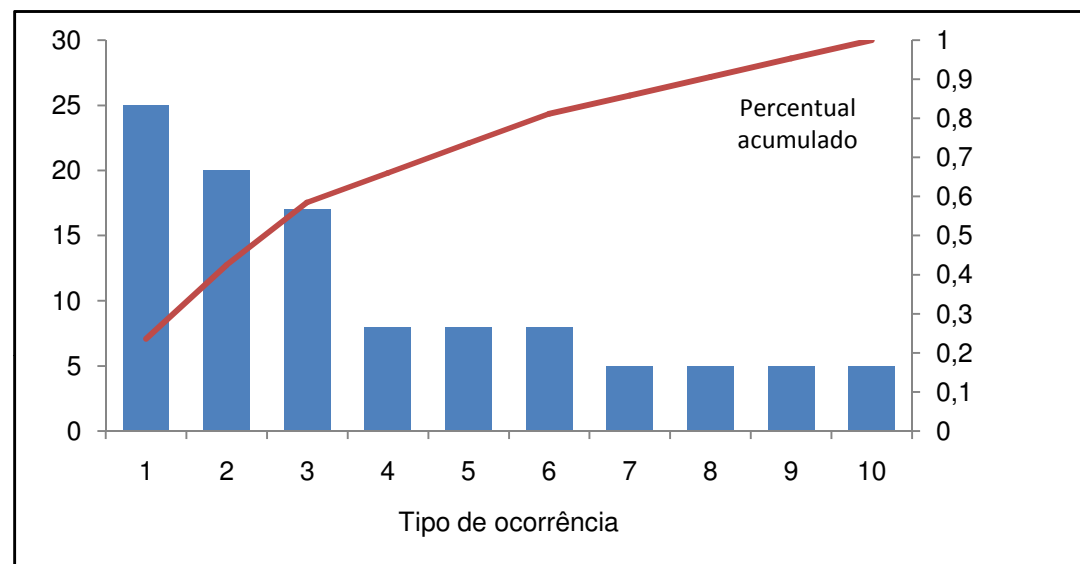


Fonte: IBGE, Censo Demográfico, 1970/2010.

### 3.2.2 Diagrama de Pareto

Um diagrama de Pareto é um tipo específico de gráfico de colunas ou histograma, ordenado por frequência de ocorrência, que mostra quantos defeitos foram gerados por tipo ou categoria de causa identificada. A técnica de Pareto é usada principalmente para identificar e avaliar não conformidades.

Gráfico 3.5 – Número de peças com defeitos encontradas em uma linha de produção em uma empresa localizada em Estrela/RS no mês de janeiro de 2014



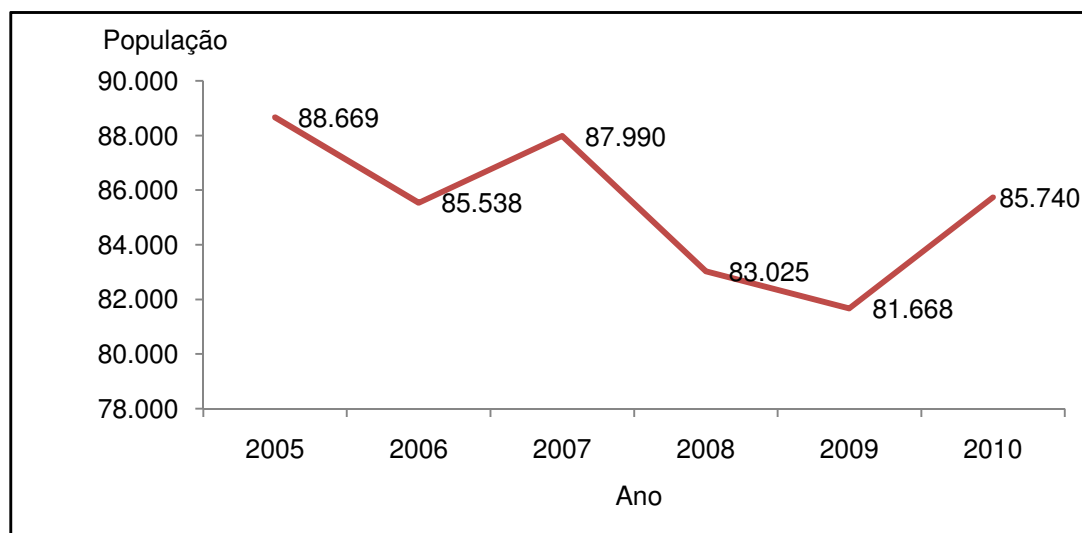
Fonte: Dados fictícios, apenas para fins ilustrativos.

Nos diagramas de Pareto, a classificação é usada para orientar as ações corretivas. A equipe do projeto deve tomar ações para resolver primeiramente os problemas que estão causando o maior número de defeitos. Os diagramas de Pareto estão conceitualmente relacionados à Lei de Pareto, que afirma que um número relativamente pequeno de causas normalmente produzirá a grande maioria dos problemas ou defeitos. Isso geralmente é chamado de princípio 80/20, em que 80% dos problemas se devem a 20% das causas.

### 3.2.3 Gráfico de linhas

Um gráfico de linhas apresenta um conjunto de pontos conectados por uma única linha. As linhas de gráfico são usadas para representar os dados que ocorrem em um período de tempo contínuo. O Gráfico 3.6 apresenta um exemplo de gráfico de uma série histórica.

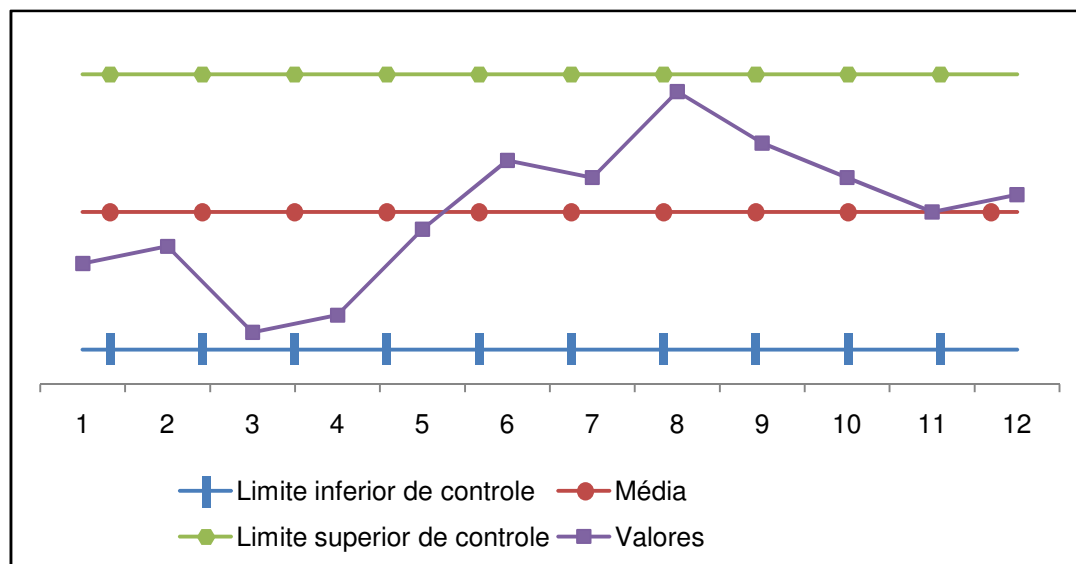
Gráfico 3.6 – Número de habitantes da região do Vale do Taquari residente na zona rural no período de 2005/ 2010



Fonte: IBGE, Censo Demográfico, 2005/2010.

O gráfico de linhas também pode ser representado pela Carta de Controle da análise dos resultados de qualidade, avaliados ao longo do tempo conforme exemplo ilustrado no Gráfico 3.7.

Gráfico 3.7 – Amostra do controle do processo produtivo da Empresa Delta

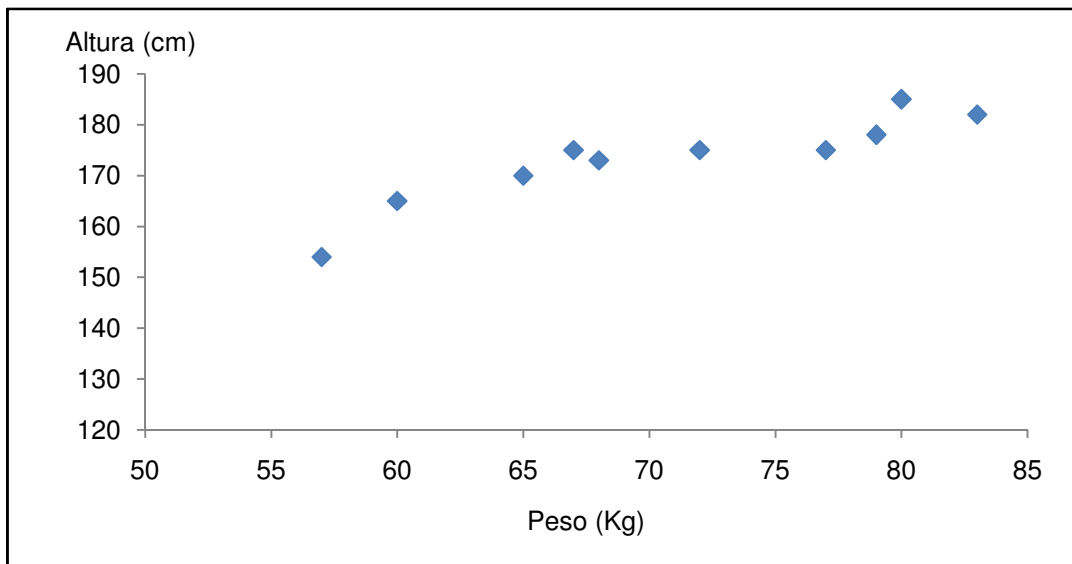


Fonte: Dados fictícios, apenas para fins ilustrativos.

### 3.2.4 Diagrama de dispersão

O diagrama de dispersão apresenta os dados levantados no espaço do sistema de eixos cartesianos, permitindo a visualização de tendências e da correlação entre as variáveis (dependente em relação à independente), mostrando como é o relacionamento (de forma linear, ou não linear ou de nenhuma relação). Em se tratando de séries estatísticas, ter-se-ão como variável dependente os dados da série e como variável independente o fator analisado (histórico, geográfico ou específico). O Gráfico 3.8 apresenta um exemplo de gráfico de dispersão.

Gráfico 3.8 – Relação entre a altura e o peso dos alunos da turma do 7º ano de uma escola municipal de Arroio do Meio/RS em 2014



Fonte: Dados fictícios, apenas para fins ilustrativos.

### 3.3 Normas de apresentação de ilustrações

A exemplo da apresentação de tabelas, as ilustrações têm por objetivo possibilitar a transmissão de dados e informações de modo mais atraente. De acordo com a NBR 14724/2011, ilustrações (gráficos, quadros, plantas, retratos, organogramas, fluxogramas, esquemas, figuras, mapas, desenhos, fotografias e outros) são elementos demonstrativos de síntese que constituem unidade autônoma e explicam ou complementam visualmente o texto, sendo, assim, diretamente relacionadas com o conteúdo da informação, pois, do contrário, não contribuirão para a análise. Assim, para facilitar, as ilustrações devem ser inseridas o mais próximo possível do trecho do texto a que se referem.

Ainda de acordo com a NBR 14724 de 2011, qualquer que seja seu tipo (desenhos, esquemas, fluxogramas, fotografias, gráficos, mapas, organogramas, plantas, quadros, retratos ou outros), sua identificação aparece na parte superior, precedida da palavra designativa (por exemplo: Gráfico...), seguida de seu número de ordem de ocorrência no texto, em algarismos arábicos e do respectivo título escrito em fonte tamanho 12. Na parte inferior, após a ilustração, indicar a fonte utilizada (elemento obrigatório, mesmo que seja produção do próprio autor), legenda, notas e outras informações necessárias à compreensão (se houver), escritas em fonte tamanho 10. Nos dois eixos do gráfico devem ser colocados os cabeçalhos, conforme forem as informações da variável dependente e independente. O corpo é representado pelo próprio gráfico.

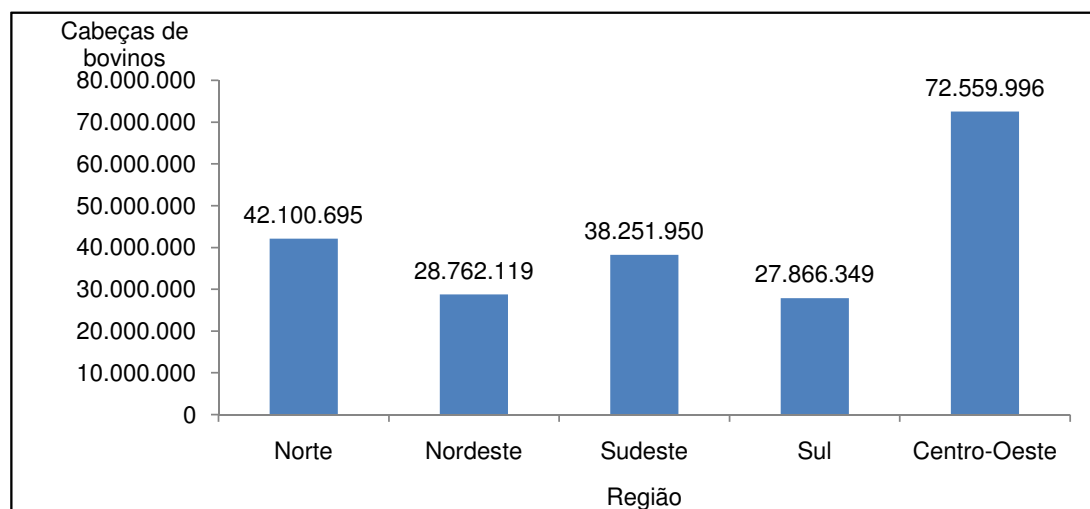
A seguir tem-se a apresentação de tabelas com os seus respectivos gráficos, conforme as séries geográfica (Tabela 3.1 e Gráfico 3.9), histórica (Tabela 3.2 e Gráfico 3.10) e específica (Tabela 3.3 e Gráfico 3.11).

Tabela 3.1 – Rebanho de bovinos nas Grandes Regiões do Brasil - 2010

Região	Cabeças de bovinos
Norte	42.100.695
Nordeste	28.762.119
Sudeste	38.251.950
Sul	27.866.349
Centro-Oeste	72.559.996
Brasil	209.541.109

Fonte: IBGE, Produção da Pecuária Municipal, 2010.

Gráfico 3.9 – Rebanho de bovinos nas Grandes Regiões do Brasil - 2010



Fonte: IBGE, Produção da Pecuária Municipal, 2010.

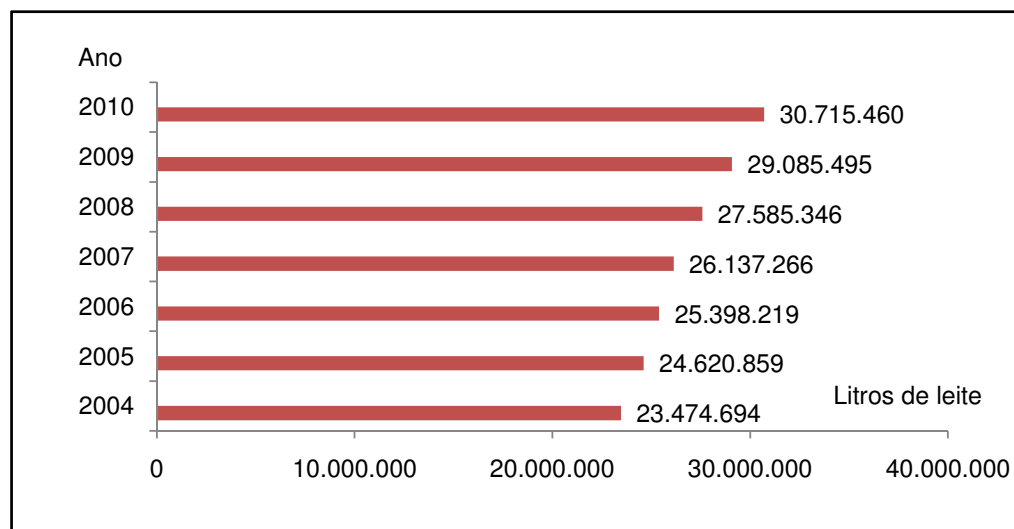
Tabela 3.2 – Produção anual de leite de vaca no Brasil (em mil litros) no período - 2004/2010

Ano	Litros de leite (mil litros)
2004	23.474.694
2005	24.620.859
2006	25.398.219
2007	26.137.266
2008	27.585.346
2009	29.085.495
2010	30.715.460

Fonte: IBGE, Produção da Pecuária Municipal, 2004/2010.



Gráfico 3.10 – Produção anual de leite de vaca no Brasil (em mil litros) no período de 2004 a 2010



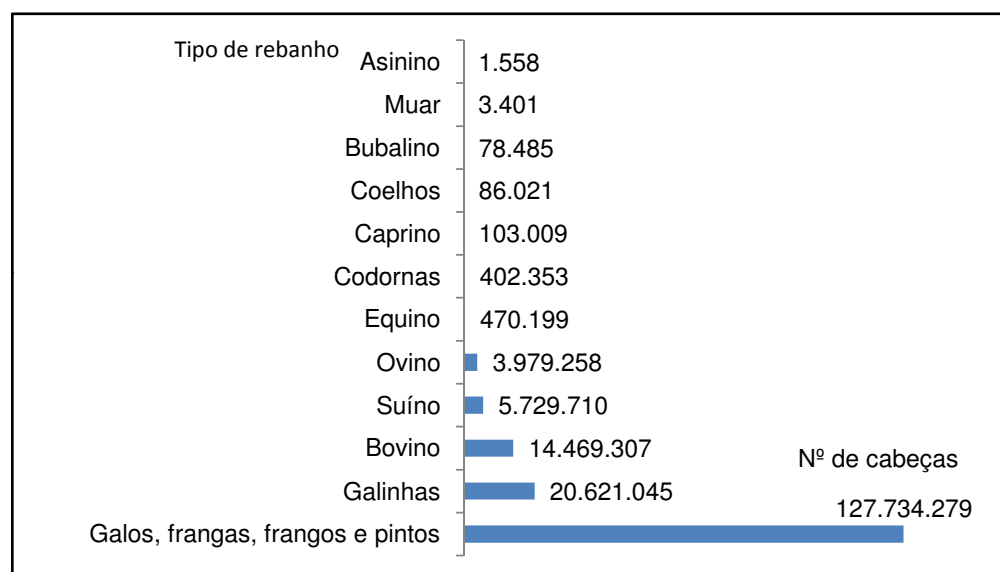
Fonte: IBGE, Pesquisa da Pecuária Municipal, 2004/2010.

Tabela 3.3 – Composição do rebanho por espécie no estado do Rio Grande do Sul - 2010

Tipo de rebanho	Número de cabeças
Galos, frangas, frangos e pintos	127.734.279
Galinhas	20.621.045
Bovino	14.469.307
Suíno	5.729.710
Ovino	3.979.258
Equino	470.199
Codornas	402.353
Caprino	103.009
Coelhos	86.021
Bubalino	78.485
Muar	3.401
Asinino	1.558

Fonte: IBGE, Produção da Pecuária Municipal, 2010.

Gráfico 3.11 - Composição do rebanho por espécie no estado do Rio Grande do Sul - 2010

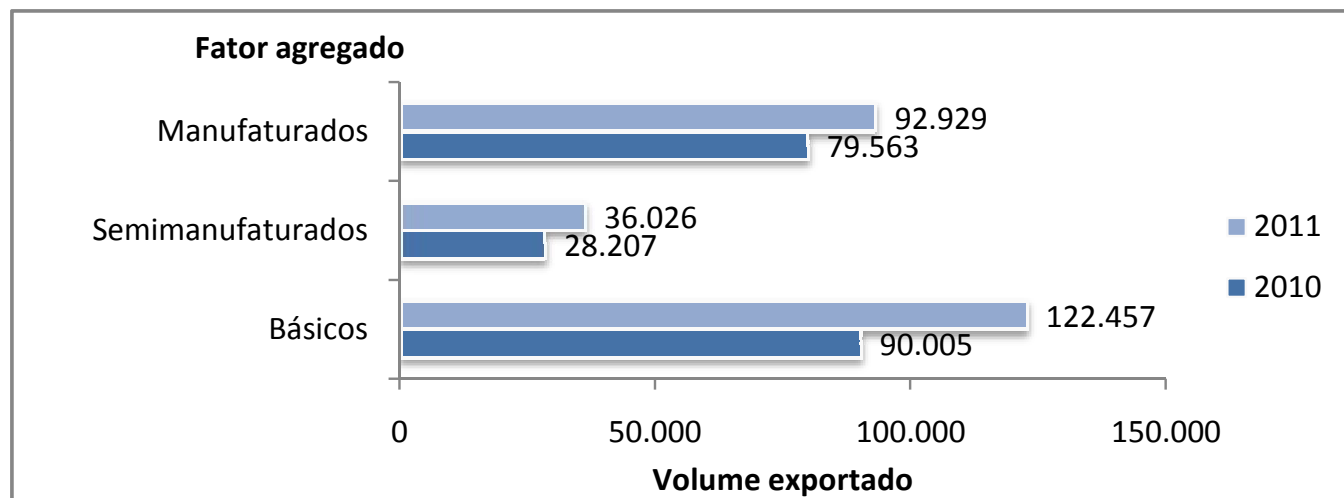


Fonte: IBGE, Produção da Pecuária Municipal, 2010.

### 3.4 Gráficos de séries complexas

Segue um exemplo de uma série específica e histórica:

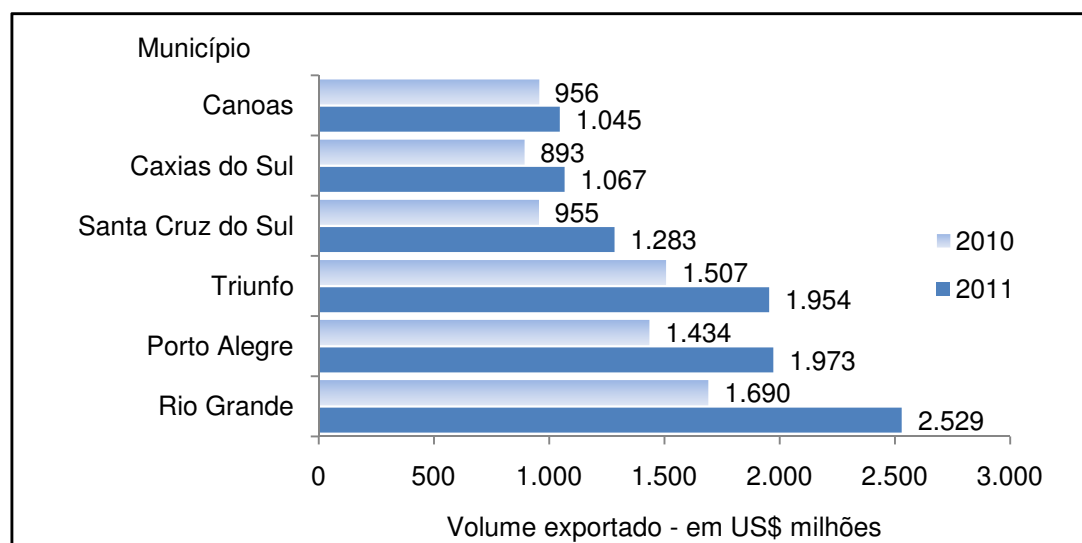
Gráfico 3.12 - Exportação brasileira por fator agregado – 2010/2011, em US\$ milhões FOB



Fonte: MDIC/ SECEX, 2010/2011.

Segue um exemplo de uma série geográfica e histórica:

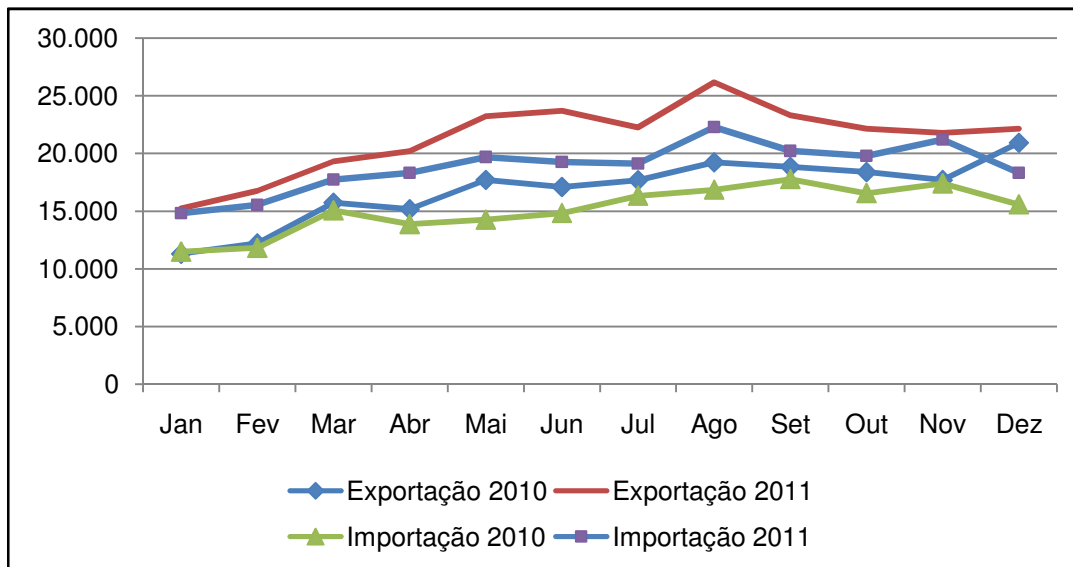
Gráfico 3.13 – Principais municípios exportadores do RS nos anos de 2010/2011, em US\$ milhões FOB



Fonte: MDIC/SECEX, 2010/2011.

Exemplo de uma série duplamente histórica:

Gráfico 3.14 - Balança Comercial de janeiro a dezembro de 2010/ 2011, em US\$ milhões FOB



Fonte: MDIC/SECEX, 2010/2011.

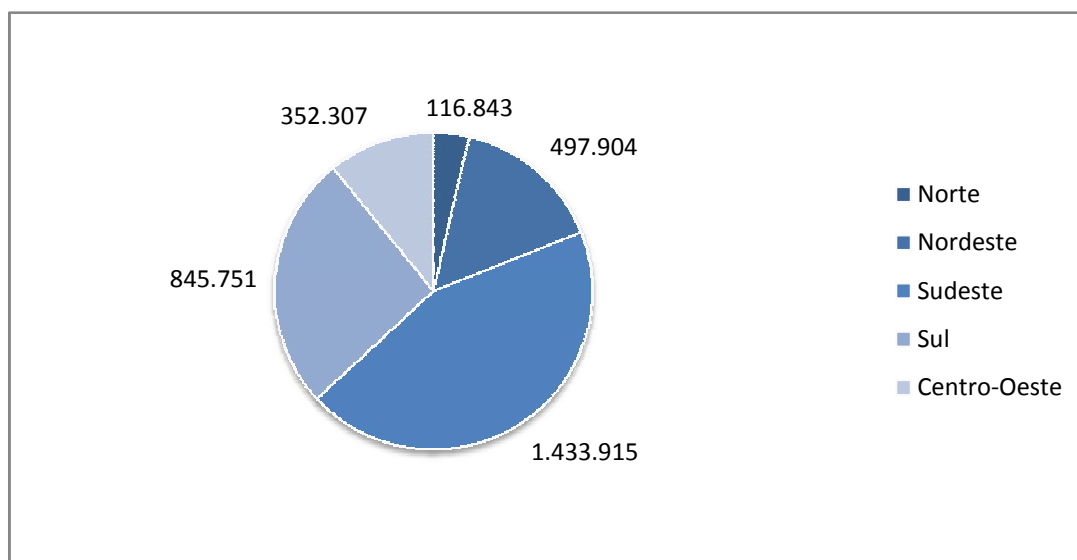
### 3.5 Outros gráficos

Além dos gráficos citados, existem outras formas utilizadas para apresentar informações estatísticas. Entre eles existem os que serão apresentados a seguir.

#### 3.5.1 Gráficos de setores ou de “pizza”

São representados por meio de setores em um círculo.

Gráfico 3.15 – Produção de ovos de galinha no Brasil (em mil dúzias) - 2010



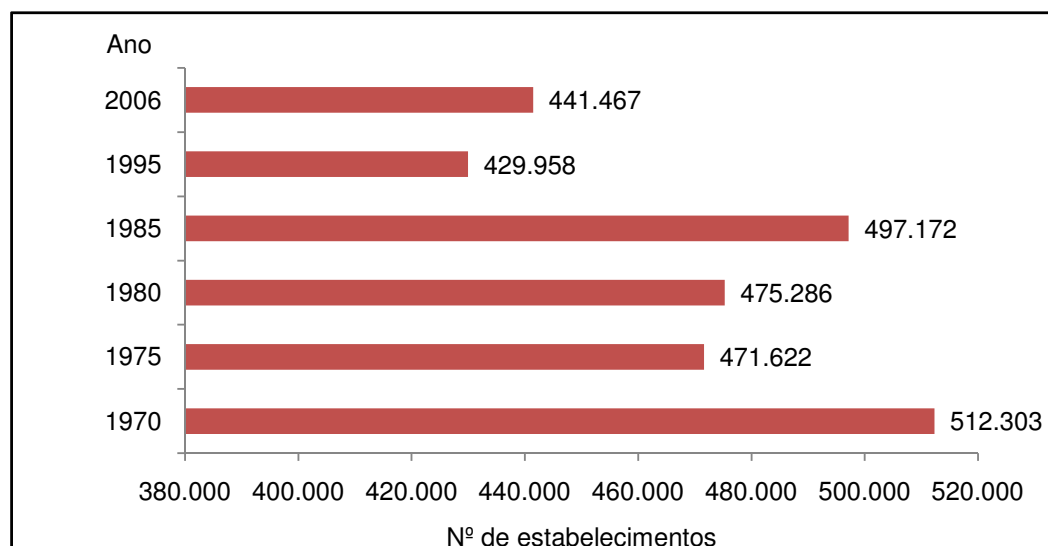
Fonte: IBGE, Produção da Pecuária Municipal, 2010.

#### 3.5.2 Gráfico de barras

São representações de retângulos ou cilindros em posição horizontal. Representam um gráfico rotativo, em que as ordenadas estão no lugar das abcissas e vice-versa.

Uma alternativa do gráfico de colunas é o gráfico de barras. São representações de retângulos em posição horizontal. Representam um gráfico rotativo, em que as ordenadas estão no lugar das abcissas e vice-versa.

Gráfico 3.16 - Número de estabelecimentos agropecuários no RS – 1970 a 2006



Fonte: IBGE, Censos Agropecuários, 1970/2006.

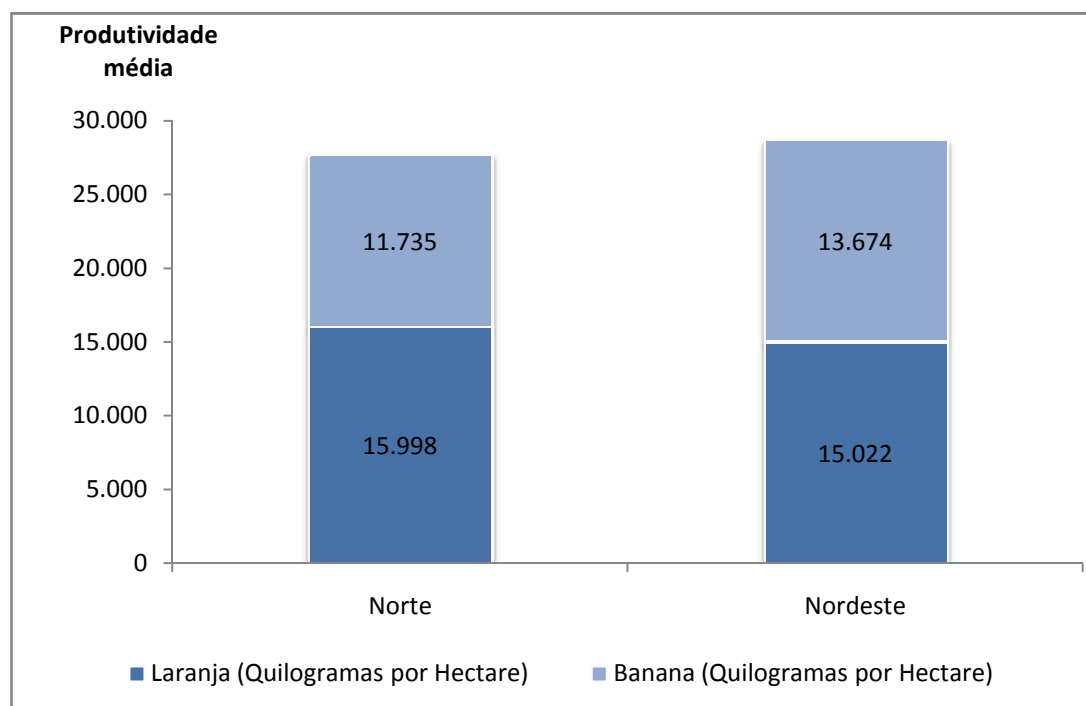
### 3.5.3 Gráficos de colunas transformadas:

Além do gráfico tradicional de colunas que pode ser apresentado em duas ou três dimensões por meio de programas como planilhas eletrônicas e *softwares* de estatística, existem outros tipos de gráficos de colunas, entre os quais os mais comuns são os seguintes:

- colunas superpostas;
- colunas contrapostas ou opostas;
- colunas compostas.

Exemplo de gráfico de colunas superpostas:

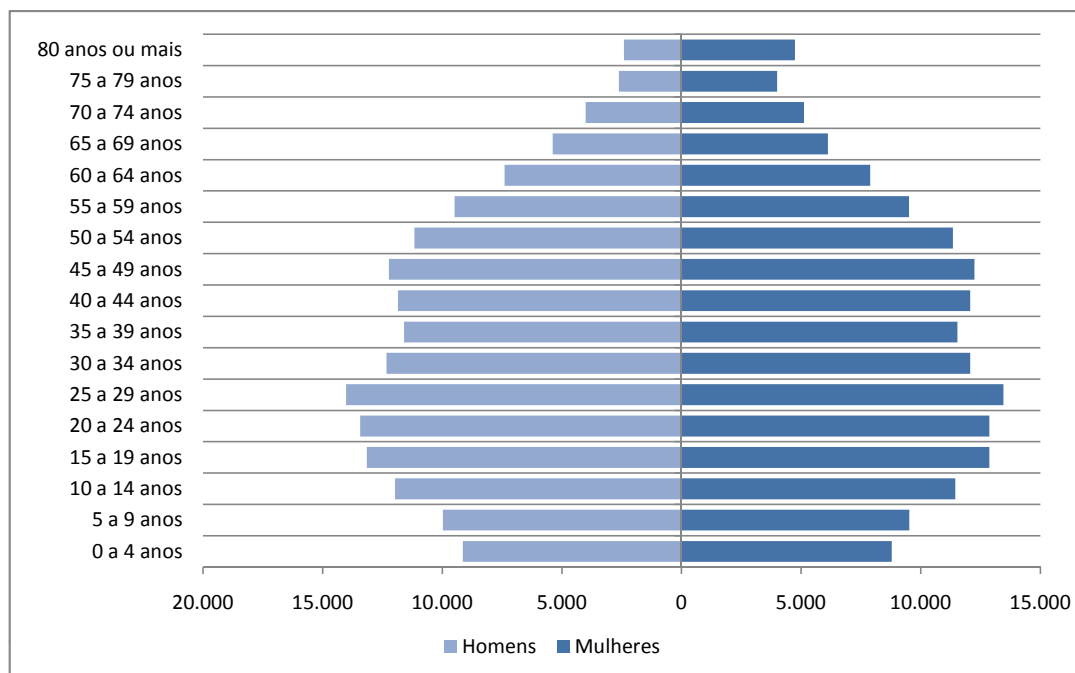
Gráfico 3.17 – Produtividade média de safra de laranja e banana das regiões Norte e Nordeste – 2011



Fonte: IBGE, Levantamento Sistemático da Produção Agrícola, 2011.

Exemplo de gráfico de colunas contrapostas ou opostas:

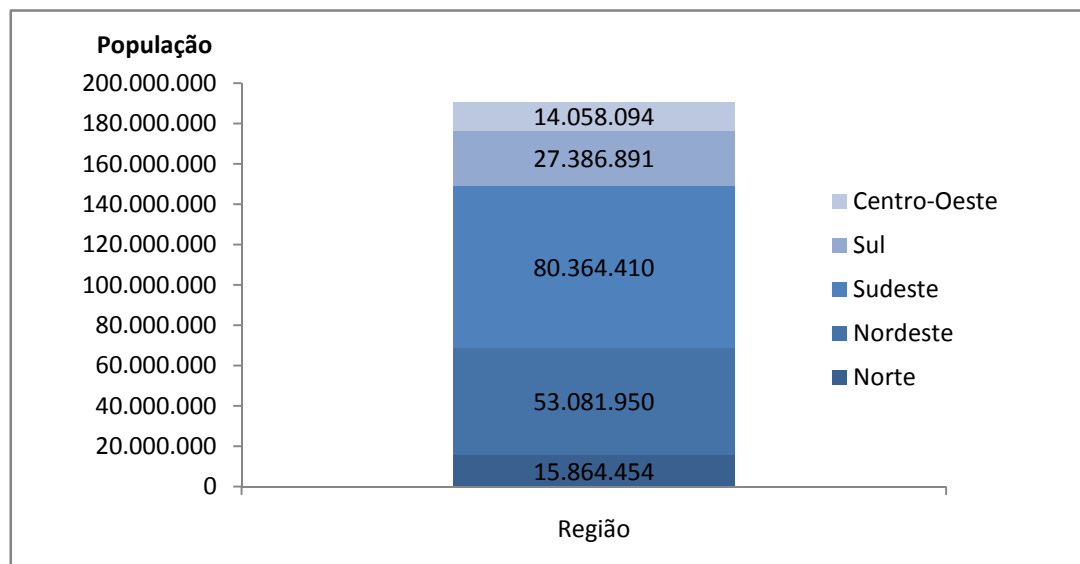
Gráfico 3.18 – Pirâmide etária do Vale do Taquari - 2010



Fonte: IBGE, Censo Demográfico, 2010.

Exemplo de gráfico de colunas compostas:

Gráfico 3.19 – População residente, por situação de domicílio, segundo as Grandes Regiões do Brasil – 2010

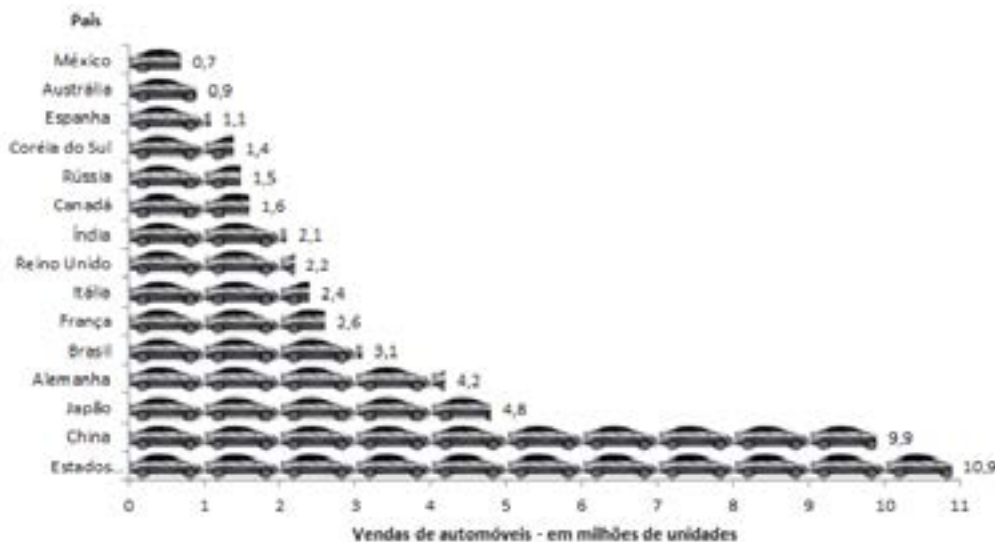


Fonte: IBGE, Censo Demográfico, 2010.

### 3.5.4 Gráficos pictóricos

São gráficos apresentados por meio de figuras que simbolizam fatos estatísticos, ao mesmo tempo que indicam as proporcionalidades.

Figura 3.1 – Vendas de automóveis no mundo – 2009



Fonte: Dos autores, adaptado de Jato Dynamics do Brasil e Fenabrave, 2009.

### 3.5.5 Cartogramas

São gráficos baseados em mapas utilizados para representar as séries geográficas.

Figura 3.2 - Distribuição dos municípios do Vale do Taquari por faixas de número de habitantes - 2006



Fonte: Banco de Dados Regional da Univates, com base nos dados da Fundação de Economia e Estatística do RS.

### 3.6 Dicas de utilização

A finalidade de um gráfico é dar uma ideia mais aproximada possível do comportamento dos fenômenos ou da relação das variáveis da série, permitindo chegar-se a conclusões sobre eles.

Não há uma única maneira de representar graficamente uma série estatística, ou seja, podem-se utilizar vários tipos de gráficos para ilustrar um mesmo objetivo.

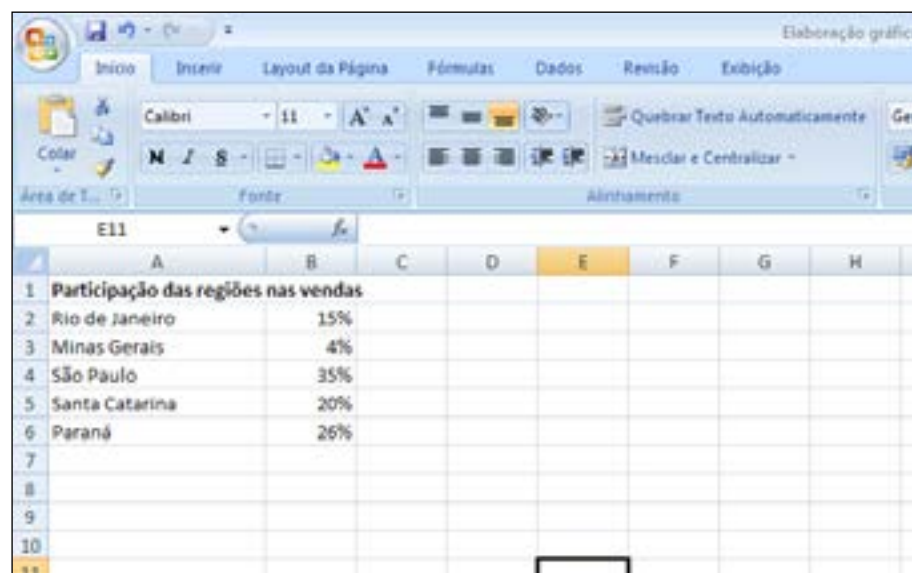
A escolha da representação gráfica ficará a critério do analista, levando em consideração elementos como simplicidade e clareza.

Para evitar erros de interpretação, sugere-se que os eixos iniciem no ponto zero (0,0).

### 3.7 Exemplo de como elaborar um gráfico de pizza

Utilizando o *software* Excel da Microsoft (ou BrOffice), devem-se estruturar as variáveis de maneira a facilitar a elaboração da representação gráfica.

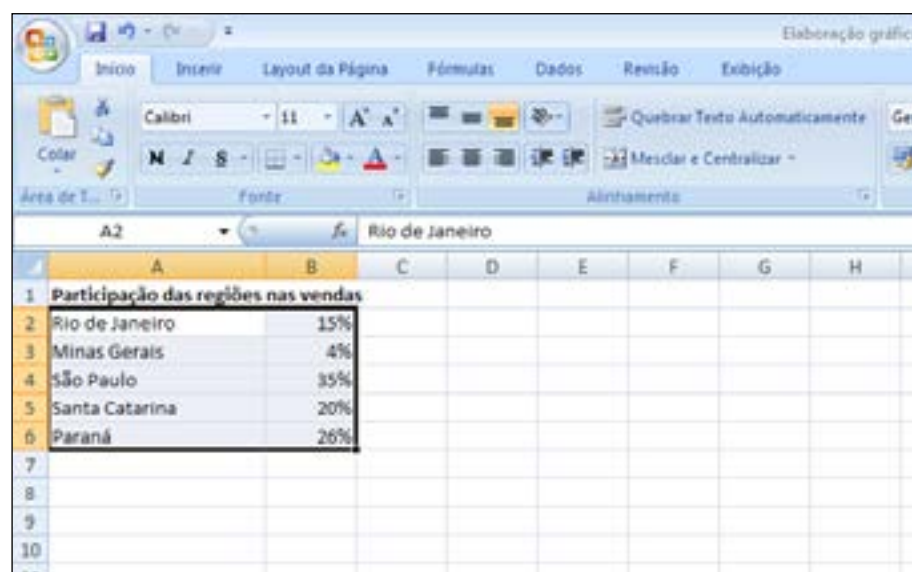
Conforme o modelo abaixo, na primeira coluna (A) foram relacionadas as categoriais e na segunda coluna (B), os valores respectivos da categoria estado.



The screenshot shows the Microsoft Excel interface with the 'Elaboração gráfico' (Chart Wizard) ribbon active. The worksheet contains the following data:

	A	B	C	D	E	F	G	H
1	Participação das regiões nas vendas							
2	Rio de Janeiro	15%						
3	Minas Gerais	4%						
4	São Paulo	35%						
5	Santa Catarina	20%						
6	Paraná	26%						
7								
8								
9								
10								
11								

Após a estruturação, devem-se selecionar as células que contêm as categorias e os dados correspondentes.

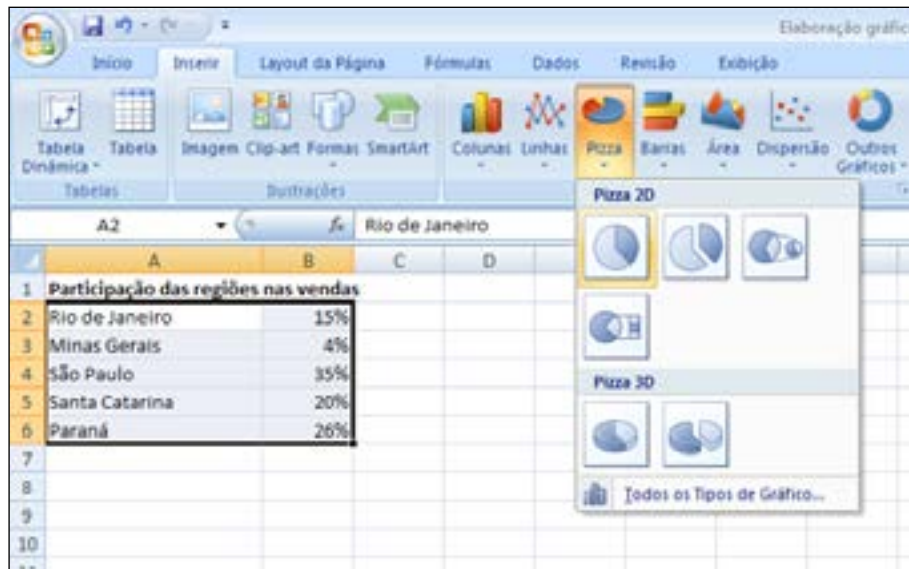


The screenshot shows the same Excel worksheet as before, but with the range A2:B6 selected. The formula bar at the top displays 'Rio de Janeiro'.

	A	B	C	D	E	F	G	H
1	Participação das regiões nas vendas							
2	Rio de Janeiro	15%						
3	Minas Gerais	4%						
4	São Paulo	35%						
5	Santa Catarina	20%						
6	Paraná	26%						
7								
8								
9								
10								
11								

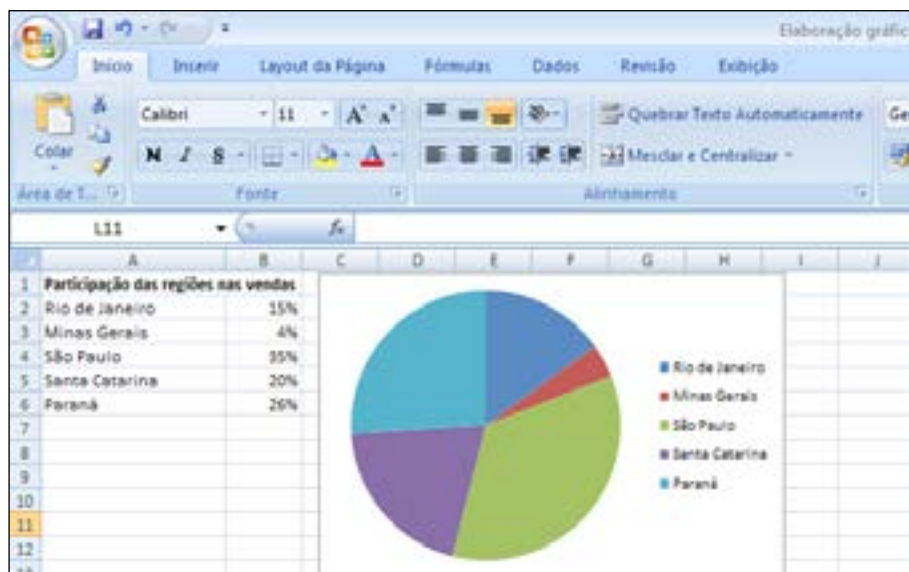


Em seguida, clicar em Inserir no menu e opção Pizza, conforme ilustração a seguir:



Após o gráfico estar visível na tela, pode-se configurá-lo conforme as preferências pessoais (cores, linhas, fundos), e ainda inserir o rótulo de dados, bastando para isso clicar com o botão direito do mouse em cima do gráfico e selecionar a opção Adicionar Rótulo de Dados.

Para os demais tipos de gráficos, devem-se estruturar as variáveis da mesma maneira (em colunas) e inserir o tipo de gráfico correspondente.



### 3.8 Resumo para fixação

Como síntese para fixação, sugerem-se os seguintes pontos em termos de principais aspectos abordados neste capítulo:

- Uma representação gráfica cartesiana tem a estrutura baseada na relação existente entre a variável dependente (Y) variando em função da variável independente (X):  $Y = f(X)$ .
- Um gráfico elaborado a partir de uma distribuição de frequências e séries permite transformar dados em informações para a tomada de decisão. Na elaboração de gráficos a partir de uma distribuição de frequências adotam-se modelos específicos (histograma ou polígono de frequências).
- O gráfico de barras ou de colunas é adequado para todos os tipos de dados. Normalmente utiliza-se o gráfico de colunas, sendo o gráfico de barras mais adequado quando há um número grande de itens ou categorias a ser apresentado, ou ainda quando a expressão da categoria for longa. Dessa forma, facilita-se a apresentação e aproveita-se melhor o espaço da página na vertical que é maior do que na horizontal.
- O gráfico de pizza é o mais indicado para apresentar as proporções e não deve ser utilizado para variáveis com respostas múltiplas. Também evitar utilizar esse tipo de gráfico quando a variável apresentar muitos atributos, bem como quando vários atributos apresentarem percentual pequeno.
- O gráfico de linhas é o mais indicado para dados de séries temporais.
- Não há apenas um modelo de gráfico que possa ser utilizado, mas se deve levar em consideração a simplicidade e a clareza no momento da elaboração.

### 3.9 Exercícios propostos

3.1. A Tabela 3.4 apresenta os resultados da renda bruta anual de uma amostra de 40 pessoas que fizeram declaração de renda no ano base de 2014, em Rincão dos Anjos.

- Faça o histograma dessa distribuição segundo as normas;
- Faça um polígono de frequências segundo as normas;
- Transforme o polígono de frequências em uma curva de frequência.

Tabela 3.4 – Renda bruta anual apurada na declaração de renda de 2014 em Rincão dos Anjos

Renda bruta anual (mil reais)	Número de declarações
05 —  10	5
10 —  15	8
15 —  20	8
20 —  25	10
25 —  30	6
30 —  35	2
35 —  40	1

Fonte: Dados fictícios, apenas para fins ilustrativos.

3.2. A empresa agrícola Verde Vida testou cinco variedades diferentes de fumo na safra de 2013-2014, que acusaram o seguinte resultado em produtividade (em kg):

Variedade	A	B	C	D	E
Produtividade	1.560	1.820	2.000	1.700	1.850

- Faça um gráfico de colunas segundo as normas;
- Faça um gráfico de barras segundo as normas.

3.3. Considere os pares de observações representados pelo preço (R\$) e a respectiva oferta (unidades) de um bem em determinado município gaúcho no mês de junho de 2014.

Preço	3	6	9	8	15	25
Oferta	5	8	10	10	12	15

- Faça o diagrama de dispersão;
- Faça um gráfico de linhas.

3.4. Um analista deseja conhecer a evolução da situação financeira de uma empresa. Para tanto, coletou uma série de dados referentes ao lucro e despesa, conforme Tabela 3.5.

- Faça um gráfico de colunas para cada série;
- Faça um gráfico de linhas colocando as duas séries no mesmo gráfico.

Tabela 3.5 – Faturamento e despesa de uma empresa de Lajeado – 2009/ 2014

Ano	Faturamento (R\$)	Despesa (R\$)
2009	230.500,00	195.100,00
2010	297.200,00	252.700,00
2011	361.900,00	340.800,00
2012	455.700,00	405.400,00
2013	560.000,00	497.200,00
2014	610.600,00	560.800,00

Fonte: Dados fictícios, apenas para fins ilustrativos.

3.5. Um administrador encomendou uma pesquisa de mercado no mês de janeiro de 2012, num município do Vale do Taquari. A pesquisa foi realizada com uma amostra de 500 consumidores, escolhidos aleatoriamente, identificando as preferências por sabor de goma de mascar e encontrou o resultado conforme apresentado na Tabela 3.6.

- Faça um diagrama de dispersão usando os dados do número de usuários representando o eixo X e os da porcentagem de usuários representando o eixo Y;
- Faça um gráfico de setores ou de pizza com os percentuais de usuários de cada categoria (Limão, Morango, Hortelã e Canela), segundo as normas.

Tabela 3.6 – Preferência por sabor de goma de mascar num município do Vale do Taquari - janeiro/2014

Sabor de goma de mascar	Número de consumidores	% de consumidores
Limão	200	40
Morango	125	25
Hortelã	50	10
Canela	125	25
Total	500	100

Fonte: Dados fictícios, apenas para fins ilustrativos.

## 4 MEDIDAS DE TENDÊNCIA CENTRAL

Uma medida de tendência central ou posição de um conjunto de dados mostra o valor em torno do qual se agrupam as observações.

Dentre as medidas de tendência central as mais importantes são: média, mediana e moda.

### 4.1 Média

Representa o centro de gravidade da distribuição da variável, sendo obtida por meio da soma de todos os valores observados dividida pelo número de observações. A sua aplicação requer uma variável quantitativa, ou a transformação de uma variável qualitativa escalar em variável quantitativa. As variações nos valores centrais afetam muito pouco o seu valor, porém é muito influenciado pelos valores extremos da variável. Embora existam três formas diferentes para se obter a média dos dados (média aritmética, média geométrica e média harmônica), a primeira é utilizada de maneira mais universal e as demais são utilizadas de maneira mais específica.

A **média aritmética** é obtida pela soma de todos os valores observados dividida pelo número de observações. A fórmula de cálculo é a seguinte:

$$\bar{X} = \frac{\sum x_i}{n} \quad (4.1)$$

onde:

$\bar{X}$  = símbolo da média aritmética amostral

$\sum$  = somatório, ou resultado da soma

$x_i$  = cada dado da amostra

$n$  = número de dados da amostra

Por exemplo, a média aritmética dos dados 5, 9 e 13 é:

$$\bar{X} = (5 + 9 + 13)/3 = 27/3 = 9$$

A média aritmética, por ser uma medida mais utilizada em estatística, será detalhada num tópico específico neste capítulo.

A **média geométrica** é definida como a raiz  $n$ -ésima do produto de um conjunto de  $n$  valores. Seu cálculo é feito utilizando a seguinte fórmula:

$$\bar{X}^G = \sqrt[n]{\prod X_i} \text{ ou seja } \sqrt[n]{(X_1)(X_1) \dots (X_n)} \quad (4.2)$$

onde:

$\bar{X}^G$  = símbolo da média geométrica

$\prod$  = produtório, ou resultado da multiplicação

$X_i$  = cada dado da amostra

$n$  = número de dados da amostra

Por exemplo, a média geométrica dos dados 5, 9 e 13 é:

$$\bar{X}^G = \sqrt[3]{(5)(9)(13)} = 8,36$$

A média geométrica também é utilizada no cálculo de taxas de crescimento de preços, vendas, população etc. Entretanto, o uso dessa média deverá utilizar índices e não variações.

Para a mesma série de dados, a média aritmética é 9. Comparativamente a média aritmética será sempre maior do que a média geométrica, para qualquer série de valores positivos, com exceção do caso em que os valores da série são todos iguais. Nesse caso, as duas médias coincidem.

A **média harmônica** é representada pelo inverso da média aritmética dos inversos dos valores das observações. O seu resultado tende a receber influências dos valores menores. É usado para determinar a média em uma seleção, em que se buscam candidatos com conhecimentos homogêneos. Pela média aritmética os candidatos poderão ter médias iguais, porém, na média harmônica, os candidatos com pouco conhecimento em determinados assuntos apresentarão médias menores.

A fórmula para determinar a média harmônica é a seguinte:

$$\overline{X}^H = \frac{1}{\frac{\sum (1/x_i)}{n}} = \frac{n}{\sum (1/x_i)} \quad (4.3)$$

onde:

$\overline{X}^H$  = símbolo da média harmônica

$x_i$  = cada dado da amostra

$n$  = número de dados da amostra

$\Sigma$  = somatório das frações

Por exemplo, a média harmônica dos três valores 5, 9 e 13 é:

$$\overline{X}^H = \frac{3}{1/5 + 1/9 + 1/13} = 7,73$$

Observa-se que o resultado, com a aplicação das três formas diferentes de cálculo da média, não é o mesmo para os mesmos dados. Enquanto o resultado da média aritmética é 9, pela média geométrica obteve-se 8,36 e pela média harmônica, 7,73.

Assim, para qualquer série de dados cujos valores não são todos os mesmos e que não incluem o zero, a média harmônica é sempre menor em relação à média aritmética e à média geométrica. Ou, em outras palavras, a média aritmética será o valor maior, a harmônica o menor e a geométrica apresentará um valor entre as duas.

A média aritmética pode ser calculada de duas formas: por meio de dados observados brutos e por meio de dados tabulados.

#### 4.1.1 Média aritmética de dados observados

A média aritmética simples é obtida pela divisão da soma dos valores da variável pelo número de observações.

As fórmulas para determinar a média aritmética de dados observados são as seguintes:

$$\overline{X} = \frac{\sum x_i}{n} \quad (\text{dados de uma amostra}) \quad (4.1)$$

ou

$$\mu = \frac{\sum x_i}{N} \quad (\text{dados de uma população}) \quad (4.4)$$

onde:

$\mu$  = símbolo da média para dados de uma **população**

$\overline{X}$  = símbolo da média para dados de uma **amostra**

$x_i$  = cada dado do conjunto levantado

$n$  = número de dados da amostra

$N$  = número de dados de uma população

Exemplo:

$$n = 5$$

$$x_i = \{2; 4; 5; 6; 8\}$$

$$\bar{X} = \frac{(2+4+5+6+8)}{5} = \frac{25}{5} = 5$$

Outra alternativa de cálculo é a média aritmética ponderada de dados observados que é uma média aritmética na qual cada valor se encontra ponderado de acordo com sua importância no grupo total.

A fórmula para determinar a média aritmética ponderada de dados observados é a seguinte:

$$\mu_w \text{ ou } \bar{X}_w = \frac{\sum x_i w_i}{\sum w_i} \quad (4.5)$$

onde:

W = o peso

$\mu$  = o símbolo da média utilizada quando os dados se referem à população

$\bar{X}$  = o símbolo da média utilizada quando os dados se referem à amostra

$\sum x_i w_i$  = somatório das multiplicações de cada ponto médio com o seu respectivo peso

$\sum w_i$  = somatório dos pesos utilizados

Exemplo:

Notas ( $x_i$ ): {2 ; 4 ; 6}

Pesos ( $w_i$ ): {2 ; 1 ; 5}

$$\bar{X}_w = \frac{(2*2) + (4*1) + (6*5)}{(2+1+5)} = \frac{4+4+30}{8} = \frac{38}{8} = 4,75$$

$$\bar{X}_w = 4,75$$

#### 4.1.2 Média aritmética para dados tabulados

Quando os dados se encontram agrupados em uma distribuição de frequência, o ponto médio de cada classe é utilizado como uma aproximação de todos os valores contidos na classe. A média para dados tabulados é uma média ponderada entre o ponto médio de cada intervalo com as respectivas frequências.

A média aritmética por dados tabulados pode ser calculada por três formas diferentes: pela frequência absoluta, pela frequência relativa e pelo método abreviado.

**Média aritmética** por meio da **frequência absoluta** é obtida pela média ponderada entre os pontos médios de cada intervalo com as respectivas frequências absolutas.

As fórmulas para determinar a média aritmética de dados tabulados pela frequência absoluta são as seguintes:

$$\mu = \frac{\sum X_i f_i}{N} \quad (\text{dados de uma população}) \quad (4.6)$$

ou

$$\bar{X} = \frac{\sum X_i f_i}{n} \quad (\text{dados de uma amostra}) \quad (4.7)$$

onde:

$\sum X_i f_i$  = somatório das multiplicações de cada ponto médio com a sua respectiva frequência

N ou  $\sum f_i$  = total de dados ou soma das frequências de dados de uma população



$n$  ou  $\sum f_i$  = total de dados ou soma das frequências de dados de uma amostra

$\mu$  = o símbolo da média utilizada quando os dados se referem à população

$\bar{X}$  = o símbolo da média utilizada quando os dados se referem à amostra

Exemplo:

Tabela 4.1 – Consumo mensal de energia elétrica (em mil kw/h) na Propriedade “A” – maio de 2012 a abril de 2014

Consumo mensal (em mil kw/h)	$x_i$	Meses ( $f_i$ )	$F_i$	$f_r$	$F_r$
30  — 35	32,5	5	5	0,1389	0,1389
35  — 40	37,5	12	17	0,3333	0,4722
40  — 45	42,5	8	25	0,2222	0,6944
45  — 50	47,5	4	29	0,1111	0,8056
50  — 55	52,5	3	32	0,0833	0,8889
55  — 60	57,5	1	33	0,0278	0,9167
60  — 65	62,5	3	36	0,0833	1,0000

Fonte: Dados fictícios, apenas para fins ilustrativos.

Observação: os símbolos têm os seguintes significados:

$X_i$  = ponto médio de cada intervalo de classe.

$F_i$  = frequência absoluta acumulada (número de meses) para cada intervalo de classe.

$f_r$  = frequência relativa (percentual de meses) para cada intervalo de dados.

$F_r$  = frequência relativa acumulada (percentual de meses) para cada intervalo de classe.

Cálculo da média aritmética de dados tabulados

$$\bar{X} = \frac{(32,5 \times 5) + (37,5 \times 12) + (42,5 \times 8) + (47,5 \times 4) + (52,5 \times 3) + (57,5 \times 1) + (62,5 \times 3)}{36}$$

$$\bar{X} = \frac{162,5 + 450 + 340 + 190 + 157,5 + 57,5 + 187,5}{36} = \frac{1.545}{36} =$$

$$\bar{X} = 42,92$$

A **média aritmética da frequência relativa** é obtida por meio da média ponderada entre os pontos médios de cada intervalo com as respectivas frequências relativas. Como a soma das frequências relativas é igual a 1, pode-se inferir que se obtém a média pela soma das multiplicações dos pontos médios com suas respectivas frequências relativas.

As fórmulas para determinar a média aritmética de dados tabulados pela frequência relativa são as seguintes:

$$\mu = \sum X_i f_r \quad (\text{dados de uma população}) \quad (4.8)$$

ou

$$\bar{X} = \sum X_i f_r \quad (\text{dados de uma amostra}) \quad (4.9)$$

onde:

$\mu$  = o símbolo da média utilizada quando os dados se referem à população

$\bar{X}$  = o símbolo da média utilizada quando os dados se referem à amostra

$\sum X_i f_r$  = somatório das multiplicações de cada ponto médio com a sua respectiva frequência relativa

Exemplo:

Utilizando os dados da tabela estatística anterior tem-se o seguinte:

$$\bar{X} = (32,5 \times 0,1389) + (37,5 \times 0,3333) + (42,5 \times 0,2222) + (47,5 \times 0,1111) \\ + (52,5 \times 0,0833) + (57,5 \times 0,0278) + (62,5 \times 0,0833)$$

$$\bar{X} = 4,51 + 12,5 + 9,44 + 5,28 + 4,38 + 1,60 + 5,21 =$$

$$\bar{X} = 42,92$$

O **método abreviado** de cálculo da **média aritmética** de dados agrupados é obtido por meio da média ponderada entre constantes arbitrárias que substituem os pontos médios de cada intervalo com as respectivas frequências absolutas. Dessa forma elimina-se o grande número de cálculos que às vezes se apresentam quando os pontos médios são de valores muito altos.

As fórmulas para determinar a média aritmética de dados tabulados pelo método abreviado são as seguintes:

$$\mu = x_0 + c_1 \frac{\sum u_i f_i}{N} \quad (\text{dados de uma população}) \quad (4.10)$$

ou

$$\bar{X} = x_0 + c_1 \frac{\sum u_i f_i}{n} \quad (\text{dados de uma amostra}) \quad (4.11)$$

onde:

$\mu$  = o símbolo da média utilizada quando os dados se referem à população

$\bar{X}$  = o símbolo da média utilizada quando os dados se referem à amostra

$X_0$  = o ponto médio do intervalo da origem do trabalho

$C_1$  = amplitude do intervalo

$u_i$  = valor de cada intervalo sendo o intervalo de origem arbitrado com o valor zero, os intervalos inferiores são em ordem decrescente negativos e os intervalos superiores são em ordem crescente positivos

$\sum u_i f_i$  = somatório das multiplicações do valor arbitrário de cada ponto médio do intervalo com a sua respectiva frequência absoluta

$N$  ou  $n$  = o total da frequência de dados, dependendo se forem de uma população ou amostra, respectivamente

Exemplo:

Calcular a média aritmética de dados tabulados pelo método abreviado do exemplo da Tabela 4.1. Ajustando os dados dessa tabela para o cálculo por esse método, obtém-se a Tabela 4.1a.

Tabela 4.1a – Consumo mensal de energia elétrica (em mil kw/h) na Propriedade “A” – maio de 2012 a abril de 2014

$x_i$	$f_i$	$u_i$	$u_i \cdot f_i$
32,5	5	-1	-5
37,5	12	0	0
42,5	8	1	8
47,5	4	2	8
52,5	3	3	9
57,5	1	4	4
62,5	3	5	15
Total	36		39

Fonte: Dados fictícios, apenas para fins ilustrativos.

Os símbolos apresentados têm os seguintes significados:

$x_i$  = ponto médio de cada intervalo de classe

$f_i$  = frequência absoluta (número de meses) para cada intervalo de classe

$u_i$  = valor arbitrário para cada intervalo de dados

$u_i f_i$  = multiplicações de cada valor arbitrário do ponto médio do intervalo com a sua respectiva frequência absoluta

Assim sendo:

$$X_0 = 37,5$$

$C_i = 5$ , onde:

$$\bar{X} = X_0 + C_i \frac{\sum u_i f_i}{n}$$

$$\bar{X} = 37,5 + 5 \frac{39}{36}$$

$$\bar{X} = 42,92$$

Observa-se, pela aplicação dos cálculos ao mesmo exemplo, pelas três diferentes formas, que os resultados são idênticos. Portanto, a opção a ser usada na solução de problemas depende das circunstâncias de cada problema e dos dados disponíveis.

#### 4.1.3 Propriedades da média aritmética

Tendo em vista que se trata de uma medida central, a média aritmética tem algumas propriedades:

1 - A soma dos desvios com relação à média é = 0

Para dados brutos:

$$\sum (X_i - \bar{X}) = 0$$

Exemplo:

$$X_i = \{2 ; 4 ; 6\}$$

$$\bar{X} = 4$$

Assim:

$$(2 - 4) + (4 - 4) + (6 - 4) = 0$$

Para dados tabulados:

$$\sum f_i (X_i - \bar{X}) = 0$$

Exemplo:

Conforme Tabela 4.1:

$$\bar{X} = 42,92$$

Assim:

$$5 (32,5 - 42,92) + 12 (37,5 - 42,92) + 8 (42,5 - 42,92) + 4 (47,5 - 42,92) + 3 (52,5 - 42,92) + 1 (57,5 - 42,92) + 3 (62,5 - 42,92) = 0$$

2 - Se uma população de N elementos se divide em p extratos ou grupos, a média da população é igual à média ponderada das médias dos grupos ou extratos.

Exemplo:

Tabela 4.2 – Resultados de uma população dividida em três grupos

Grupo	$X_i$	$f_i$	$X_i \cdot f_i$
10  — 20	15	20	300
20  — 30	25	45	1125
30  — 40	35	13	455
S		78	1880

Fonte: Dados fictícios, apenas para fins ilustrativos.

Solução:

Utilizando a fórmula 4.5 da média ponderada, obtém-se o seguinte resultado:

$$\bar{X}_w = \frac{(15 \cdot 20) + (25 \cdot 45) + (35 \cdot 13)}{(20 + 45 + 13)} = \frac{300 + 1125 + 455}{78} = \frac{1880}{78} = 24,1$$

$$\bar{X}_w = 24,1$$

Calculando a média da população utilizando a fórmula 4.5, observa-se que se chega ao mesmo resultado:

$$\mu = \frac{1880}{78} = 24,1$$

## 4.2 Mediana

A mediana divide a distribuição ou o conjunto de dados de uma amostra em duas partes iguais. Para aplicar a medida da mediana é necessário que a variável possa ser ordenada em forma de um rol. A sua aplicação requer uma variável quantitativa, ou a transformação de uma variável qualitativa escalar em variável quantitativa.

### 4.2.1 Mediana para dados brutos

Número ímpar de observações

A mediana de um grupo de itens é o valor do item central quando todos os itens do grupo forem dispostos, em termos de valor, em ordem crescente ou decrescente.

Exemplo: 1 ; 2 ; 5 ; 7 ; 9 ; 13 ; 15 ; 15 ; 18 ; 22 ; 22.

O número 13 é o elemento central, portanto é a mediana.

Número par de observações

O conjunto de dados é formado por número par de elementos, têm-se dois elementos centrais e a mediana será a média aritmética dos dois elementos centrais.

Exemplo: 5 ; 8 ; 8 ; 11 ; 11 ; 11 ; 14 ; 16

Os elementos centrais são: 11 e 11

A mediana será igual:

$$\text{Med} = \frac{11 + 11}{2} = 11$$

### 4.2.2 Mediana para dados agrupados

Para dados agrupados, a determinação da mediana é feita em duas fases. Primeiro se determina a classe que contém o valor mediano, que é a primeira classe cuja frequência acumulada é igual ou maior que a metade do número de observações. Assim, a classe que contém o valor mediano tem frequência absoluta acumulada  $\geq n/2$ . Na segunda etapa é determinada, por interpolação, a posição da mediana dentro da classe. A determinação da mediana para dados tabulados é dada pela seguinte fórmula:

$$Med = li_k + A_k \frac{\frac{n}{2} - F_{k-1}}{f_k} \quad (4.12)$$

onde:

Med = valor da mediana

$li_k$  = limite inferior do intervalo que contém a mediana

$A_k$  = amplitude do intervalo que contém a mediana

$n$  = nº de observações

$f_k$  = frequência absoluta da classe que contém a mediana

$F_{k-1}$  = frequência absoluta acumulada anterior à classe que contém a mediana

Exemplo:

Utilizando as informações da Tabela 4.1, determinar a mediana:

Tabela 4.1 – Consumo mensal de energia elétrica (em mil kw/h) na Propriedade “A” – maio de 2012 a abril de 2014

Consumo mensal (kw/h)	$X_i$	Meses ( $f_i$ )	$F_i$	$f_r$	$F_r$
30  — 35	32,5	5	5	0,1389	0,1389
35  — 40	37,5	12	17	0,3333	0,4722
40  — 45	42,5	8	25	0,2222	0,6944
45  — 50	47,5	4	29	0,1111	0,8056
50  — 55	52,5	3	32	0,0833	0,8889
55  — 60	57,5	1	33	0,0278	0,9167
60  — 65	62,5	3	36	0,0833	1,0000

Fonte: Dados fictícios, apenas para fins ilustrativos.

Solução:

1º passo: Localizar o ponto que contém a mediana, sendo:

$n = 36$

Ponto que contém a mediana:  $n/2 = 36/2 = 18$

A mediana é dada pela décima oitava observação que é encontrada no 3º intervalo, onde:

$F_i = 25$ , sendo  $25 > 18$

2º passo: determinar o valor da mediana no intervalo escolhido

Com base na fórmula de cálculo 4.12, dispõe-se dos seguintes dados:

$f_k = 8$

$li_k = 40$

$F_{k-1} = 17$

$A_k = 5$

$$n = 36$$

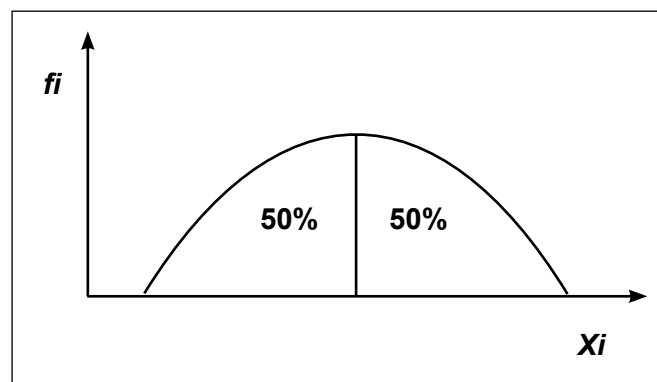
Assim,

$$\text{Med} = 40,625$$

O resultado significa que em 50% do período o consumo foi menor de 40,625 kw/h e em 50% do período o consumo foi igual ou superior a 40,625 kw/h.

Graficamente a mediana divide a curva em duas partes exatamente iguais (FIGURA 4.1).

Figura 4.1 – Ilustração gráfica da mediana



Fonte: Dos autores.

#### 4.2.3 Outras formas de separatrizes

Além da mediana, existem outras formas de separatrizes que subdividem a distribuição de frequências em partes iguais. As formas mais conhecidas são: quartil, decil e centil.

Os quartis, decis e centis são muito similares à mediana, uma vez que também subdividem a distribuição de medidas de acordo com a proporção das frequências observadas.

Enquanto que a mediana divide a distribuição em duas metades, os quartis dividem-na em quatro quartos, os decis em 10 décimos e os pontos centis dividem a distribuição em 100 partes.

A determinação é feita de forma similar à da mediana, substituindo  $n/2$  pela respectiva forma:

$$\text{Quartil: } Q1 = n/4; Q2 = 2n/4; Q3 = 3n/4; Q4 = 4n/4$$

$$\text{Decil: } D1 = n/10; D2 = 2n/10; D3 = 3n/10; \dots\dots\dots D5 = 5n/10; \dots\dots\dots; D8 = 8n/10; \text{ etc.}$$

$$\text{Centil: } C1 = n/100; \dots\dots\dots; C10 = 10n/100; \dots\dots\dots;$$

$$C30 = 30n/100; \text{ etc.}$$

Exemplos:

Utilizando as informações da Tabela 4.1, calcular:

a) o 3º quartil.

b) o 3º decil.

c) o 95º centil.

Solução a):

1. Determinação do intervalo de localização do 3º quartil:

$$\frac{3n}{4} = \frac{3 \times 36}{4} = 27$$

$27 \leq F_i$ ;  $27 \leq 29$ : portanto,

o 3° Quartil localiza-se no intervalo 45 |— 50

2. Determinação do valor do 3° quartil:

$$Q_3 = li_k + A_k \frac{\frac{3n}{4} - F_{k-1}}{f_k} \quad (4.13)$$

sendo:

$$f_k = 4$$

$$li_k = 45$$

$$F_{k-1} = 25$$

$$A_k = 5$$

$$n = 36$$

Assim:

$$Q_3 = 45 + 5 \frac{27 - 25}{4} = 45 + 2,5 = 47,5$$

Resposta:

Em 75% do período o consumo foi menor do que 47.500 kw/h ou em 25% do período o consumo foi de 47.500 kw/h ou mais.

Solução b):

1. Determinação do intervalo em que está localizado o 3° decil:

$$\frac{3n}{10} = \frac{3 \times 36}{10} = 10,8$$

$10,8 \leq F_i$ ;  $10,8 \leq 17$ : portanto, o 3° decil localiza-se no intervalo 35 |— 40

2. Determinação do valor do 3° decil:

$$D_3 = li_k + A_k \frac{\frac{3n}{10} - F_{k-1}}{f_k} \quad (4.14)$$

sendo:

$$f_k = 12$$

$$li_k = 35$$

$$F_{k-1} = 5$$

$$A_k = 5$$

$$n = 36$$

Assim:

$$D_3 = 35 + 5 \frac{10,8 - 5}{12} = 35 + 2,92 = 37,42$$

Resposta:

Em 30% do período o consumo foi menor do que 37.420 kw/h ou em 70% do período o consumo foi de 37.420 kw/h ou mais.

Solução c):



1. Determinação do intervalo que contém o 95° centil:

$$\frac{95n}{100} = \frac{95 \times 36}{100} = 34,2$$

$34,2 \leq F_i$ ;  $34,2 \leq 36$ ; portanto, o 95° centil localiza-se no intervalo 60 |— 65

2. Determinação do valor do 95° centil:

$$C_{95} = li_k + A_k \frac{\frac{95n}{100} - F_{k-1}}{f_k} \quad (4.15)$$

sendo:

$$f_k = 3$$

$$li_k = 60$$

$$F_{k-1} = 33$$

$$A_k = 5$$

$$n = 36$$

Assim:

$$C_{95} = 60 + 5 \frac{34,2 - 33}{3} = 60 + 2,0 = 62,00$$

Resposta: em 95% do período o consumo foi menor do que 62.000 kw/h ou em 5% do período o consumo foi de 62.000 kw/h ou mais.

### 4.3 Moda

Moda ou valor modal é aquele que se apresenta com maior frequência dentro de uma série. É uma das medidas de tendência central das mais simples, sendo aplicável a qualquer variável, seja quantitativa ou qualitativa. Uma série poderá ser:

- Unimodal: quando possui um único valor para a moda;
- Bimodal: quando possui dois valores como moda para a série;
- Multimodal: quando possui mais de dois valores como moda;
- Amodal: quando não possui nenhum valor modal.

#### 4.3.1 Moda para dados brutos

Para determinar o valor da moda de uma série de observações, devemos colocar os dados em forma de rol e verificar qual o número que aparece com maior frequência.

Exemplos:

1) 1 ; 3 ; 3 ; 3 ; 3 ; 5 ; 5 ; 7 ; 7 ; 9 ; 12

A série é unimodal e o valor da moda é 3.

2) 3 ; 3 ; 3 ; 3 ; 4 ; 4 ; 4 ; 4

A série é bimodal, sendo 3 e 4 as respectivas modas.

3) 1 ; 3 ; 3 ; 3 ; 5 ; 7 ; 7 ; 7 ; 8 ; 8 ; 9 ; 9 ; 9 ; 10

A série é multimodal, sendo 3, 7 e 9 as respectivas modas.

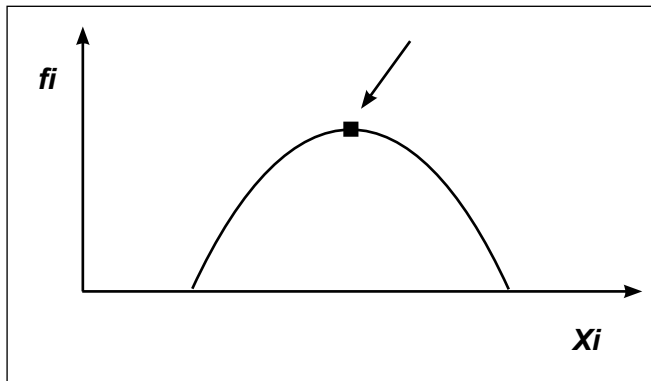
4) 1 ; 2 ; 3 ; 4 ; 5 ; 6 ; 7 ; 8 ; 10 ; 12

A série é amodal, ou seja, não possui moda.

Normalmente num conjunto de dados a tendência é ter uma única moda. Quando há duas ou mais modas, o mais indicado é a verificação de qual fator provoca cada moda. Por exemplo, numa base de dados pode haver uma moda das pessoas da faixa de renda mais alta e outra moda da faixa de renda mais baixa. Nesse caso, o indicado é analisar separadamente cada grupo que apresenta a sua moda.

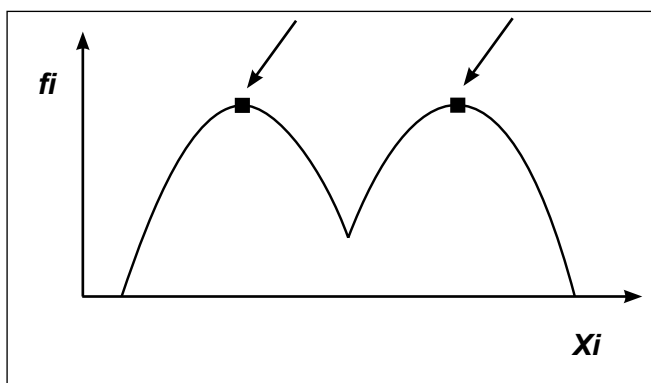
Graficamente pode-se observar facilmente a moda. O valor modal localiza-se na parte mais alta da curva, na qual a frequência tem o maior valor.

Figura 4.2 - Moda – Unimodal



Fonte: Dos autores.

Figura 4.3 - Moda Bimodal



Fonte: Dos autores.

### 4.3.2 Moda para dados agrupados

Para dados agrupados, em uma distribuição de frequência com intervalos de classe com iguais amplitudes, a determinação da moda é feita em duas fases.

Primeiro se determina a classe que contém o valor modal, identificando a classe com maior número de observações.

Na segunda etapa é determinada por interpolação a posição da moda dentro da classe.

A determinação da moda para dados tabulados é dada pela fórmula de Czuber:

$$Mod = li_k + A_k \frac{\Delta 1}{\Delta 1 + \Delta 2} \quad (4.16)$$

onde:

Mod = Valor da moda

$\Delta 1$  = diferença entre a frequência modal e a anterior ( $f_k - f_{k-1}$ )

$\Delta 2$  = diferença entre a frequência modal e a posterior ( $f_k - f_{k+1}$ )

$f_k$  = frequência absoluta da classe que contém a moda

$f_{k-1}$  = frequência absoluta anterior à classe que contém a moda

$f_{k+1}$  = frequência absoluta posterior à classe que contém a moda

$li_k$  = limite inferior do intervalo que contém a moda

$A_k$  = amplitude do intervalo que contém a moda

Exemplo:

Utilizando a Tabela 4.1 b, determinar o valor da moda.

Tabela 4.1 b – Consumo mensal de energia elétrica (em mil kw/h) na Propriedade “A” – maio de 2012 a abril de 2014

Consumo mensal (kw/h)	$X_i$	Meses ( $f_i$ )	$F_i$
30  — 35	32,5	5	5
35  — 40	37,5	12	17
40  — 45	42,5	8	25
45  — 50	47,5	4	29
50  — 55	52,5	3	32
55  — 60	57,5	1	33
60  — 65	62,5	3	36

Fonte: Dados fictícios, apenas para fins ilustrativos.

Os símbolos têm os seguintes significados:

$X_i$  = ponto médio de cada intervalo de classe.

$f_i$  = frequência absoluta (meses) para cada intervalo de dados.

$F_i$  = frequência absoluta acumulada (número de meses) para cada intervalo de classe.

### 1. Determinação do intervalo que contém a moda

O intervalo 35 |— 40 tem o maior número de observações, portanto, é neste intervalo que se encontra a moda.

### 2. Determinação do valor da moda

$$Mod = li_k + A_k \frac{\Delta 1}{\Delta 1 + \Delta 2}$$

sendo:

Mod = Valor da moda a ser calculada

$$\Delta 1 = (f_k - f_{k-1}) = 12 - 5 = 7$$

$$\Delta 2 = (f_k - f_{k+1}) = 12 - 8 = 4$$

$$f_k = 12$$

$$f_{k-1} = 5$$

$$f_{k+1} = 8 \quad li_k = 35 \quad A_k = 5$$

Assim:

$$Mod = 35 + 5 \frac{7}{7 + 4} = 35 + 3,18 = 38,18$$

Resposta: O consumo de energia elétrica com maior frequência no período é de 38.180 kw/h.

A exemplo dos dados brutos, nos dados tabulados também pode existir mais de uma moda. Nesse caso, deve-se calcular o valor de cada moda separadamente.

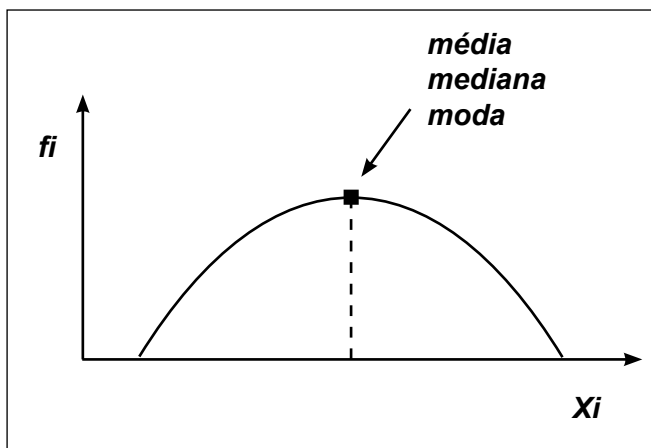
#### 4.4 Relação entre a média, a mediana e a moda

Para os dados agrupados representados por uma curva de frequência, as diferenças entre os valores da média, mediana e da moda são indicadores do formato da curva em termos de assimetria.

Nessa relação podemos encontrar três tipos de curva:

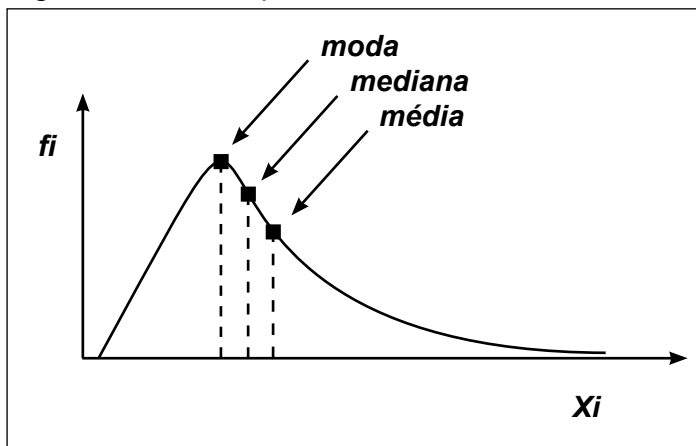
- curva simétrica ou normal: ocorre quando os valores da média, da mediana e da moda são iguais. Tem o formato de sino e os dois lados a partir do centro são iguais (FIGURA 4.4);
- curva assimétrica à direita: ocorre quando o valor da moda é menor do que os valores da mediana e da média (FIGURA 4.5);
- curva assimétrica à esquerda: ocorre quando o valor da moda é maior do que os valores da mediana e da média (FIGURA 4.6);

Figura 4.4 - Exemplo de curva simétrica



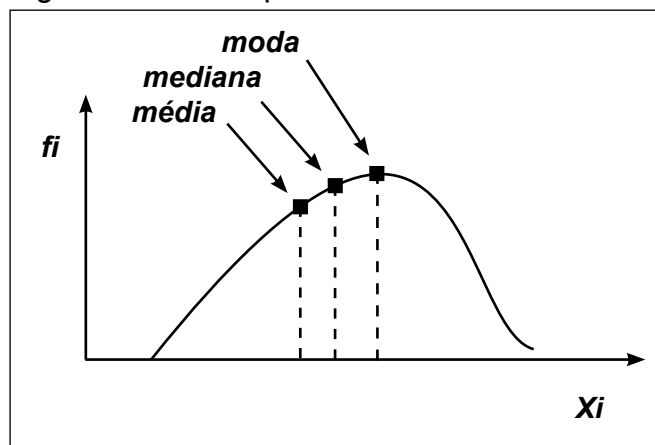
Fonte: Dos autores.

Figura 4.5 - Exemplo de curva assimétrica à direita



Fonte: Dos autores.

Figura 4.6 - Exemplo de curva assimétrica à esquerda



Fonte: Dos autores.

#### 4.5 Resumo para fixação

Como síntese para fixação, sugerem-se os seguintes pontos em termos de principais aspectos abordados neste capítulo:

- As medidas de tendência central mostram o valor em torno do qual se agrupam as observações. As medidas mais utilizadas são a média (aritmética, geométrica e harmônica), a mediana e a moda;
- A média geométrica é mais indicada para cálculos de média de índices de preços, crescimentos populacionais etc.;
- A média harmônica é usada principalmente em processos de seleção, pois pela média aritmética os candidatos poderão ter médias iguais, entretanto pela média harmônica os candidatos com pouco conhecimento em determinada área apresentarão médias menores;
- A aritmética é a média mais utilizada no dia a dia. Ela pode ser simples (somatório dos valores dividido pelo número de observações) ou ponderada (nela cada valor se encontra ponderado de acordo com sua importância ou peso);
- De maneira geral a média aritmética será a maior, a harmônica a menor e a geométrica com valor entre ambas;
- A mediana divide a distribuição ou o conjunto de dados em duas partes iguais;
- A moda pode ser aplicada em variáveis qualitativas e quantitativas. Representa o valor ou variável que mais se repete no conjunto de informações.

## 4.6 Exercícios propostos

4.1. O resultado de tempo gasto por indivíduo de dois grupos para uma determinada tarefa foi o seguinte:

I) 14, 3, 1, 7, 4, 3, 29, 41, 16, 8

II) 15, 2, 7, 4, 19, 4, 43, 31, 13

Usando os valores brutos:

- Calcular a média do tempo para cada grupo;
- Qual foi a mediana do tempo para cada grupo;
- Qual foi a moda do tempo para cada grupo.

4.2. Um professor de Fundamentos de Estatística atribuiu pesos às avaliações a serem realizadas na disciplina, ou seja, a primeira e a segunda avaliação tiveram peso 3 (cada) e a terceira avaliação teve peso 4. Diante disso, calcule a média final na disciplina, tendo que um aluno tirado 6, 7, 5, respectivamente, na 1ª, 2ª e 3ª avaliação.

4.3. Calcule a média harmônica para o conjunto de dados: 8; 4; 3; 10; 23; 3; 9; 11; 17; 22; 1.

4.4. Calcule a média geométrica de uma categoria de trabalhadores que tiveram um aumento salarial de 20% após um mês, 12% após dois meses e 7% após três meses. Qual o percentual médio mensal de aumento dessa categoria?

4.5. Dada a Tabela 4.3, determine:

- Média da duração das válvulas;
- A mediana;
- A moda;
- O 30º centil;
- Faça a curva da frequência, localize a média, mediana e moda e avalie a simetria da curva.

Tabela 4.3 - Duração das válvulas fabricadas pela Empresa Alfa no mês de março de 2014

Duração – Horas	Número de válvulas
300  — 400	14
400  — 500	46
500  — 600	58
600  — 700	76
700  — 800	68
800  — 900	62
900  — 1000	48
1000  — 1100	22
1100  — 1200	6

Fonte: Dados fictícios, apenas para fins ilustrativos.



4.6. Calcular a média, a moda e a mediana da seguinte amostra correspondente à distribuição por tamanhos das vendas de pares de calçados em uma loja:

Tabela 4.4 – Vendas de pares de calçados conforme o tamanho, na Loja Beta, durante o primeiro semestre de 2014

Tamanho	Pares vendidos
37	15
38	70
39	192
40	293
41	260
42	133
43	38
44	7

Fonte: Dados fictícios, apenas para fins ilustrativos.

4.7. Na Tabela 4.5 são apresentados os números de acidentes com veículos ocorridos num município do Vale do Taquari. Determine:

- Qual a percentagem de meses com até 14 acidentes?
- Qual a percentagem de meses com mais de 12 acidentes?
- Qual é a média mensal de acidentes?
- Nos 10% dos meses que houve o menor número de acidentes, qual foi o máximo de acidentes?
- Qual foi a quantidade de acidentes que ocorreu em maior número de meses?

Tabela 4.5 – Número de acidentes com veículos no ano de 2012 num município do Vale do Taquari

Número de acidentes com veículos	Número de meses
-0,5 — 2,5	3
2,5 — 5,5	8
5,5 — 8,5	10
8,5 — 11,5	20
11,5 — 14,5	11
14,5 — 17,5	5
17,5 — 20,5	3

Fonte: Dados fictícios, apenas para fins ilustrativos.

## 5 MEDIDAS DE DISPERSÃO OU DE VARIABILIDADE

As medidas de dispersão ou de variabilidade dizem respeito à descrição de um grupo de valores em termos da variabilidade existente entre os itens incluídos dentro do grupo, tanto em termos absoluto ou relativo dos dados. É comum encontrar séries que, apesar de terem medidas de tendência central iguais, são compostas de maneira distinta. Assim, essas medidas indicam quão dispersos estão os valores da amostra em torno da medida de centralidade.

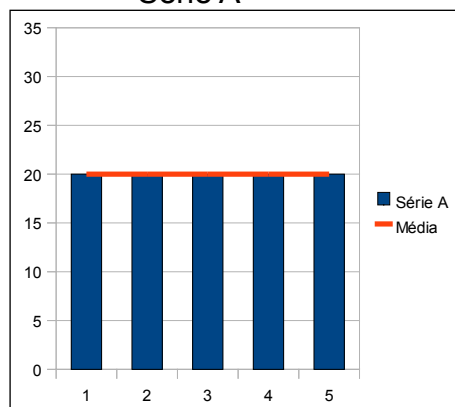
Exemplos:

- Série A: 20 ; 20 ; 20 ; 20 ; 20
- Série B: 20 ; 21 ; 20 ; 19 ; 20
- Série C: 15 ; 10 ; 20 ; 25 ; 30

As séries A, B e C acima apresentadas têm a mesma média e mediana, cujo valor é igual a 20, mas elas apresentam características diferentes. A série A não apresenta nenhuma variação. Já a série B apresenta pouca dispersão, enquanto os valores da série C estão bastante dispersos em torno da média. Assim, podemos observar que, quanto maiores os valores encontrados para as medidas de dispersão, mais afastados estarão os dados entre si e menor será a importância das medidas de posição central para a tomada de decisão.

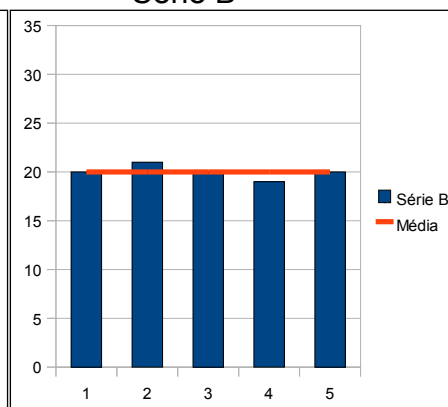
Graficamente, as diferentes variabilidades são perceptíveis.

Gráfico 5.1 – Amostra  
Série A



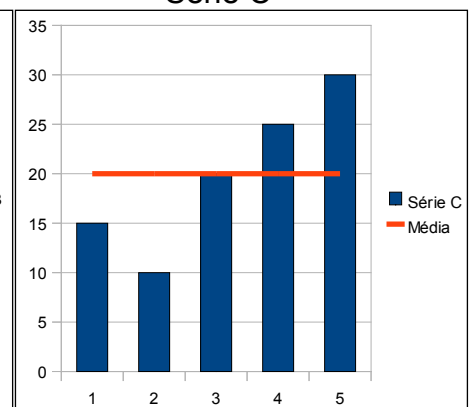
Fonte: Dos autores.

Gráfico 5.2 – Amostra  
Série B



Fonte: Dos autores.

Gráfico 5.3 – Amostra  
Série C



Fonte: Dos autores.

No Gráfico 5.1 observa-se que há uma homogeneidade dos dados, enquanto nos dados da série B (Gráfico 5.2) já se percebe uma variabilidade. Por fim, no Gráfico 5.3 observa-se uma variabilidade bem maior, mesmo que as três séries tenham a mesma média.

Para medir a dispersão de valores são usadas diversas formas, entre as quais as mais importantes são: desvio médio, variância, desvio-padrão, coeficiente de variabilidade, variância relativa e amplitude interquartilica. A seguir essas medidas serão apresentadas individualmente.

### 5.1 Desvio médio

O desvio médio analisa a dispersão dos dados em torno da média. Em outras palavras, trata-se da média aritmética dos desvios dos dados observados em relação à média, considerados em módulos (valor absoluto). Ele é utilizado para comparar duas distribuições com igual média e saber qual das duas está mais ou menos dispersa, pois representa o afastamento médio dos dados em torno da média.

O desvio médio é obtido pelo somatório das diferenças, consideradas em termos de módulos, de cada um dos valores de um conjunto de informações e sua média aritmética, cujo valor deverá

ser dividido pelo número de informações desse conjunto, conforme pode ser observado nas seguintes fórmulas:

Para dados brutos:

Série de dados oriundos de uma população:

$$DM = \frac{\sum |x_i - \mu|}{N} \quad (5.1)$$

onde:

$\sum$  = somatório

$x_i$  = cada dado do conjunto levantado

$\mu$  = média aritmética da população

$N$  = número de elementos da população

Série de dados oriundos de uma amostra:

$$DM = \frac{\sum |x_i - \bar{X}|}{n} \quad (5.2)$$

Onde:

$x_i$  = cada dado do conjunto levantado

$\bar{X}$  = média aritmética da amostra

$n$  = número de elementos da amostra

Exemplo:

O desvio médio da série A: {9, 7, 6, 1, 12} pode ser calculado da seguinte forma:

$$\bar{X} = \frac{\sum x_i}{n} = \frac{9+7+6+1+12}{5} = 7$$

$$DM = \frac{\sum |x_i - \bar{X}|}{N} = \frac{|9-7|+|7-7|+|6-7|+|1-7|+|12-7|}{5} = \frac{2+0+1+6+5}{5} = 2,8$$

Resposta:

O valor encontrado para o desvio médio da série A foi igual a 2,8. Isso significa que os dados se afastam em média 2,8 da média aritmética da série analisada.

Como o valor do desvio médio é expresso em termos absolutos, ele não pode ser utilizado para realizar comparações com outras variáveis, pois ele apenas indica quanto cada variável está dispersa em torno de sua média. Além disso, o desvio médio absoluto é considerado uma medida viesada, pois o desvio médio de amostras não condiz com o desvio médio da população. Por esses motivos, o desvio médio é uma medida menos utilizada, se comparado ao desvio-padrão.

Para dados agrupados:

Tabela de frequência com dados de uma população:

$$DM = \frac{\sum |X_i - \mu| f_i}{N} \quad (5.3)$$

onde:

$X_i$  = ponto médio de cada intervalo de classe

$\mu$  = média aritmética da população

$N$  = número de elementos da população

$f_i$  = frequência de cada intervalo de classe

Tabela de frequência com dados de uma amostra:

$$DM = \frac{\sum |X_i - \bar{X}| f_i}{n} \quad (5.4)$$

onde:

$X_i$  = ponto médio de cada intervalo de classe

$\bar{X}$  = média aritmética da amostra

$n$  = número de elementos da amostra

$f_i$  = frequência de cada intervalo de classe

Exemplo:

Usando os dados da Tabela 5.1, para determinar o desvio médio precisa-se conhecer os seguintes dados:

$\bar{X}$

$|X_i - \bar{X}|$

$|X_i - \bar{X}| \cdot f_i$

Tabela 5.1 – Salário pago (valor por hora) a uma amostra de funcionários da empresa AB

Salário	Nº de pessoas	$X_i$	$X_i \cdot f_i$	$ X_i - \bar{X} $	$ X_i - \bar{X}  \cdot f_i$
10  — 20	5	15	75	15-37=22	22x5=110
20  — 30	8	25	200	25-37=12	12x8=96
30  — 40	11	35	385	35-37=2	2x11=22
40  — 50	9	45	405	45-37=8	8x9=72
50  — 60	4	55	220	55-37=18	18x4=72
60  — 70	3	65	195	65-37=28	28x3=84
$\Sigma$	40		1480		456

Fonte: Dados fictícios, apenas para fins ilustrativos.

Sendo a média dada pela seguinte fórmula:

$$\bar{X} = \frac{\sum X_i f_i}{\sum f_i} = \frac{1480}{40} = 37$$

Assim, a partir desses dados determina-se o desvio médio:

$$DM = \frac{\sum |X_i - \bar{X}| f_i}{n} = \frac{456}{40} = 11,40$$

Resposta:

Os valores estão dispersos a uma distância média de 11,40 em relação à média da distribuição.

## 5.2 Variância

A variância representa a média aritmética dos quadrados dos desvios em relação à média aritmética em termos absolutos. Em outras palavras, ela é obtida pela soma das diferenças entre cada elemento da população ou amostra e sua média aritmética, elevadas ao quadrado, cujo valor deverá ser dividido pela quantidade de elementos do conjunto.

Símbolos:

$\sigma^2$  para dado de uma população;

$s^2$  para dados de uma amostra.

## Fórmulas para dados brutos:

Série de dados oriundos de uma população:

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N} \quad (5.6)$$

onde:

$x_i$  = cada dados do conjunto levantado

$\mu$  = média aritmética da população

$N$  = número de elementos da população

Desenvolvendo essa fórmula, obter-se-á a resultante que permite maior facilidade operacional nos cálculos:

$$\sigma^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{N}}{N} \quad (5.7)$$

onde:

$x_i$  = cada dado do conjunto levantado

$x_i^2$  = o quadrado de cada dado do conjunto levantado

$N$  = número de elementos da população

Série de dados oriundos de uma amostra:

$$s^2 = \frac{\sum (x_i - \bar{X})^2}{n - 1} \quad (5.8)$$

onde:

$x_i$  = elemento  $i$  do conjunto

$\bar{X}$  = média aritmética da amostra

$n$  = número de elementos da amostra

Desenvolvendo essa fórmula, obter-se-á a resultante que permite maior facilidade operacional nos cálculos:

$$s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n - 1} \quad (5.9)$$

onde:

$x_i$  = cada dado do conjunto levantado

$x_i^2$  = o quadrado de cada dado do conjunto levantado

$n$  = número de elementos da amostra

Na amostra, a variância não é exatamente equivalente à variância da população, visto que é introduzido um fator de correção ( $n-1$ ) nessa fórmula, de tal maneira que a variância da amostra seja um estimador não tendencioso da variância da população. Esse fator é utilizado pois a média usada para efeito de cálculo da variância já é estimada da população por meio da mesma amostra. Cada vez que, ao determinar-se um novo cálculo, é introduzido um parâmetro estimado pela mesma amostra, deve-se descontar uma unidade do número de dados levantados.

Fórmulas para dados agrupados:

Tabela de frequência com dados de uma população

$$\sigma^2 = \frac{\sum f_i (X_i - \mu)^2}{N} \quad (5.10)$$

onde:

$X_i$  = ponto médio de cada intervalo de classe

$\mu$  = média aritmética da população

$N$  = número de elementos da população

$f_i$  = frequência de cada intervalo de classe

Desenvolvendo a fórmula 5.10 obtém-se 5.11, que permite maior facilidade operacional nos cálculos:

$$\sigma^2 = \frac{\sum X_i^2 f_i - \frac{(\sum X_i f_i)^2}{N}}{N} \quad (5.11)$$

onde:

$X_i$  = ponto médio de cada intervalo de classe

$X_i^2$  = quadrado do ponto médio de cada intervalo de classe

$N$  = número de elementos da população

$f_i$  = frequência de cada intervalo de classe

Tabela de frequência com dados de uma amostra

$$s^2 = \frac{\sum f_i (X_i - \bar{X})^2}{n-1} \quad (5.12)$$

onde:

$X_i$  = ponto médio de cada intervalo de classe

$\bar{X}$  = média aritmética da amostra

$n$  = número de elementos da amostra

$f_i$  = frequência de cada intervalo de classe

Desenvolvendo essa fórmula, obter-se-á a resultante que permite maior facilidade operacional nos cálculos:

$$s^2 = \frac{\sum X_i^2 f_i - \frac{(\sum X_i f_i)^2}{n}}{n-1} \quad (5.13)$$

onde:

$X_i$  = ponto médio de cada intervalo de classe

$X_i^2$  = quadrado do ponto médio de cada intervalo de classe

$n$  = número de elementos da amostra

$f_i$  = frequência de cada intervalo de classe

A variância é uma medida de análise estatística, mas não tem efeito de aplicação prática, sendo para isso utilizado o desvio-padrão. Ou seja, para se encontrar o desvio-padrão é necessário calcular a variância.

### 5.3 Desvio-padrão

A medida de dispersão mais utilizada é o desvio-padrão. Ele pode ser definido como uma medida da variação dos valores em torno da média.

As características do desvio-padrão de uma amostra, segundo TRIOLA (2008, p.77), são:

- o desvio-padrão é uma medida da variação de todos os valores a partir da média.
- o valor do desvio-padrão  $s$  é usualmente positivo. É zero apenas quando todos os valores dos dados são o mesmo número. Ele nunca é negativo. Também maiores valores do desvio-padrão indicam maior variação.
- o valor do desvio-padrão  $s$  pode crescer drasticamente com a inclusão de um ou mais *outliers* (valores de dados que estão muito afastados dos demais).
- as unidades do desvio-padrão  $s$  (tais como minutos, libras, pés) são as mesmas unidades dos dados originais.

Esses conceitos também podem ser utilizados para o desvio-padrão de uma população.

No cálculo da variância, usa-se o somatório dos quadrados dos desvios, ou seja, o resultado é uma média quadrática. Dessa forma, para poder-se comparar com a unidade da variável que é simples, deve-se extrair a raiz quadrada da variância. A raiz quadrada resultante é chamada de desvio-padrão.

Desvio-padrão da População

$$\sigma = \sqrt{\sigma^2} \quad (5.14)$$

onde:

$\sigma^2$  = variância da população

Desvio-padrão da Amostra

$$s = \sqrt{s^2} \quad (5.15)$$

onde:

$s^2$  = variância da amostra

O valor numérico obtido descreve o maior ou menor grau de dispersão da distribuição com respeito à média aritmética. Ou seja, quando os dados de uma série forem próximos entre si, o desvio-padrão será pequeno. Por outro lado, quanto maior for a diferença entre os dados de uma série, maior será o desvio-padrão.

Exemplos de cálculo da variância e do desvio-padrão

Para dados brutos:

Determine usando os dados da série {5 ; 10 ; 12 ; 15 ; 20}

a) A variância;

Usando a fórmula 5.9 obtém-se o seguinte resultado:

$$s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n - 1}$$

Sendo,

$$\sum x_i = (5 + 10 + 12 + 15 + 20) = 62$$



$$\sum x_i^2 = (5^2 + 10^2 + 12^2 + 15^2 + 20^2) = 894$$

Resposta: A variância da série é 31,30 unidades ao quadrado.

b) O desvio-padrão.

Extrai-se a raiz quadrada da variância.

Assim,

$$s = \sqrt{s^2} \rightarrow s = \sqrt{31,30} \rightarrow s = 5,59$$

Resposta:

Os dados estão dispersos em média em seis unidades em relação à média da distribuição.

Para dados tabulados:

Tendo por base uma amostra da altura de uma parcela da população apresentada na Tabela 5.2, determinar:

- A variância das alturas;
- O desvio-padrão das alturas.

Tabela 5.2 – Estatura de uma amostra de uma população A

Altura (cm)	Nº de pessoas
150  — 158	5
158  — 166	18
166  — 174	42
174  — 182	27
182  — 190	8
$\Sigma$	100

Fonte: Dados fictícios, apenas para fins ilustrativos.

Solução

a) A variância das alturas

Usando a fórmula 5.13 obtém-se o seguinte resultado:

$$s^2 = \frac{\sum X_i^2 f_i - \frac{(\sum X_i f_i)^2}{n}}{n-1}$$

Para calcular a variância, necessita-se conhecer as informações a seguir, cujos valores estão calculados na Tabela 5.3:

$$\sum X_i f_i$$

$$\sum X_i^2 f_i$$

Tabela 5.3 – Tabela auxiliar da Tabela 5.2

Altura (cm)	Nº de pessoas	$X_i$	$X_i f_i$	$X_i^2 f_i$
150  — 158	5	154	770	118.580
158  — 166	18	162	2916	472.392
166  — 174	42	170	7140	1.213.800
174  — 182	27	178	4806	855.468
182  — 190	8	186	1.488	276.768
$\Sigma$	100		17.120	2.937.008

Fonte: Dados fictícios, apenas para fins ilustrativos.

Substituindo os valores na fórmula 5.13 obtém-se os seguintes resultados:

$$s^2 = \frac{2937008 - \frac{17120^2}{100}}{100 - 1}$$

$$s^2 = \frac{2937008 - 2930944}{100 - 1} = \frac{6064}{99}$$

$$s^2 = 61,25$$

Resposta:

A variância das alturas é de 61,25 cm<sup>2</sup>

b) O desvio-padrão das alturas

Solução

Extrai-se a raiz quadrada da variância.

Assim,

$$s = \sqrt{s^2}$$

$$s = \sqrt{61,25}$$

$$s = 7,8263$$

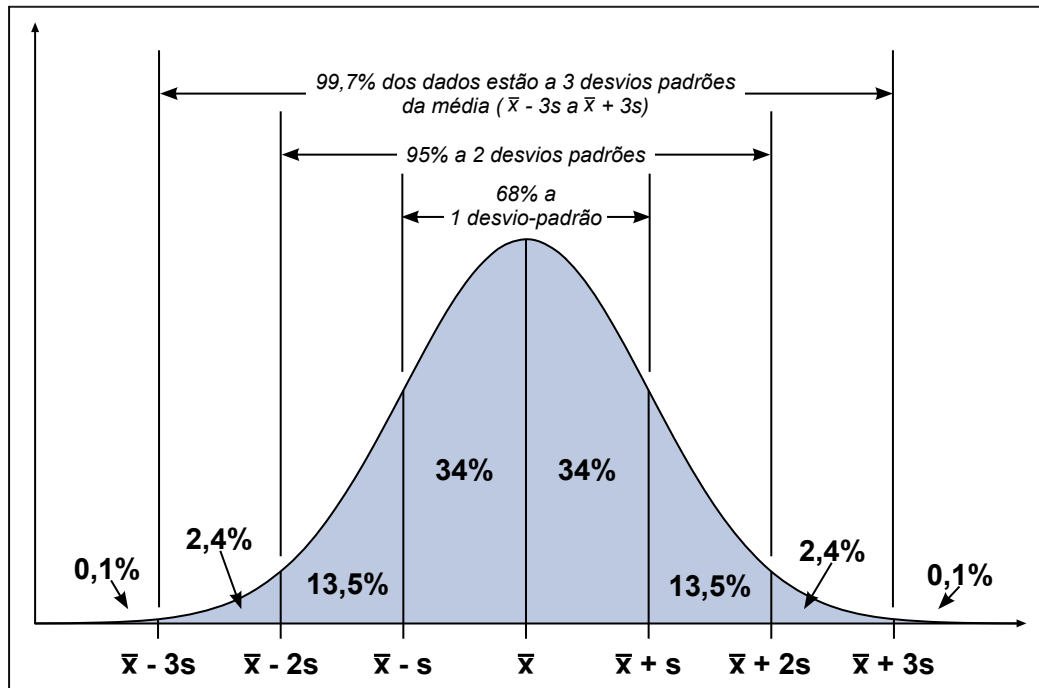
Resposta:

As estaturas das pessoas estão dispersas em média 7,83 cm em relação à média da distribuição.

### 5.3.1 Utilização do Desvio-padrão em Estatística

Uma importante utilização do desvio-padrão em estatística é na determinação do escore **reduzido** da variável **z** na Distribuição Normal de Frequência Padronizada. O escore reduzido **z** mede o afastamento de um valor  $x_i$  em relação à média da distribuição, em unidades do desvio-padrão (FIGURA 5.4).

Figura 5.4 – Curva normal padrão



Fonte: Triola (2008).

O valor de  $z$  é fornecido pela seguinte fórmula:

Para dados de uma população:

$$z = \frac{x_i - \mu}{\sigma} \quad (5.16)$$

onde:

$x_i$  = qualquer dado de um conjunto

$\mu$  = média aritmética da população

$\sigma$  = desvio-padrão da população

ou

Para dados de uma amostra

$$z = \frac{x_i - \bar{X}}{s} \quad (5.17)$$

onde:

$x_i$  = qualquer dado de um conjunto

$\bar{X}$  = média aritmética da amostra

$s$  = desvio-padrão da amostra

A relação na curva normal, mostrada graficamente na Figura 5.4, pode ser analisada em termos de Distribuição Normal, da seguinte forma:

$(\mu + \sigma)$  e  $(\mu - \sigma)$  → ocupa 68,27% da área da curva

$(\mu + 2\sigma)$  e  $(\mu - 2\sigma)$  → ocupa 95,44% da área da curva

$(\mu + 3\sigma)$  e  $(\mu - 3\sigma)$  → ocupa 99,73% da área da curva

Isso significa que, em um conjunto de dados que se distribuem seguindo uma distribuição normal, com média de 100 e desvio-padrão igual a 7, por exemplo, temos cerca de 68,27% de todos os valores da amostra distantes a 1 desvio-padrão da média (para mais e para menos),

ou seja, entre 93 e 107 cerca de 95,44% dos valores estão a 2 desvios-padrões, e 99,7% dos valores estão a 3 desvios-padrões da média. Neste caso, 99,7% dos valores da amostra estariam entre 79 e 121.

Essa análise aplica-se a distribuições simétricas e mesocúrticas e será ampliada no capítulo 9 Distribuição básica de probabilidade de variáveis contínuas, cujas medidas serão detalhadas no capítulo 6, Medidas de Assimetria e Curtose.

### 5.3.2 Desigualdade de Chebyshev

O Teorema de Chebyshev diz que, para qualquer conjunto de dados, o intervalo entre a média menos  $k$  desvios-padrão e a média mais  $k$  desvios-padrão contém sempre pelo menos  $100 \times (1 - 1/k^2)\%$  dos dados.

Assim, ele também permite identificar como os dados estão distribuídos. Embora apresente resultados próximos aos anteriores da utilização do desvio-padrão ou do tópico anterior, ele aplica-se a qualquer conjunto de dados, não havendo a necessidade de esses seguirem uma distribuição normal.

Exemplo:

Um conjunto de dados apresenta como média 25. Quantos por cento dos dados ficam, no mínimo, no intervalo entre 25 mais 2 desvios-padrão e 25 menos 2 desvios-padrão?

Resposta:

$$Perc. = 100 \left( 1 - \frac{1}{2^2} \right) = 100 \left( 1 - \frac{1}{4} \right) = 100(1 - 0,25) = 0,75$$

Percentual = pelo menos 75%

### 5.4 Coeficiente de variabilidade ou de variação

O coeficiente de variação indica a magnitude relativa do desvio-padrão quando comparado com a média da distribuição das medidas. Essa medida é utilizada com frequência quando temos o objetivo de comparar duas ou mais distribuições que possuem médias diferentes, ou estão expressas em unidade de medidas diferentes.

Para calcular o coeficiente é utilizada a seguinte fórmula:

Para dados de uma população:

$$\gamma = \frac{\sigma}{\mu} \quad (5.18)$$

onde:

$\mu$  = média aritmética da população

$\sigma$  = desvio-padrão

ou

Para dados de uma amostra

$$CV = \frac{s}{\bar{X}} \quad (5.19)$$

onde:

$\bar{X}$  = média aritmética da amostra

$s$  = desvio-padrão

### 5.4.1 Características

1 - Não possui unidades, por isso, existe a possibilidade de se compará-lo com diversas distribuições de frequência.

2 - Deixa de ser útil quando a média tende a zero. Nesse caso, o resultado não tem utilidade para a análise da variabilidade.

Exemplo:

Dadas as distribuições da Tabela 5.4, determine:

a) A distribuição mais homogênea.

b) A distribuição mais heterogênea.

Tabela 5.4 – Dados de distribuições

Distribuição A	Distribuição B	Distribuição C
N = 200	N = 50	$\mu = 8$
$\sum X_i f_i = 4000$	$\sum X_i f_i = 500$	$\sum X_i f_i = 3200$
$\sum f_i (X_i - \mu)^2 = 5000$	$\sum X_i^2 f_i = 5450$	$\sum X_i^2 f_i = 32000$

Fonte: Dos autores.

Para responder essas questões, precisa-se calcular o coeficiente de variabilidade, conforme equações 5.18 e 5.19. Portanto, inicialmente, devem-se encontrar a média e o desvio-padrão de cada distribuição, para após poder calcular o coeficiente de variabilidade.

Distribuição A:

$$\mu = \frac{\sum X_i f_i}{N} = \frac{4000}{200} = 20$$

$$\sigma^2 = \frac{\sum f_i (X_i - \mu)^2}{N} = \frac{5000}{200} = 25$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{25} = 5$$

$$CV = \frac{\sigma}{\mu} = \frac{5}{20} = 0,25$$

Distribuição B:

$$\mu = \frac{\sum X_i f_i}{N} = \frac{500}{50} = 10$$

$$\sigma^2 = \frac{\sum X_i^2 f_i - \frac{(\sum X_i f_i)^2}{N}}{N} = \frac{5450 - \frac{500^2}{50}}{50} = \frac{5450 - 5000}{50} = 9$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{9} = 3$$

$$CV = \frac{\sigma}{\mu} = \frac{3}{10} = 0,3$$

Distribuição C:

$$\mu = 8 \text{ como: } \mu = \frac{\sum X_i f_i}{N}, \text{ logo } N = \frac{\sum X_i f_i}{\mu} = \frac{3200}{8} = 400$$

$$\sigma^2 = \frac{\sum X_i^2 f_i - \frac{(\sum X_i f_i)^2}{N}}{N} = \frac{32000 - \frac{3200^2}{400}}{400} = \frac{32000 - 25600}{400} = 16$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{16} = 4$$

$$CV = \frac{\sigma}{\mu} = \frac{4}{8} = 0,5$$

Resposta:

- a) A distribuição mais homogênea é a A com  $CV = 25\%$ , ou seja, a que tem o menor coeficiente de variabilidade.
- b) A distribuição mais heterogênea é a distribuição C com  $CV = 50\%$ , ou seja, a que tem o maior coeficiente de variabilidade.

Quando se eleva o coeficiente de variabilidade ao quadrado, obtém-se a variância relativa. Ela é obtida pela elevação ao quadrado do resultado do coeficiente de variabilidade ou dividindo a variância pelo quadrado da média.

### 5.5 Amplitude interquartílica

A amplitude interquartílica é uma medida de dispersão absoluta que representa a diferença entre o terceiro e o primeiro quartil da distribuição. A amplitude interquartílica é expressa na mesma unidade de medida da variável considerada.

$$AIQ = Q_3 - Q_1 \quad (5.20)$$

onde:

AIQ = valor da amplitude interquartílica

$Q_3$  = valor do terceiro quartil

$Q_1$  = valor do primeiro quartil

Essa medida representa a diferença existente entre os 50% dos dados centrais da amostra. Eles representam uma consistência maior dos dados.

Uma variante dessa medida é a amplitude semi-interquartílica ou desvio quartílico, que representa a média da diferença entre os valores do terceiro quartil e o primeiro quartil, ou seja, é a metade do valor da amplitude interquartílica, conforme equação a seguir:

$$Q = \frac{AIQ}{2} = \frac{Q_3 - Q_1}{2} \quad (5.21)$$

onde:

Q = valor da amplitude semi-interquartílica

AIQ = valor da amplitude interquartílica

$Q_3$  = valor do terceiro quartil

$Q_1$  = valor do primeiro quartil

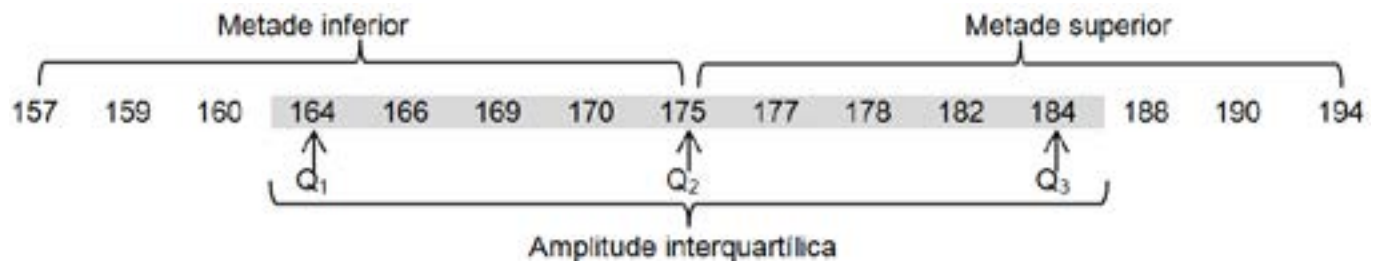
Essas medidas permitem analisar a amplitude total da dispersão eliminados os casos extremos.

Exemplo:

A altura (em centímetros) de uma amostra dos estudantes da Universidade XYZ foi: {157, 159, 160, 164, 166, 169, 170, 175, 177, 178, 182, 184, 188, 190, 194}. A partir dela, determine a amplitude interquartílica.

Para facilitar o entendimento da amplitude interquartílica é apresentada a Figura 5.5, para ilustrar o problema com os dados apresentados:

Figura 5.5 – Figura ilustrativa da amplitude interquartílica



Fonte: Dados fictícios, apenas para fins ilustrativos.

a) Amplitude interquartílica

$$AIQ = 184 - 164 = 20 \text{ cm}$$

Resposta:

A altura dos 50% dos alunos centrais da amostra apresenta uma amplitude de variação de 20cm.

b) Amplitude semi-interquartílica

$$Q = \frac{184 - 164}{2} = \frac{20}{2} = 10 \text{ cm}$$

Resposta:

Sabendo que o segundo quartil 2 (ou mediana) é igual a 175, pode-se afirmar que 50% dos estudantes têm alturas compreendidas entre 165 e 185, ou seja,  $175 \pm 10 \text{ cm}$ .



## 5.6 Resumo para fixação

Para auxiliar, como resumo para fixação dos conteúdos abordados no capítulo sugerem-se os seguintes aspectos:

- As medidas de dispersão ou variabilidade indicam quão dispersos estão os valores da amostra em torno da medida de centralidade, em termos absolutos e relativos.
- Quando o dado está agrupado, a ponderação deve ocorrer pela sua frequência.
- O desvio médio trata-se da média aritmética dos desvios dos dados observados em relação à média aritmética, considerados em módulos (valor absoluto).
- A variância representa a média aritmética dos quadrados dos desvios em relação à média aritmética. Em consequência, seu resultado é expresso em unidades de medida ao quadrado. Por exemplo, se a variável for medida em metros, a sua variância será expressa em metros quadrados.
- A variância amostral é obtida dividindo-se o somatório dos quadrados dos desvios em relação à média por  $n-1$ , o qual indica os graus de liberdade da amostra. Esse procedimento faz-se necessário para evitar que a projeção da população a partir da amostra seja não tendenciosa.
- O desvio-padrão de um conjunto de dados é definido como uma medida da variação dos valores em torno da média. É obtido extraíndo-se a raiz quadrada da variância.
- A medida de dispersão mais importante e utilizada é o desvio-padrão, pois pela sua natureza expressa valores maiores que o desvio médio.
- O coeficiente de variabilidade indica a magnitude relativa do desvio-padrão quando comparado com a média da distribuição das medidas. É utilizado para comparar diferentes variáveis.
- O coeficiente de variabilidade elevado ao quadrado é conhecido como variância relativa.
- A amplitude interquartílica é uma medida de dispersão absoluta que identifica os 50% dos dados centrais, sendo obtida pela diferença entre o terceiro e o primeiro quartil da distribuição. É expressa na mesma unidade de medida da variável considerada.

## 5.7 Exercícios propostos

5.1 A produção mensal de manteiga em toneladas, nos últimos seis meses, em uma indústria de laticínios foi de: {11, 8, 4, 10, 9, 12}. A partir dessas informações calcule:

- a) a média;
- b) o desvio médio;
- c) o desvio-padrão.

5.2 As notas de um determinado aluno de Administração foram iguais a {6, 3, 7, 4}. Pergunta-se, qual foi o desvio médio?

5.3 Se a média diária de consumo de calorias por um indivíduo é igual a 1.900 calorias, com um desvio-padrão de 350 calorias, estime o coeficiente de variação desses dados.

5.4 As notas finais de uma amostra dos alunos nos trabalhos de Ciências Contábeis foram: {7, 5, 8, 9, 5, 10}. A partir desses dados calcule:

- a) média aritmética e a média harmônica;
- b) amplitude amostral;
- c) desvio médio;
- d) variância absoluta;
- e) coeficiente de variabilidade.

5.5 Um grupo de 85 alunos tem estatura média de 176,4 cm, com desvio-padrão igual a 6,3 cm. Outro grupo de 125 alunos tem uma estatura média de 171,5cm, sendo o desvio-padrão igual a 4,9 cm. Qual é o grupo mais homogêneo?

5.6 A partir da Tabela 5.5, responda:

- a) Qual é o salário médio pago aos funcionários?
- b) De quanto é o desvio-padrão dos salários? E a variância?
- c) Qual é o desvio médio dos salários?
- d) Qual é o coeficiente de variabilidade?

Tabela 5.5 – Distribuição de frequência dos salários recebidos na empresa “Bom Negócio”

Classe (salários mínimos)	Nº de funcionários (frequência)
0  — 2	5
2  — 4	10
4  — 6	7
6  — 8	2
8  — 10	1

Fonte: Dados fictícios, apenas para fins ilustrativos.

5.7 A distribuição de frequências das idades dos funcionários da empresa “Bom Negócio”, em 01/01/2014 está descrita na Tabela 5.6

Tabela 5.6 – Distribuição de frequência da idade dos funcionários da empresa “Bom Negócio”

Idade (anos)	Número de funcionários
20  — 25	6
25  — 30	13
30  — 35	7
35  — 40	3
40  — 45	1

Fonte: Dados fictícios, apenas para fins ilustrativos.

A partir da tabela, responda:

- Qual é a idade média dos funcionários?
- Qual é o desvio médio?
- Qual é o desvio-padrão?
- Qual é a amplitude interquartílica?

## 6 MEDIDAS DE ASSIMETRIA E CURTOSE

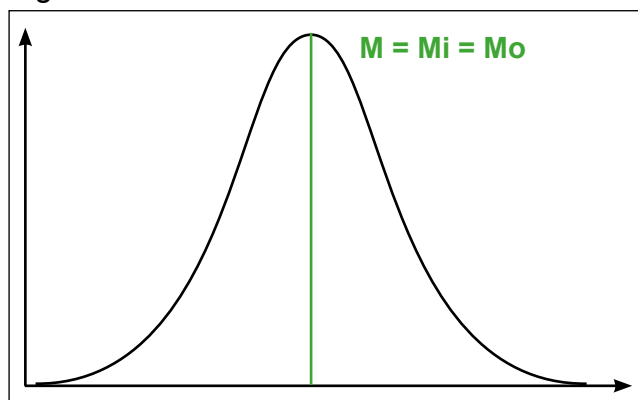
As medidas de assimetria e curtose, assim como as medidas de tendência central e as medidas de dispersão, auxiliam para a descrição e compreensão de distribuição de frequência e são úteis para comparar a forma das distribuições das variáveis em análise.

### 6.1 Medidas de assimetria

As medidas de assimetria são expressas em coeficientes, que permitem analisar a forma da distribuição de frequência dos dados, calculados com base nas relações entre as medidas de moda, média, mediana e demais separatrizes e o desvio-padrão. Assim, pelas medidas de assimetria ou também de forma gráfica, pode-se concluir se a distribuição é simétrica, assimétrica à direita ou assimétrica à esquerda (FIGURAS 6.1; 6.2 e 6.3).

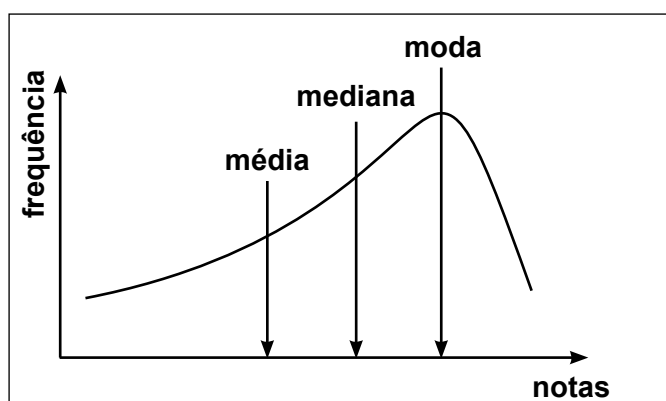
Em uma distribuição simétrica, coincidem a média, a moda e a mediana e os quartis ficam equidistantes da média. Na distribuição assimétrica não ocorre isso. A distribuição é assimétrica à direita quando o valor da moda é menor do que os respectivos valores da mediana e da média. Isso significa que ocorre a uma maior repetição de valores no grupo dos 50% de valores menores. Quando a distribuição é assimétrica à esquerda, ocorre o inverso, sendo o valor da moda maior que os respectivos valores da mediana e da média.

Figura 6.1 – Curva simétrica



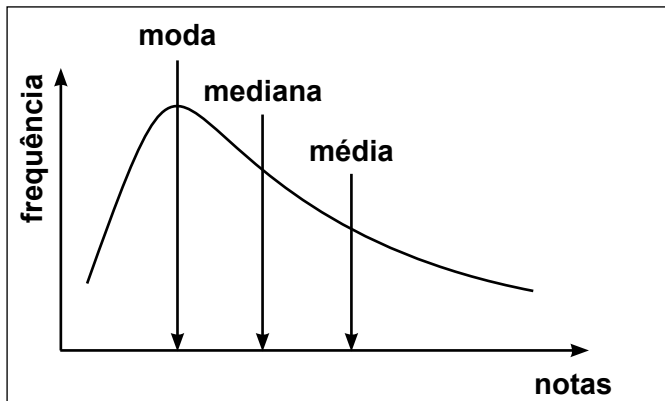
Fonte: Dos autores.

Figura 6.2 – Curva assimétrica à esquerda



Fonte: Dos autores.

Figura 6.3 – Curva assimétrica à direita



Fonte: Dos autores.

Vários autores tentaram definir regras para avaliar a assimetria das distribuições, estabelecendo assim coeficientes de assimetria.

### 6.1.1 Coeficiente de Assimetria

O Coeficiente de Assimetria permite comparar duas ou mais distribuições diferentes e avaliar qual delas é mais assimétrica. Quanto maior for o Coeficiente de Assimetria mais assimétrica é a curva.

Dependendo das informações de que se dispõe, pode-se calcular de quatro formas o Coeficiente de Assimetria: coeficiente de Pearson “A”, coeficiente de Pearson “B”, coeficiente de Bowley e coeficiente de Kelley.

Quando se dispõe de valores da média e do desvio-padrão, pode-se calcular o Coeficiente de Assimetria de duas formas:

1ª forma: Coeficiente de Pearson “A”

O coeficiente de Assimetria de Pearson “A” mede o afastamento da simetria expressando a diferença entre a média e a mediana em relação ao desvio-padrão do grupo de medidas.

Para o cálculo desse coeficiente, usa-se a seguinte fórmula:

$$A_{Med} = \frac{3 (\bar{X} - Med)}{s} \quad (6.1)$$

onde:

$A_{Med}$  = Coeficiente de Assimetria de Pearson “A”

Med = Mediana

$\bar{X}$  = Média aritmética da amostra

s = Desvio-padrão da amostra

Após o cálculo do coeficiente deve-se analisar o resultado encontrado, considerando a seguinte convenção:

$A_{Med} = 0$ ; a distribuição é simétrica.

$A_{Med} > 0$ ; a distribuição é assimétrica positiva.

$A_{Med} < 0$ ; a distribuição é assimétrica negativa.

Para uma distribuição simétrica, o coeficiente de assimetria será sempre zero, pois, nesse caso, o valor da média será igual ao valor da mediana, o que anula a expressão. Para uma distribuição positivamente assimétrica ou assimétrica à direita, a média é sempre maior que a mediana, em consequência o coeficiente será positivo. Para uma distribuição negativamente

assimétrica ou assimétrica à esquerda, a média é sempre menor que a mediana, em consequência o coeficiente será negativo.

Exemplo:

Sendo o salário médio de uma empresa igual a R\$ 1.194,50, a mediana de R\$ 1.134,40 e o desvio-padrão de R\$ 124,35. O coeficiente de assimetria será:

$$A_{med} = \frac{3(1.194,50 - 1.134,40)}{124,35} = + 1,45$$

Resposta: A curva é positivamente assimétrica ou assimétrica à direita.

## 2ª forma: Coeficiente de Pearson “B”

O Coeficiente de Pearson “B” mede o afastamento da simetria expressando a diferença entre a média e a moda em relação ao desvio-padrão do grupo de medidas.

O resultado desse coeficiente é obtido com o uso da seguinte fórmula:

$$A_{Mod} = \frac{\bar{X} - Mod}{s} \quad (6.2)$$

onde:

$A_{Mod}$  = Coeficiente de Assimetria de Pearson “B”

Mod = Moda

$\bar{X}$  = Média aritmética da amostra

s = Desvio-padrão da amostra

Da mesma forma que na solução anterior, quando:

$A_{mod} = 0$ ; a distribuição é simétrica.

$A_{mod} > 0$ ; a distribuição é assimétrica positiva.

$A_{mod} < 0$ ; a distribuição é assimétrica negativa.

Quando não se dispõe dos valores calculados da média e do desvio-padrão, mas têm-se os dados dos quartis, utiliza-se a forma a seguir.

## 3ª forma: Coeficiente de Assimetria de Bowley ou Coeficiente Quartílico de Assimetria

Este coeficiente mede o afastamento da simetria relacionando o 1º Quartil e o 3º Quartil com o 2º Quartil (ou a mediana) do grupo de medidas.

O resultado desse coeficiente é obtido com o uso da seguinte fórmula:

$$A_Q = \frac{Q_3 + Q_1 - 2 Med}{Q_3 - Q_1} \quad (6.3)$$

onde:

$A_Q$  = Coeficiente de Assimetria de Bowley

Med = Mediana

$Q_3$  = Valor do terceiro quartil

$Q_1$  = Valor do primeiro quartil

Quando a distribuição for simétrica, então  $Q_1$  e  $Q_3$  serão equidistantes de  $Q_2$  ou da mediana, o que vale dizer que a soma de  $Q_1$  e  $Q_3$  é igual a 2 vezes  $Q_2$  (ou mediana). Dessa forma, quando a distribuição for simétrica, a relação se anula, ou seja, o resultado é igual a zero.

Quando a distribuição é assimétrica, então  $Q_1$  e  $Q_3$  não são equidistantes de  $Q_2$ , o que resultará em um valor diferente de zero na expressão. Se a distância de  $Q_3$  em relação a  $Q_2$  for

maior que a distância de Q1 em relação a Q2, ter-se-á um resultado positivo na expressão, o que indica uma assimetria à direita, e, se for o inverso, haverá uma assimetria à esquerda.

Numa distribuição simétrica é válida a relação:

$$Q_2 = \frac{Q_1 + Q_3}{2}$$

onde:

$Q_2$  = Valor do segundo quartil (ou mediana)

$Q_3$  = Valor do terceiro quartil

$Q_1$  = Valor do primeiro quartil

Numa distribuição assimétrica a relação válida é:

$$Q_2 \neq \frac{Q_1 + Q_3}{2}$$

Assim sendo, valem as conclusões das formas anteriores, ou seja:

$A_Q = 0$ ; a distribuição é simétrica.

$A_Q > 0$ ; a distribuição é assimétrica positiva.

$A_Q < 0$ ; a distribuição é assimétrica negativa.

#### 4ª forma: Coeficiente de Assimetria de Kelley ou Coeficiente de Assimetria dos Centis

Esta forma é uma variante da Assimetria de Bowley, substituindo os Quartis por Centis, relacionando o 90º Centil e o 10º Centil com o 50º Centil (ou mediana) do grupo de medidas.

Para obter o resultado desse coeficiente usa-se a seguinte fórmula:

$$A_C = \frac{C_{90} + C_{10} - 2 \text{ Med}}{C_{90} - C_{10}} \quad (6.4)$$

onde:

$A_C$  = Coeficiente de Assimetria de Kelley

Med = Mediana

$C_{90}$  = Valor do nonagésimo centil

$C_{10}$  = Valor do décimo centil

As conclusões são idênticas às demais formas de análise de assimetria das distribuições.

Assim, também valem as conclusões das formas anteriores, ou seja:

$A_C = 0$ ; a distribuição é simétrica.

$A_C > 0$ ; a distribuição é assimétrica positiva.

$A_C < 0$ ; a distribuição é assimétrica negativa.

Como os pontos de análise da curva não são os mesmos para as quatro fórmulas apresentadas, os resultados entre si, dependendo do grau de desuniformidade da distribuição, podem variar. Assim, fórmulas de Pearson, "A" e "B", que se baseiam nas medidas centrais da distribuição, são mais seguras e devem ser usadas prioritariamente. O Coeficiente de Kelly, que leva em conta dados mais extremos da curva, décimo e nonagésimo centis, é o mais vulnerável em termos de sensibilidade da curva.

#### Exemplos

Utilizando a Tabela 6.1, determine o coeficiente de assimetria por meio das quatro fórmulas apresentadas:

Tabela 6.1 – Valores recebidos por hora por moradores de um bairro “A”, no ano de 2011

Salários por hora	Nº de pessoas	$x_i$	$F_i$
50  — 60	15	55	15
60  — 70	20	65	35
70  — 80	30	75	65
80  — 90	20	85	85
90  — 100	15	95	100
Total	100		-----

Fonte: Dados fictícios, apenas para fins ilustrativos.

Solução:

1 - Coeficiente de Pearson “A”

Para encontrar a solução, precisa-se calcular a média, a mediana e o desvio-padrão, para o que será usada a Tabela 6.1 b com os dados de cálculos.

$$A_{Med} = \frac{3 (\bar{X} - Med)}{s}$$

Tabela 6.1 b – Valores recebidos por hora por pessoas de um bairro “A”, no ano de 2011

Salários por hora	Nº de pessoas	$x_i$	$F_i$	$x_i * f_i$	$x_i^2$	$x_i^2 * f_i$
50  — 60	15	55	15	825	3.025	45.375
60  — 70	20	65	35	1300	4.225	84.500
70  — 80	30	75	65	2250	5.625	168.750
80  — 90	20	85	85	1700	7.225	144.500
90  — 100	15	95	100	1425	9.025	135.375
Total	100	-	-	7500	-	578.500

Fonte: Dados fictícios, apenas para fins ilustrativos.

Onde:

$x_i$  – ponto médio do intervalo de frequência

$F_i$  – frequência do intervalo

Com base nos dados da Tabela 6.2 e utilizando as fórmulas de cálculo da média, da mediana e do desvio-padrão, obtêm-se os seguintes resultados:

$$\bar{X} = \frac{\sum X_i f_i}{n} = \frac{7500}{100} = 75$$

$$Med = li_k + A_k \frac{\frac{n}{2} - F_{k-1}}{f_k} \quad Med = 70 + 10 \frac{50 - 35}{99} = 75$$

$$s^2 = \frac{578500 - \frac{(7500)^2}{100}}{100 - 1} = \frac{578500 - 562500}{99} =$$

$$s^2 = \frac{16000}{99} = 161,61616$$

$$s = 12,71$$

$$A_{Med} = \frac{3 (75 - 75)}{12,71} = 0$$

Resposta:

Como o valor calculado do coeficiente de Pearson “A” é igual a zero, a distribuição é simétrica.



## 2 - Coeficiente de Pearson “B”

Para encontrar a solução, precisa-se calcular a média, a moda e o desvio-padrão. Para isso, serão usados os dados da Tabela 6.1 b.

$$A_{Mod} = \frac{\bar{X} - Mod}{s}$$

Com base nos dados da Tabela 6.1.b e utilizando-se os valores já calculados da média e do desvio-padrão, precisa-se ainda calcular a moda por meio da fórmula a seguir, obtendo os seguintes resultados:

$$Mod = li_k + A_k \frac{\Delta_1}{\Delta_1 + \Delta_2}$$

A moda encontra-se no intervalo 70 |— 80, que tem a maior frequência.

Onde:

$$\Delta_1 = 30 - 20 = 10$$

$$\Delta_2 = 30 - 20 = 10$$

$$li_k = 70$$

$$A_k = 10$$

Assim,

$$Mod = 70 + 10 \frac{10}{10 + 10} = 70 + 5 = 75$$

E, substituindo a fórmula do coeficiente

$$A_{Mod} = \frac{\bar{X} - Mod}{s} = \frac{75 - 75}{12,71} = 0$$

Resposta:

Da mesma forma como no cálculo da fórmula anterior, o coeficiente de Pearson “B” é igual a zero, portanto a distribuição é simétrica.

## 3 - Coeficiente de Assimetria de Bowley

$$A_Q = \frac{Q_3 + Q_1 - 2 Med}{Q_3 - Q_1}$$

Para calcular esse coeficiente, precisa-se dos valores do terceiro quartil ( $Q_3$ ), primeiro quartil ( $Q_1$ ) e da Mediana. Para isso, usam-se as respectivas fórmulas:

$$Q_3 = li_k + A_k \frac{\frac{3n}{4} - F_{k-1}}{f_k} \quad Q_1 = li_k + A_k \frac{\frac{n}{4} - F_{k-1}}{f_k}$$

$$Med = li_k + A_k \frac{\frac{n}{2} - F_{k-1}}{f_k}$$

Utilizando-se os valores da Tabela 6.1 b, obtêm-se os seguintes resultados:

$$Q_3 = 80 + 10 \frac{75 - 65}{20} = 80 + 5 = 85$$

$$Q_1 = 60 + 10 \frac{25 - 15}{20} = 60 + 5 = 65$$

$$Med = 70 + 10 \frac{50 - 35}{30} = 70 + 5 = 75$$

$$A_Q = \frac{85 + 65 - 2(75)}{85 - 65} = \frac{150 - 150}{20} = 0$$

Resposta:

Confirma-se o resultado anterior, calculado agora pelo coeficiente de Bowley.

#### 4 - Coeficiente de Assimetria de Kelley

$$A_Q = \frac{C_{90} + C_{10} - 2 Med}{C_{90} - C_{10}}$$

Para calcular este coeficiente, precisa-se dos valores centil 90 (C90), centil 10 (C10) e Mediana, utilizando as seguintes fórmulas:

$$C_{90} = li_k + A_k \frac{\frac{90n}{100} - F_{k-1}}{f_k} \quad C_{10} = li_k + A_k \frac{\frac{10n}{100} - F_{k-1}}{f_k}$$

Utilizando-se os valores da Tabela 6.1 b e os valores das medidas calculadas anteriormente, encontram-se os seguintes valores para as medidas necessárias e o coeficiente de Kelley:

$$C_{90} = 90 + 10 \frac{90 - 85}{15} = 90 + 3,33 = 93,33$$

$$C_{10} = 50 + 10 \frac{10 - 0}{15} = 50 + 6,67 = 56,67$$

$$A_C = \frac{93,33 + 56,67 - 2(75)}{93,33 - 56,67} = \frac{150 - 150}{36,66} = 0$$

Resposta:

Da mesma forma, confirma-se o resultado calculado pelos outros métodos.

## 6.2 Medidas de curtose

Entende-se por curtose o grau de achatamento de uma distribuição, em comparação com uma distribuição padrão. Com referência ao grau de achatamento de uma curva, pode-se ter três tipos de curva: mesocúrtica, leptocúrtica e platicúrtica.

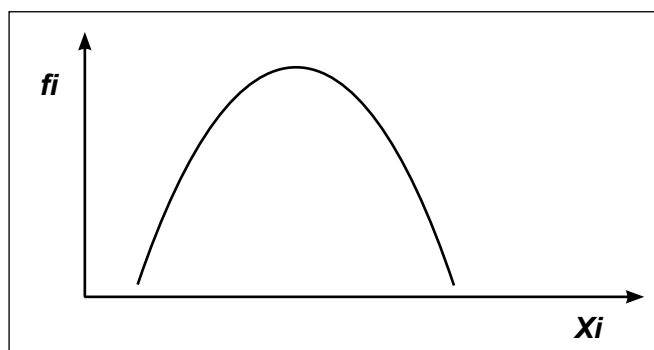
A curva **mesocúrtica** é a curva básica que apresenta um grau de achatamento padrão equivalente ao da curva normal.

Já a curva **leptocúrtica** apresenta um alto grau de afilamento, superior ao normal, sendo uma curva mais fechada ou empinada.

Por outro lado, a curva **platicúrtica** apresenta um alto grau de achatamento, superior ao normal, sendo mais aberta.

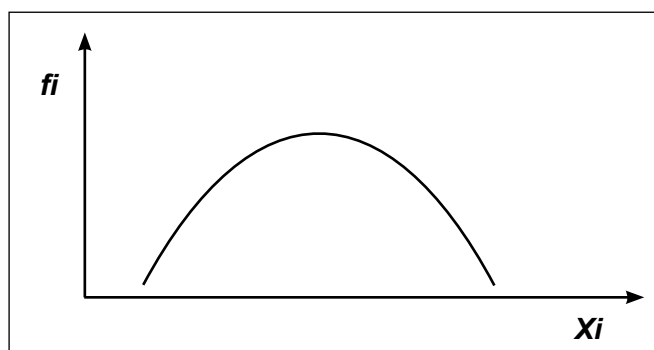
As três possibilidades estão ilustradas graficamente Figuras 6.4, 6.5 e 6.6.

Figura 6.4- Curva leptocúrtica



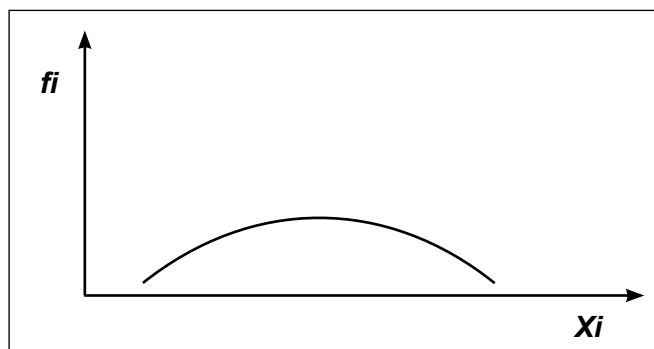
Fonte: Dos autores

Figura 6.5 - Curva mesocúrtica



Fonte: Dos autores

Figura 6.6 - Curva platicúrtica



Fonte: Dos autores

Para medir o grau de curtose utiliza-se a seguinte fórmula para calcular o coeficiente de Curtose:

$$K = \frac{Q_3 - Q_1}{2 (C_{90} - C_{10})} \quad (6.5)$$

onde:

$C_{90}$  = 90° Centil

$C_{10}$  = 10° Centil

$Q_3$  = 3° Quartil

$Q_1$  = 1° Quartil

Com o valor calculado pela fórmula de Assimetria acima, conclui-se sobre a realidade de curtose de seguinte forma:

Se  $K = 0,263$ , a curva corresponde a uma distribuição de frequência **mesocúrtica**;

Se  $K > 0,263$ , a curva corresponde a uma distribuição de frequência **platicúrtica**;

Se  $K < 0,263$ , a curva corresponde a uma distribuição de frequência **leptocúrtica**.

Exemplo:

Conclua a respeito do tipo de curva da distribuição da Tabela 6.2 quanto à curtose.

Tabela 6.2 – Distribuição de dados de uma variável “B” de determinada época

Intervalo de classe	Frequência	$F_i$
3  — 8	5	5
8  — 13	15	20
13  — 18	20	40
18  — 23	10	50

Fonte: Dados fictícios, apenas para fins ilustrativos.

Para o cálculo do Coeficiente K, necessita-se calcular  $Q_1$ ,  $Q_3$ ,  $C_{10}$  e  $C_{90}$ :

$$Q_1 = li_k + A_k \frac{\frac{n}{4} - F_{k-1}}{f_k}$$

$$Q_1 = 8 + 5 \frac{12,5 - 5}{15} = 8 + 2,5 = 10,5$$

$$Q_3 = li_k + A_k \frac{\frac{3n}{4} - F_{k-1}}{f_k}$$

$$Q_3 = 13 + 5 \frac{37,5 - 20}{20} = 13 + 4,38 = 17,38$$

$$C_{10} = li_k + A_k \frac{\frac{10n}{100} - F_{k-1}}{f_k}$$

$$C_{10} = 3 + 5 \frac{5 - 0}{5} = 3 + 5 = 8$$

$$C_{90} = li_k + A_k \frac{\frac{90n}{100} - F_{k-1}}{f_k}$$

$$C_{90} = 18 + 5 \frac{45 - 40}{10} = 18 + 2,5 = 20,5$$

$$K = \frac{Q_3 - Q_1}{2(C_{90} - C_{10})} = \frac{17,38 - 10,5}{2(20,5 - 8)} = \frac{6,88}{25} = 0,27$$

Resposta:

Como  $K > 0,263$ , logo, a curva correspondente é suavemente **platicúrtica**.

### 6.3 Resumo para fixação

Para fixar os principais aspectos abordados no capítulo, são apresentados os seguintes pontos como resumo:

- A assimetria analisa a curva de frequência de forma horizontal, relacionando a sua característica com a configuração de uma distribuição simétrica.
- Na distribuição simétrica, coincidem os valores da média, moda e a mediana, sendo o valor dos quartis equidistantes da mediana.
- A distribuição assimétrica à direita (ou positiva) apresenta a moda com menor valor, a média com valor maior e a mediana com valor intermediário.
- Já a distribuição assimétrica à esquerda (ou negativa) tem como valor maior a moda, menor a média e intermediário a mediana.
- O grau da assimetria é dado pelos Coeficientes de Assimetria, que quanto maiores forem os seus valores maior será o grau da assimetria.
- No cálculo dos coeficientes de assimetria deve-se dar preferência aos de Pearson A e B, que são baseados nas medidas centrais da distribuição, em que a curva é menos irregular, sendo os seus valores mais seguros.
- A curtose analisa a curva de frequência de forma vertical, relacionando a sua característica com a configuração de uma distribuição mesocúrtica, característica de uma Distribuição Normal.
- As medidas de assimetria e curtose são medidas independentes e que não se influenciam mutuamente.
- As medidas de curtose exigem dados referentes aos quartis e centis, enquanto que as medidas de assimetria podem ser definidas por valores da média, moda, mediana e desvio-padrão.
- Para medir a semelhança da curva de distribuição com a característica de uma curva mesocúrtica, usa-se o coeficiente de curtose.
- A curva mesocúrtica tem como padrão de curtose o valor de 0,263.
- Coeficiente de curtose com valor ABAIXO de 0,263 indica que a curva é mais afilada (empinada) que uma curva mesocúrtica, sendo chamada de leptocúrtica.
- Coeficiente de curtose com valor ACIMA de 0,263 indica que a curva é mais achatada que uma curva mesocúrtica, sendo chamada de platicúrtica.

## 6.4 Exercícios propostos

6.1 Considerando a afirmação, assinale a resposta correta. Uma curva simétrica se caracteriza pelo seguinte atributo:

- a) a curva é assimétrica à esquerda.
- b) o valor da moda é maior que os da mediana e da média.
- c) as medidas moda, mediana e média são iguais.
- d) o valor do desvio-padrão é maior que os valores da mediana e da moda.
- e) os valores dos decis são equivalentes ao valor da média

6.2 Dados os resultados:  $\text{Mod} = 30$ ,  $\text{Med} = 28$  e  $\bar{X} = 22$ , pode-se afirmar que a curva de frequência é (assinale a resposta correta)

- a) mesocúrtica.
- b) simétrica.
- c) assimétrica negativa.
- d) assimétrica positiva.
- e) assimétrica leptocúrtica.

6.3 O coeficiente de curtose para uma determinada distribuição de frequências apresentou o seguinte resultado:  $k = 0,278$ . Nesse caso pode-se afirmar que a curva é:

- a) assimétrica positiva.
- b) leptocúrtica.
- c) mesocúrtica.
- d) simétrica.
- e) platicúrtica.

6.4 O coeficiente de curtose para uma determinada distribuição de frequências apresentou o seguinte resultado:  $k = 0,358$ . Nesse caso pode-se afirmar que a curva é:

- a) assimétrica positiva.
- b) leptocúrtica.
- c) mesocúrtica.
- d) simétrica.
- e) platicúrtica.

6.5 Um estudante de estatística acompanhou no mês de agosto de 2014 o tempo gasto em horas em uma tarefa na empresa Testando Sempre. Os dados coletados são os seguintes: 14; 3; 27; 4; 6; 7; 8; 1; 16; 4.

- a) Qual é o tipo de distribuição em termos de assimetria?
- b) Qual é o tipo de distribuição em termos de curtose?

6.6 Na Tabela 6.3 são apresentadas as notas obtidas pelos alunos de estatística.

Tabela 6.3 – Notas obtidas por alunos de Estatística na instituição “A” em junho de 2014

Notas	Número de alunos
0  — 2	4
2  — 4	12
4  — 6	15
6  — 8	13
8  — 10	6

Fonte: Dados fictícios, apenas para fins ilustrativos.

- Qual é o tipo de distribuição em termos de assimetria?
- Qual é o tipo de distribuição em termos de curtose?

## 7 TEORIA ELEMENTAR DE PROBABILIDADE

### 7.1 Introdução

O termo probabilidade é usado de modo muito amplo na conversação diária. A sua teoria pertence ao campo da matemática pura. As raízes da teoria de probabilidade se encontram nos jogos de azar, muito em voga no século XVII, em especial entre a nobreza da época. Assim, a partir de 1700, os progressos no campo da Estatística tornaram-se importantes.

Em 1713, James Bernoulli publica a *Ars Conjectandi*. A importância dessa obra deve-se ao fato de que ela estabelece a base da relação entre probabilidade e frequência relativa.

Em 1718 surge a primeira edição de *The Doctrine of Changes*, obra de Moivre, abrindo os horizontes do cálculo de probabilidades.

Em 1812, Laplace publica *Theorie Analytique des Probabilités*, com várias aplicações do cálculo de probabilidade.

A probabilidade sugere certo grau de incerteza sobre o que aconteceu no passado, o que ocorrerá no futuro ou o que está ocorrendo no presente (FIGURA 7.1). Sua compreensão é absolutamente necessária para o desenvolvimento da estatística indutiva ou inferência estatística.

A teoria da probabilidade nasceu por meio dos matemáticos, na busca de alternativas de jogadas nos jogos de azar.

Figura 7.1 – Utilização da probabilidade



Fonte: Jornal Zero Hora, s.d.

Assim, tem aplicação na Administração de Empresas, Economia, Psicologia, Biologia e outros ramos da Ciência. Se, por exemplo, um empresário deseja lançar um produto novo no mercado, ele necessita conhecer a probabilidade de sucesso desse produto.

### 7.2 Conceitos

O conceito de probabilidade depende do ponto de vista de análise e, com base na solução matemática adotada, pode ser classificado em três grupos: conceito clássico, conceito moderno e conceito subjetivo.

#### 7.2.1 Clássico ou conceito de Laplace

É o conceito mais primitivo e sua definição parte do princípio de que todos os eventos de um experimento têm iguais possibilidades. Assim, pode-se enunciar este conceito como: a probabilidade é o quociente entre o número de casos favoráveis do evento com o número de casos possíveis, desde que todos os eventos sejam equiprováveis.



Esse conceito se aplica quando os experimentos podem ser repetidos sob as mesmas condições e permite determinar matematicamente o valor esperado para cada evento.

Os conceitos aplicam-se aos experimentos, como, por exemplo:

- retirada de uma carta de um baralho com reposição;
- jogo de xadrez;
- retirada de bolas de um balaio com reposição;
- sorteio de um número com reposição.

### 7.2.2 Conceito moderno

Modernamente, surgiu o conceito estatístico de probabilidade que considera as repetições de cada evento em um grande número de tentativas e mede as frequências relativas da ocorrência dos eventos. Esse conceito não parte do princípio de que os eventos têm a mesma possibilidade de ocorrerem, mas relata as ocorrências de cada evento.

Numa relação entre os conceitos clássico e moderno aplicados a um mesmo experimento, podemos considerar que a probabilidade é o limite para o qual tende a frequência relativa quando o número de observações direciona-se para o infinito. O conceito moderno, quando utilizado numa amostragem, na medida em que se amplia a quantidade de amostras, a tendência é de que se aproxime do conceito clássico, pois o conceito clássico origina-se da ideia de que o experimento tende ao infinito.

Tanto o enfoque “clássico” como o “moderno” geram valores objetivos de probabilidade, no sentido de que tais valores indicam a taxa relativa de ocorrência do evento no longo prazo.

Exemplos ilustrativos são alguns jogadores de moeda, apresentados a seguir.

O conde Buffon (1707 – 1788), naturalista francês, jogou uma moeda 4.040 vezes. O resultado foi 2.048 caras, ou uma proporção de 0,5069 caras.

Por volta de 1900, o estatístico inglês Karl Pearson heroicamente lançou uma moeda 24.000 vezes, obtendo 12.012 caras, ou uma proporção de 0,5005.

Enquanto foi prisioneiro dos alemães na Segunda Guerra Mundial, o matemático sul-africano John Kerrich lançou uma moeda 10.000 vezes, obtendo 5.067 caras, ou uma proporção de 0,5067.

### 7.2.3 Conceito subjetivo

Neste conceito, a definição da medida matemática de cada evento depende do grau de certeza de cada observador. Dessa forma, podemos considerar como conceito que a probabilidade de certo evento é medida pelo grau de crença que cada pessoa atribui à ocorrência do evento.

Com essa conceituação, temos diferentes probabilidades para o mesmo evento. Esse enfoque tem sido desenvolvido recentemente por Bayes e está relacionado com a análise Bayesiana de decisão.

Nesta publicação não será abordado esse conceito, mas a seguir é apresentado um exemplo da aplicação da análise Bayesiana, que é o Teorema de Bayes.

## 7.3 Conceitos básicos

### 7.3.1 Experimento aleatório

Experimento aleatório é todo o fenômeno ou ação que geralmente pode ser repetido sob as mesmas condições indefinidamente e cujo resultado é casual ou aleatório. Como símbolo será usada a letra “E”.

Um exemplo de experimento aleatório é jogar um dado. Como o dado tem seis faces, esse experimento terá como resultado um dos valores:  $\{1, 2, 3, 4, 5, 6\}$ .

Além desse, há outros experimentos, como:

- retirar uma carta de um baralho;
- retirar bolas de uma urna;
- contar o número de peças defeituosas da produção diária de uma máquina;
- jogar uma moeda;
- sortear um aluno.

### 7.3.2 Espaço amostral

O conjunto de todos os resultados possíveis de um experimento aleatório é chamado de espaço amostral e é indicado pelo símbolo “S”.

Exemplo:

No lançamento de um dado tem-se o seguinte espaço amostral:

$$S = \{1, 2, 3, 4, 5, 6\}.$$

### 7.3.3 Evento

Os itens que formam o conjunto de resultados de um experimento são denominados de eventos. Os eventos são indicados por letras maiúsculas: A, B, C etc.

Todos os eventos de um experimento são chamados de casos possíveis, e os analisados são considerados casos favoráveis (que interessam) ao pesquisador.

Exemplo:

- Experimento:  $E = \text{lançar um dado}$
- Espaço amostral:  $S = \{1, 2, 3, 4, 5, 6\}$
- Evento A: Sair um número par:  $A = \{2, 4, 6\}$

Os eventos são classificados em duas categorias: evento simples e evento composto.

Evento simples é qualquer subconjunto do espaço amostral unitário.

Exemplo:

Evento A: Sair o número 2 no lançamento de um dado.

Evento composto é qualquer subconjunto do espaço amostral formado por mais de um elemento.

Exemplo:

Evento A: Sair um número par no lançamento de um dado. Nesse caso existem três possibilidades que satisfazem o evento A: 2, 4 e 6.

Quanto à forma de realização dos eventos, é possível classificá-las em duas categorias: simultaneidade da ocorrência e independência de ocorrência.

A simultaneidade da ocorrência pode ocorrer de duas formas: eventos mutuamente exclusivos ou eventos não mutuamente exclusivos.

A situação de eventos mutuamente exclusivos ocorre quando dois eventos, A e B, não puderem ocorrer simultaneamente, ou seja, quando um evento ocorrer, o outro não pode ocorrer.

Exemplo de ocorrência de eventos mutuamente exclusivos:

- Experimento: lançamento de um dado.
- Evento A: sair um número par.

- Evento B: sair um número ímpar.

Esses dois eventos não poderão ocorrer ao mesmo tempo. Por isso são denominados mutuamente exclusivos.

Os eventos não mutuamente exclusivos ocorrem quando um evento não interfere no outro, ou quando os dois podem ocorrer ao mesmo tempo. Como, por exemplo, sair o número 2 e um número par. Pode ocorrer de sair o número 2, que também é par e, assim, atender aos dois eventos, ou pode sair o número 4 e atender apenas a um dos eventos. Nos eventos não mutuamente exclusivos, deve-se ter o cuidado na avaliação das probabilidades, para não as superestimár.

Em termos de independência de ocorrência, os eventos podem ser independentes ou dependentes.

Eventos independentes apresentam como característica a independência em relação à realização de eventos, ou seja, a realização ou a não realização de um dos eventos não afeta a probabilidade da realização do outro ou vice-versa.

Exemplo de eventos independentes:

- Experimento: lançamento de dois dados.
- Evento A: sair o número 6 no primeiro dado.
- Evento B: sair o número 5 no segundo dado.

Os dois eventos são independentes, pois a realização de um evento não afeta o resultado do outro.

Assim, a ocorrência de eventos independentes está intrínseca à natureza de realização, como, por exemplo, a jogada de um dado ou resultante da reposição dos elementos, como no caso de retirada de uma carta do baralho.

No caso de eventos dependentes, a ocorrência de um evento está condicionada à ocorrência de outro, ou à realização do evento precedente. Essa situação ocorre, por exemplo, no sorteio de dois números, em que o segundo sorteio já não contará mais com o primeiro número sorteado.

## 7.4 Definições axiomáticas de probabilidade

Nos cálculos das probabilidades são utilizadas as seguintes expressões matemáticas:

- $P(A)$ : probabilidade de ocorrência do evento A;
- $P(B)$ : probabilidade de ocorrência do evento B;
- $P(A+B)$ : probabilidade de ocorrência de um evento A ou evento B (regra do “ou”), que podem ser eventos mutuamente exclusivos ou eventos não mutuamente exclusivos;
- $P(AB)$ : probabilidade de ocorrência dos eventos A e B simultaneamente (regra do “e”), que podem ser eventos dependentes ou eventos independentes;
- $P(A/B)$ : probabilidade de ocorrer A depois que tenha ocorrido o evento B (eventos dependentes);
- $P(B/A)$ : probabilidade de ocorrer B depois que tenha ocorrido o evento A (eventos dependentes).

## 7.5 Axiomas

Os axiomas expressam os valores resultantes das probabilidades. Existem sete axiomas que definem o cálculo da probabilidade elementar, que são os seguintes:

1º a probabilidade de um evento P qualquer é um número compreendido entre 0 e 1, com ambos os limites incluídos:

$$0 \leq P \leq 1$$

2º a probabilidade de um evento certo é igual a 1;

3º a probabilidade de um evento impossível é igual a 0;

4º a probabilidade de ocorrer um entre dois eventos é igual à soma das probabilidades de eles ocorrerem individualmente, menos a probabilidade de ambos ocorrerem simultaneamente:

$$P(A+B) = P(A) + P(B) - P(AB) \quad (7.1)$$

5º se  $\bar{A}$  é o evento complementar de A, então

$$P(\bar{A}) = 1 - P(A) \quad (7.2)$$

6º a probabilidade de dois eventos ocorrerem simultaneamente é igual ao produto da probabilidade de um evento pela probabilidade do outro, sob a condição de o primeiro já ter sido realizado.

$$P(AB) = \{P(A) \cdot P(B/A)\} \quad (7.3)$$

ou

$$P(BA) = \{P(B) \cdot P(A/B)\} \quad (7.4)$$

7º probabilidade condicional: a probabilidade condicional do evento A, tendo ocorrido o evento B, é indicada por  $P(A/B)$ , e definida pela relação dos elementos do evento A em relação aos elementos do evento B.

Exemplo 1:

- Experimento: retirada de uma carta de um baralho;
- Evento A: retirar um rei;
- Evento B: retirar uma carta de espada;
- A probabilidade do Evento A é igual a:  $4/52$ ;
- A probabilidade do Evento B depois que ocorreu A é igual a:  $1/4$ ;

Assim, num baralho de 52 cartas existem quatro reis e entre os quatro reis há uma carta de espada.

Exemplo 2:

- Experimento: sortear dois números sem reposição de um conjunto de 10;
- Evento A: Sortear o primeiro número;
- Evento B: Sortear o segundo número;
- A probabilidade do Evento A é igual a:  $1/10$ ;
- A probabilidade do Evento B depois que ocorreu A é igual a:  $1/9$ ;

## 7.6 Teoremas

1. Se dois eventos A e B são mutuamente exclusivos, a probabilidade de suceder ou um ou outro é igual à soma das probabilidades dos eventos ocorrerem separadamente:

$$P(A+B) = P(A) + P(B) - P(AB)$$

Em sendo  $P(AB) = 0$ , portanto tem-se:

$$P(A+B) = P(A) + P(B)$$

2. Dois eventos se dizem independentes se a probabilidade de ocorrer um dos eventos não depende da ocorrência do outro. Portanto, se dois eventos são independentes, a probabilidade de eles ocorrerem simultaneamente, é o produto das probabilidades de eles ocorrerem separadamente:

$P(A/B) = P(A)$ , pois o resultado do Evento A não depende do Evento B.

$P(B/A) = P(B)$ , pois o resultado do Evento B não depende do Evento A, assim, tem-se:

$$P(AB) = P(A) \cdot P(B)$$

## 7.7 Avaliação de probabilidades

Para avaliar a probabilidade dos eventos, admite-se que todos os elementos do espaço amostral tenham a mesma chance (baseado no conceito clássico de probabilidade), ou seja, os resultados são igualmente prováveis. Isso significa que, se “n” for o número de elementos de S (espaço amostral), a probabilidade de cada evento simples será dada por  $1/n$ .

Exemplo:

- Experimento = lançamento de um dado;
- Espaço amostral (S) = {1, 2, 3, 4, 5, 6}, portanto  $n=6$ ;
- Evento A = sair o número 3, sendo portanto um caso favorável;
- $P(A) = 1/6$ .

Assim sendo, a probabilidade do evento A é dada pelo quociente entre o número de casos favoráveis ao evento A e o número total de casos possíveis:

$$P(A) = \frac{\text{Número de casos favoráveis}}{\text{Número de casos possíveis } (n)} \quad (7.5)$$

Exemplo:

1. Qual é a probabilidade de aparecer uma face ímpar no lançamento de um dado?

Dados:

- Evento A = {1, 3, 5} (número de casos favoráveis = 3)
- Espaço Amostral (S) = {1, 2, 3, 4, 5, 6} (número de casos possíveis = 6)

$$P(A) = \frac{3}{6} = \frac{1}{2} = 0,5 \text{ ou } 50\%$$

Logo, a probabilidade de aparecer um número ímpar no lançamento de um dado é 0,5 ou 50%.

2. Qual é a probabilidade de se tirar um “rei” de um baralho de 52 cartas?

Dados:

- A = {4 reis} (número de casos favoráveis = 4)
- S = {52 cartas} (número de casos possíveis = 52)

$$P(A) = \frac{4}{52} = \frac{1}{13} = 0,0769 \text{ ou } 7,69\%$$

A probabilidade, portanto, de se tirar um “rei” de um baralho de 52 cartas é  $1/13$  ou 7,69%.

Para facilitar a resolução de problemas na avaliação de probabilidades elementares de variáveis discretas, é apresentado, no Quadro 7.1, um resumo das regras existentes.

Quadro 7.1 – Resumo da aplicação de regras na análise elementar de probabilidades de variáveis discretas

Número da regra	Aplicação da regra	Fórmulas de cálculo
Regra 1	Definição do valor da probabilidade de um evento	$P(A) = \frac{\text{Número de casos favoráveis}}{\text{Número de casos possíveis } (n)}$
Regra 2	“ou” (soma de probabilidades em eventos mutuamente exclusivos)	$P(A+B) = P(A) + P(B)$
Regra 3	“ou” modificada (soma de probabilidades em eventos não mutuamente exclusivos)	$P(A+B) = P(A) + P(B) - P(AB)$
Regra 4	“e” (multiplicação de probabilidades em eventos independentes ou com reposição)	$P(AB) = P(A) \times P(B)$
Regra 5	“e” modificada (multiplicação de probabilidades em eventos condicionais ou sem reposição)	$P(AB) = P(A) \times P(B/A)$
Regra 6	Definição do valor da probabilidade para um evento complementar	$P(\bar{A}) = 1 - P(A)$

Fonte: Dos autores.

É importante observar que na solução de problemas de probabilidade elementar normalmente não é usada uma regra somente. Na maioria das situações, usa-se a conjugação de duas ou mais regras. Existe também a possibilidade de solucionar um determinado problema por várias formas, pela conjugação das regras apresentadas.

## 7.8 Exercícios resolvidos

- Calcular a probabilidade de, retirando-se uma certa carta de um baralho de 52 cartas, obter-se:
  - uma carta, sendo simultaneamente uma carta “rei” e uma carta de espada;
  - uma carta, sendo ou uma carta “rei” ou uma carta de espada.

Solução de “a”:

- Evento A = retirar uma carta “rei”;
- Evento B = retirar uma carta de “espada”;
- Calcular P (AB): regra “e”;
- $P(A) = 4/52$ ;
- $P(B) = 13/52$ .

Como os eventos não são independentes, pois é retirada apenas uma carta que precisa apresentar as duas situações (rei e espada), tem-se que o evento B depende da ocorrência do evento A ou A depende da ocorrência do evento B. Assim sendo, tem-se:

$$P(AB) = P(A) \cdot P(B/A) \text{ ou } P(AB) = P(B) \cdot P(A/B)$$

Num baralho há 52 cartas e quatro reis. Assim,  $P(A) = 4/52$ .

Retirada a carta rei, há quatro possibilidades do naipe, das quais uma é espada:  $P(B/A) = 1/4$

Dessa forma:

$$P(AB) = P(A) \cdot P(B/A)$$

$$P(AB) = 4/52 \times 1/4 = 1/52$$

Ou, considerando a segunda possibilidade:

$$P(AB) = P(B) \cdot P(A/B)$$

Num baralho há 13 cartas de espadas, assim

$$P(B) = 13/52$$

Retirada a carta espada, há uma possibilidade de ser rei:

$$P(A/B) = 1/13$$

$$P(AB) = 13/52 \times 1/13 = 1/52$$

De fato, existe um “rei” de “espada” em um baralho de 52 cartas.

Solução de b:

- Evento A = retirar uma carta “rei”;
- Evento B = retirar uma carta de “espada”;
- Calcular  $P(A+B)$ : regra “ou”;
- $P(A) = 4/52$ ;
- $P(B) = 13/52$ .

Nesse exemplo aceita-se retirar tanto uma carta “rei” quanto uma do naipe de espada. Como existem ao mesmo tempo uma carta “rei” e, também, uma de “espada”, os eventos não são mutuamente exclusivos. Assim, deve-se considerar essa condição para calcular a probabilidade correta. Dessa forma tem-se:

$$P(A+B) = P(A) + P(B) - P(AB), \text{ sendo:}$$

$$P(A) = 4/52 \text{ e } P(B) = 13/52$$

$$P(AB) = 1/52 \text{ (ver resultado da parte a), tem-se:}$$

$$P(A+B) = 4/52 + 13/52 - 1/52 = 16/52 \text{ ou } 4/13$$

A probabilidade, portanto, de se tirar um “rei” ou uma carta de “espada” de um baralho de 52 cartas é 4/13 ou 30,77%.

2. Jogando duas vezes um dado ou dois dados em uma vez, calcular a probabilidade de os pontos obtidos somarem 7.

Solução:

O evento “obter a soma 7” é um evento composto de subconjuntos compostos. O evento é independente e mutuamente exclusivo.

Eventos possíveis:

Para obter um determinado número no primeiro dado e um determinado número no segundo dado, para que a soma seja 7, é necessário usar a regra “e”:

$$P(AB) = P(A) \cdot P(B), \text{ onde,}$$

$$P(A) = 1/6$$

$$P(B) = 1/6. \text{ Assim,}$$

$$P(AB) = 1/6 \times 1/6 = 1/36$$

Existem, portanto, 36 possibilidades diferentes de se obter uma determinada soma.

Eventos favoráveis:

Como existem várias combinações de números entre o resultado do primeiro dado com o resultado do segundo dado, cuja soma é igual a sete, que, portanto, são alternativas, há de se considerar a regra “ou”, ou seja, a soma delas.

$$P(A + B) = \text{soma das diferentes alternativas}$$

Para calcular os eventos favoráveis, podem-se conjugar as possibilidades apresentadas na Tabela 7.1

Tabela 7.1 – Possibilidade da obtenção da soma 7 no lançamento simultâneo de dois dados

Resultados favoráveis no lançamento do primeiro dado	Resultados favoráveis no lançamento do segundo dado	Soma dos resultados dos dois dados
1	6	7
2	5	7
3	4	7
4	3	7
5	2	7
6	1	7

Fonte: Dos autores.

Cada alternativa tem um valor de probabilidade igual a  $1/36$  e como existem seis alternativas, tem-se:

$$P(A + B) = 1/36 + 1/36 + 1/36 + 1/36 + 1/36 + 1/36 = 6/36 =$$

$$P(A + B) = 1/6$$

A probabilidade, portanto, de se obter sete na soma de dois dados é  $1/6$  ou 16,67%.

3. Determinar a probabilidade de aparecer um quatro pelo menos uma vez em dois lançamentos de um dado honesto:

Evento A = aparecimento do 4 no 1o lançamento;

Evento B = aparecimento do 4 no 2o lançamento;

$$P(A + B) = P(A) + P(B) - P(AB).$$

Solução:

Como o quatro pode aparecer nos dois lançamentos, os eventos não são mutuamente exclusivos. Como um lançamento não interfere no outro, os eventos são independentes. Então,

$$P(A) = 1/6$$

$$P(B) = 1/6$$

$$P(AB) = 1/6 \times 1/6 = 1/36$$

$$P(A + B) = 1/6 + 1/6 - 1/36 = 2/6 - 1/36 = 12/36 - 1/36 = 11/36$$

$$P(A + B) = 11/36 \text{ ou } 30,56\%$$

4. Uma bolsa contém quatro bolas brancas e duas pretas. Outra bolsa contém três brancas e cinco pretas. Se for retirada uma bola de cada bolsa, calcular a probabilidade de:

a) ambas serem brancas;

b) ambas serem pretas;

c) uma ser branca e outra ser preta.

Solução:

Evento A: retirar uma bola branca da primeira bolsa;

Evento B: retirar uma bola branca da segunda bolsa;

Evento C: retirar uma bola preta da primeira bolsa;

Evento D: retirar uma bola preta da segunda bolsa.

Dessa forma:

$$P(A) = 4/6$$

$$P(B) = 3/8$$

$$P(C) = 2/6$$

$$P(D) = 5/8$$



Como é retirada uma bola de cada bolsa, os eventos são independentes e mutuamente exclusivos. Assim:

Solução a:

Regra “e” (multiplicação):

$$P(AB) = P(A) \cdot P(B)$$

$$P(AB) = 4/6 \times 3/8 = 12/48 = 1/4 = 0,25 = 25\%$$

Solução b:

Regra “e” (multiplicação):

$$P(CD) = P(C) \cdot P(D)$$

$$P(CD) = 2/6 \times 5/8 = 10/48 = 5/24 = 0,208 = 20,8\%$$

Solução c:

Regra “e” (multiplicação) e regra “ou” (soma das alternativas):

$$P(AD) = P(A) \cdot P(D)$$

ou

$$P(CB) = P(C) \cdot P(B)$$

$$P(AD) = 4/6 \times 5/8 = 20/48$$

$$P(CB) = 2/6 \times 3/8 = 6/48$$

$$P(AD) + P(CB) = 20/48 + 6/48 = 26/48 = 13/24$$

$$P(AD) + P(CB) = 13/24 \text{ ou } 54,17\%$$

5. A probabilidade de que certa porta esteja chaveada é igual a 0,8. A chave correspondente a tal porta está em um chaveiro que contém cinco chaves. Se uma pessoa seleciona ao acaso uma das chaves, qual é a probabilidade de que a porta seja aberta na primeira tentativa?

Dados:

Eventos:

PF = Porta Fechada

PA = Porta Aberta

CC = Chave Certa

CE = Chave Errada

Probabilidades de cada evento:

$$P(PF) = 0,8 \quad P(CC) = 1/5 \text{ ou } 20\%$$

$$P(PA) = 0,2 \quad P(CE) = 4/5 \text{ ou } 80\%$$

Casos possíveis:

(PF.CE), (PF.CC), (PA.CE), (PA.CC)

Casos favoráveis:

(PF.CC), (PA.CC), (PA.CE)

Casos desfavoráveis:

(PF.CE)

Cálculo:

Probabilidades favoráveis = 1 – probabilidades desfavoráveis:

$$P(PF.CC) + P(PA.CC) + P(PA.CE) = 1 - P(PF.CE)$$

$$P(PF.CE) = 0,8 \times 4/5 = 8/10 \times 4/5 = 32/50$$

$$\text{Probabilidades favoráveis} = 1 - 32/50 = 50/50 - 32/50 = 18/50$$

Resposta:

A probabilidade de que a porta seja aberta na primeira tentativa é 18/50 ou 9/25 ou ainda 36%.

## 7.9 Cálculos acessórios

Pelo enfoque clássico, para calcular a probabilidade, precisa-se conhecer os casos favoráveis ao evento e os resultados possíveis. Em problemas mais complexos, são necessários os métodos de permutações e combinações para determinar o número de resultados possíveis.

### 7.9.1 Permutações

O número de permutações de “n” objetos é o número de maneiras pelas quais os objetos podem ser arranjados em termos de ordem.

A definição do número de permutações de “n” objetos é igual ao fatorial de “n”, ou seja,  $n! = n(n-1) \times \dots \times 2 \times 1$ .

Exemplo de cálculo de  $n!$  (fatorial) = 3

$$3! = 3 \times 2 \times 1 = 6$$

Quando se quer determinar o número de permutações de “n” objetos tomados “r” de cada vez, onde “r” é menor do que “n”, usa-se a seguinte fórmula:

$${}_nPr = \frac{n!}{(n-r)!} \quad (7.6)$$

onde:

${}_nPr$  = número de permutações de “n” objetos tomados “r” de cada vez

$n!$  = fatorial de “n”

n = número de dados

r = número de dados agrupados por vez

$(n-r)!$  = fatorial da diferença entre “n” e “r”

### 7.9.2 Combinações

As combinações indicam o número de diferentes agrupamentos de objetos que podem ocorrer sem levar em consideração a ordem. A combinação define o número de diferentes subgrupos que se pode formar com “n” objetos.

O número de combinações de “n” objetos tomados “r” de cada vez é calculado pela seguinte fórmula:

$${}_nC_r = \frac{n!}{r!(n-r)!} \quad (7.7)$$

onde:

${}_nC_r$  = Número de combinações de n objetos tomados r de cada vez

$n!$  = fatorial de n

n = número de dados

r = número de dados agrupados por vez

$r!$  = fatorial do número de dados agrupados por vez

$(n-r)!$  = fatorial da diferença entre n e r

### 7.9.3 Combinação especial

Uma forma especial de combinação é conhecida como princípio da multiplicação para resultados sequenciais. Essa situação ocorre quando dois eventos são independentes, podendo o primeiro ocorrer de  $n_1$  maneiras e o segundo, de  $n_2$  maneiras. Nesse caso, o total de maneiras diferentes que os dois eventos podem ocorrer em combinação é :  $n_1 \times n_2$ .

### 7.10 Teorema de Bayes

O teorema de Bayes representa a relação da combinação da regra da multiplicação de eventos mutuamente exclusivos favoráveis com a combinação da regra da multiplicação dos eventos mutuamente exclusivos possíveis ou probabilidade total.

Dessa forma, o teorema de Bayes constitui um quadro lógico importante para a revisão ou a atualização de probabilidades previamente estabelecidas, à luz de novas informações. Pelo teorema de Bayes existe a possibilidade do cálculo da probabilidade a *posteriori*, considerando as probabilidades associadas aos resultados do primeiro estágio, dado o resultado do segundo estágio.

Assim, dado um experimento em dois estágios e conhecidas as probabilidades “a priori”,  $P(A_i)$ , e as probabilidades condicionais,  $P(B/A_i)$ , o teorema de Bayes permite calcular as probabilidades “a posteriori”,  $P(A_i/B)$ , isto é, as probabilidades associadas aos resultados do primeiro estágio, dado o resultado do segundo estágio, pela seguinte fórmula:

$$P(A_i / B) = \frac{P(A_i) * P(B/A_i)}{P(A_1) * P(B/A_1) + P(A_2) * P(B/A_2) + \dots + P(A_n) * P(B/A_n)} \quad (7.8)$$

Exemplo:

Há duas caixas, havendo na primeira três bolas brancas e sete pretas, e na segunda, uma bola branca e cinco pretas, das quais foi retirada uma bola ao acaso. Retirada a bola, observa-se que ela é preta. Qual é a probabilidade de que a caixa de onde foi extraída a bola seja a primeira?

$P(A_1) = 1^{\text{a}} \text{ caixa} = 7/10$ , ou seja, 7 pretas em 10 bolas = 0,70

$P(A_2) = 2^{\text{a}} \text{ caixa} = 5/6$ , ou seja, 5 pretas em 6 bolas = 0,8333

$P(B/A_1) = 1/2$ , ou seja, tirar uma bola da 1ª caixa = 0,5

$P(B/A_2) = 1/2$ , ou seja, tirar uma bola da 2ª caixa = 0,5

$$P(A_i / B) = \frac{P(A_i) * P(B/A_i)}{P(A_1) * P(B/A_1) + P(A_2) * P(B/A_2) + \dots + P(A_n) * P(B/A_n)}$$

$$P(A_1 / B) = \frac{\frac{7}{10} \times \frac{1}{2}}{\frac{7}{10} \times \frac{1}{2} + \frac{5}{6} \times \frac{1}{2}} = \frac{\frac{7}{20}}{\frac{7}{20} + \frac{5}{12}}$$

$$P(A_1 / B) = \frac{\frac{7}{20}}{\frac{46}{60}} = \frac{7}{20} \times \frac{60}{46} = \frac{21}{46}$$

Ou seja,

$P(A_1/B) = 21/46$  ou 45,65%

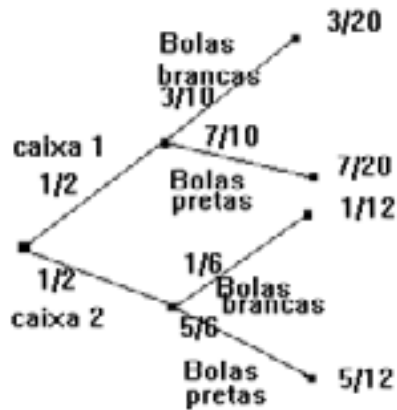
Ou, usando os valores das probabilidades “a priori” e condicionais, a solução apresenta-se da seguinte forma:

$$P(A_1/B) = \frac{0,70 \cdot 0,5}{0,70 \cdot 0,5 + 0,8333 \cdot 0,5} = \frac{0,35}{0,50 + 0,4166} = 0,4566 \text{ ou } 45,66\%$$

A probabilidade de que a caixa de onde foi extraída a bola seja a primeira é de 45,65%.

O exemplo apresentado acima também pode ser resolvido se for construída “a árvore de possibilidade”, conforme Figura 7.2. Cada ramo mostra uma situação possível e, no fim de cada ramo, estão colocados os valores das respectivas probabilidades.

Figura 7.2 – Árvore de possibilidades de probabilidade



Fonte: Dos autores.

Onde:

Casos favoráveis: 7/20

Casos possíveis: 7/20 + 5/12

$$P(A_1 / B) = \frac{\frac{7}{20}}{\frac{7}{20} + \frac{5}{12}} = \frac{\frac{7}{20}}{\frac{46}{60}} = \frac{7}{20} \times \frac{60}{46} = \frac{21}{46} = 0,4565 \text{ ou } 45,65$$

## 7.11 Distribuição de probabilidade

Nota-se que a probabilidade é análoga a uma distribuição de frequências relativas, com estas substituídas pelas probabilidades. Por essa razão, podem-se imaginar as distribuições de probabilidade como uma forma teórica ou de limite ideal das distribuições de frequências relativas, quando o número de observações feitas tornar-se muito grande. Assim, pode-se concluir que as distribuições de probabilidade referem-se à população, ao passo que as distribuições de frequências relativas referem-se a amostras dela extraídas.

A distribuição de probabilidade pode ser representada graficamente, mediante alocação de  $P(X)$  em relação a  $X_i$ , da mesma forma que a distribuição da frequência relativa  $fr$ .

Mediante a acumulação das probabilidades obtêm-se distribuições de probabilidades acumuladas, análogas às de frequência relativa acumulada.

Dependendo da variável, se for discreta ou contínua, do tipo do evento, da quantidade de dados disponíveis e do tamanho da amostra, existem diversas formas de distribuições de probabilidades teóricas, entre as quais as mais comuns são:

- Binomial ou Distribuição de Bernoulli – para eventos independentes e mutuamente exclusivos em experimentos estáticos;
- Distribuição de Poisson – para eventos independentes e mutuamente exclusivos em experimentos dinâmicos;
- Distribuição Hipergeométrica – para eventos condicionais e mutuamente exclusivos;
- Distribuição Normal – para dados contínuos e grandes amostras.
- Distribuição “t” de Student – para dados contínuos e pequenas amostras;
- Distribuição Qui-Quadrado – para variáveis discretas qualitativas com duas ou mais categorias.

### 7.12 Resumo para fixação

Para auxiliar na fixação dos principais aspectos abordados no capítulo, são apresentados os seguintes pontos como resumo:

- A probabilidade ajuda a lidar com a incerteza dos eventos.
- O resultado do cálculo das probabilidades ficará entre 0 e 1, ou seja, 0 quando o evento não ocorrer e 1 quando o evento tem a certeza de ocorrência.
- As regras mais utilizadas para calcular as probabilidades são:
  - Probabilidade de um evento: número de casos favoráveis dividido pelo número de casos possíveis;
  - Probabilidade de um ou outro evento independente e mutuamente exclusivo:  $P(A+B) = P(A) + P(B)$ ;
  - Probabilidade de um ou outro evento independente e não mutuamente exclusivo:  $P(A+B) = P(A) + P(B) - P(AB)$ ;
  - Probabilidade de ambos eventos independentes (ou com reposição):  $P(AB) = P(A) \times P(B)$ ;
  - Probabilidade de ambos eventos dependentes (ou sem reposição):  $P(AB) = P(A) \times P(B/A)$ .

### 7.13 Exercícios propostos

7.1 Na casa de dona Maria, há um casal de cães cujos filhotes nascem brancos, pretos ou malhados. Considerando o experimento aleatório do nascimento de dois cães da casa de dona Maria, responder:

- a) Qual o espaço amostral do experimento? (Denotar B para branco, P para preto e M para malhado)
- b) Qual a probabilidade de nascer dois cães malhados?
- c) Qual a probabilidade de nascer no mínimo, um cão malhado?

7.2 Na turma A de estatística há 25 colorados e 20 gremistas, enquanto que na turma B há 20 colorados e 30 gremistas. Caso seja escolhido, ao acaso, um aluno de cada turma, a probabilidade de ambos serem gremistas é aproximadamente:

- (a) 26,67%      (b) 22,22%      (c) 52,63%      (d) 47,37%      (e) nenhuma das anteriores

7.3 Na localidade de Arroio do Mato há 250 moradores, 140 praticam atividade física, 40 fazem artesanato e 10 praticam atividade física e fazem artesanato. Se um dos moradores dessa localidade for selecionado aleatoriamente, a probabilidade de que o morador pratique atividade física ou faça artesanato é:

- (a) 56,0%      (b) 76,0%      (c) 68,0%      (d) 16,0%      (e) nenhuma das anteriores

7.4 Um inspetor testa duas lâmpadas de um lote de 25 unidades. Dado que o lote tem quatro lâmpadas queimadas, qual é a probabilidade de ele não achar nenhuma com defeito (21 lâmpadas boas)?

- a) Considere com reposição;      b) Considere sem reposição.

7.5 Em um lote de 12 peças, três são defeituosas (e nove são boas), sendo retiradas, aleatoriamente, duas peças, sem reposição, calcule:

- a) a probabilidade de ambas serem defeituosas;      c) a probabilidade de uma ser defeituosa.
- b) a probabilidade de ambas não serem defeituosas;

7.6 Você é o responsável pela área de produção numa fábrica. Há quatro máquinas na linha de produção e, conforme controle histórico, as probabilidades de falha em determinado dia são 1%, 2%, 5% e 10%. Diante disso, calcule as probabilidades:

- a) de todas falharem em determinado dia;      c) de apenas a terceira falhar.
- b) de nenhuma falhar;

7.7 Uma empresa garante, na embalagem de seu produto, que apenas 1% das peças produzidas por ela são defeituosas. Se forem adquiridas duas peças produzidas por essa empresa, considerando os eventos independentes, qual é a probabilidade de que:

- a) ambas sejam defeituosas?      d) uma seja defeituosa e a outra boa?
- b) pelo menos uma seja defeituosa?      e) pelo menos uma seja boa?
- c) ambas sejam boas?

7.8 Uma urna possui seis bolas azuis, 10 bolas vermelhas e quatro bolas amarelas. Tirando-se duas bolas sem reposição, calcule as probabilidades de:

- a) saírem uma bola azul e uma amarela;      c) saírem duas bolas amarelas;
- b) saírem duas bolas vermelhas;      d) sair pelo menos uma bola vermelha.

## 8 DISTRIBUIÇÕES BÁSICAS DE PROBABILIDADE DE VARIÁVEIS DISCRETAS

### 8.1 Distribuição de probabilidade

Quando se definem os valores de probabilidade de todas as possibilidades de uma variável aleatória  $X$ , tanto por uma listagem como por uma função matemática, o resultado é uma distribuição de probabilidade dessa variável.

O conhecimento de certas distribuições teóricas de probabilidade permite a determinação de probabilidades dos eventos aparentemente complexos, com certa facilidade, e é essencial para a resolução de problemas de inferência estatística.

Entende-se por variável aleatória discreta em uma distribuição de probabilidades aquela cujos valores são determinados por processos acidentais, ao acaso, que não estão sob o controle do observador. Desta forma, uma variável aleatória é uma variável que tem associada a si as probabilidades de assumir seus diferentes valores (no caso de ser uma variável discreta) ou se encontrar dentro de intervalos determinados (no caso de variável contínua).

Assim, as distribuições de probabilidade serão calculadas de acordo com o tipo de variável aleatória: discreta ou contínua.

Uma variável aleatória discreta admite um número finito de valores (ou seja, pode-se contar o espaço amostral), sendo possível listar todos os possíveis valores da variável em uma tabela de probabilidades correspondentes, cuja soma de todos os resultados possíveis será igual a um. Nesse caso, é comum calcular a probabilidade de ocorrência pela distribuição binomial, quando baseada em eventos estáticos, ou pela distribuição de Poisson, quando baseada em eventos dinâmicos. Além dessas duas formas de cálculo, ainda existem outros tipos de distribuições discretas de probabilidade, mas que não serão abordadas neste livro.

Por outro lado, quando a variável aleatória for contínua, ela admite um número infinito de valores, não sendo possível listá-los, apenas definir os intervalos nos quais se encontrarão as variáveis. A principal maneira de calcular a probabilidade, nesse caso, é por meio da distribuição normal, que será analisada no capítulo 9.

A seguir, são apresentadas, de forma mais detalhada, a distribuição binomial e a distribuição de *Poisson*.

### 8.2 Distribuição binomial

A distribuição binomial é uma das distribuições discretas mais utilizadas. Essa distribuição baseia-se no processo de amostragem do matemático Jacques Bernoulli, que enunciou e demonstrou a “Lei dos Grandes Números”<sup>1</sup> que uniu a abstração do cálculo matemático com a realidade expressa pelos números.

Essa distribuição destina-se a eventos mutuamente exclusivos (em que a ocorrência de um evento impossibilita a ocorrência de outro evento) e independentes (quando a probabilidade de um evento não interfere na probabilidade do outro evento ocorrer). Dessa forma, a essência dessa distribuição está de acordo com a probabilidade elementar, pois também parte do princípio de que todos os eventos de um experimento têm iguais possibilidades de ocorrer. No entanto, elas divergem quanto ao número de experimentos. Em geral, quando há até três experimentos, utiliza-

<sup>1</sup> A Lei dos Grandes Números afirma que a frequência de determinados acontecimentos, observados em um grande número de casos, tende a se estabilizar à medida que aumenta o número de casos observados, de modo a aproximar-se do valor previsto pela teoria das probabilidades. Assim, ela associa a probabilidade de um evento  $A$  à sua frequência quando o número de experimentos é muito grande. Como exemplo pode-se citar um jogo de cara ou coroa.



se a probabilidade elementar, e, quando forem mais experimentos, utiliza-se a probabilidade binomial.

Assim, as características da distribuição binominal são as seguintes:

- a) em cada tentativa existem dois resultados possíveis **mutuamente exclusivos**, chamados de **sucesso e fracasso**;
- b) as séries de tentativas ou observações são constituídas de **eventos independentes**;
- c) a probabilidade de sucesso permanece constante de tentativa em tentativa, ou seja, o processo é estacionário, pois os eventos são independentes ou com reposição.

A palavra **sucesso** está relacionada ao êxito estatístico do evento em análise, não representando necessariamente algo bom ou positivo. Em outras palavras, situações de sucesso caracterizam-se por encontrar aquilo que se pretende estudar ou se deseja encontrar. Por exemplo, quando se pretende testar a probabilidade de encontrar peças com defeitos em um processo produtivo de uma empresa, o sucesso estará atribuído à probabilidade de encontrar peças defeituosas, enquanto o fracasso do evento estará relacionado à probabilidade de encontrar peças sem defeitos, mesmo que intuitivamente parece ser ruim encontrar peças com defeito e bom encontrar peças sem defeito.

### 8.2.1 Conceito

A distribuição binomial é a probabilidade de “**x**” êxitos em uma série de “**n**” experimentos de evento com probabilidade de “**p**” que suceda em uma só vez e de “**q**” que não suceda em um só intento.

Em outras palavras, a distribuição binomial aplica-se a experimentos independentes, que podem ser repetidos, e para os quais somente são possíveis dois resultados: sucesso (p) ou fracasso (q).

Assim, para calcular a probabilidade por meio da distribuição binomial, três valores são necessários: número de sucessos (x), número de tentativas ou observações (n) e probabilidade de sucesso em cada tentativa (p).

Abaixo serão desenvolvidas três formas de calcular a probabilidade de x sucessos em n tentativas de um experimento binomial, utilizando o método de Bernoulli, a fórmula de cálculo e a tabela de probabilidade binomial.

### 8.2.2 Método de Bernoulli

A distribuição binomial tem sua origem nos termos do desenvolvimento do binômio de Newton  $(p + q)^n$ , usado por Bernoulli.

Exemplo:

Análise das possibilidades de soma no lançamento de dois dados:

Tabela 8.1 Probabilidades de resposta em um lançamento de dois dados

Valores possíveis de soma	2	3	4	5	6	7	8	9	10	11	12	$\Sigma$
Número de Sucessos	1	2	3	4	5	6	5	4	3	2	1	36
Probabilidade de Sucesso	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	1

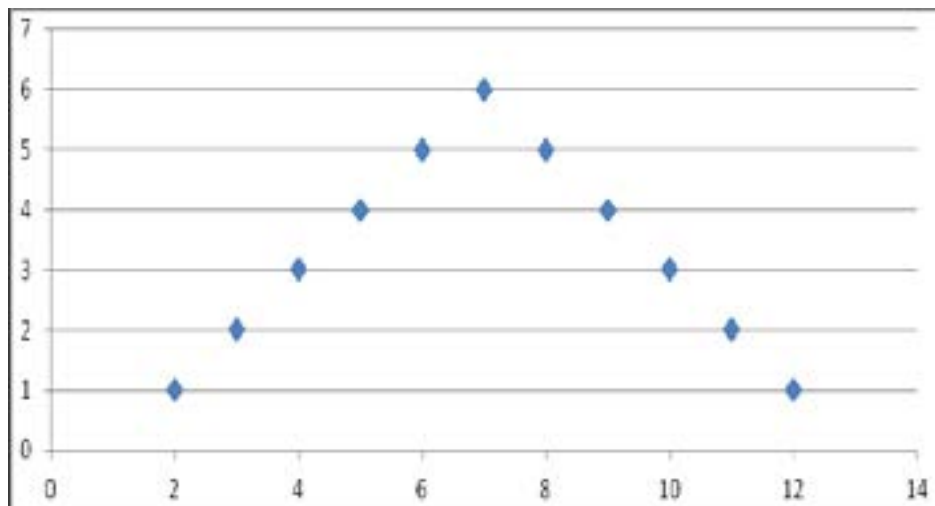
Fonte: Dos autores.

Em cada lançamento de um dado, existem seis resultados possíveis (1, 2, 3, 4, 5, 6). Assim, quando se lançam simultaneamente dois dados, podem-se obter diferentes resultados de soma dos dois valores, conforme é apresentado na primeira linha da Tabela 8.1, ou seja, a soma mínima dos dois dados é dois (quando em ambos, sair o número 1) e a máxima é 12 (quando



em ambos os dados se obter o número 6). No entanto, esses dados podem se combinar de diferentes maneiras. A segunda linha da tabela acima apresenta o número de sucessos de se obter as somas apresentadas na primeira linha, isto é, para obter-se a soma 2, somente existe uma combinação possível {1, 1}; para obter-se a soma 3, os resultados podem ser {1, 2} ou {2, 1}; para obter-se a soma 4, as combinações podem ser as seguintes: {1, 3}, {3, 1} ou {2, 2}; ... sendo possível obter um total de 36 diferentes combinações no lançamento de 2 dados. E, na última linha da tabela são expressas as probabilidades de sucesso de acordo com os resultados obtidos nas linhas 1 e 2. Esses resultados podem ser visualizados no Gráfico 8.1

Gráfico 8.1- Probabilidades de resposta em um lançamento de dois dados



Fonte: Dos autores.

Assim, o Gráfico 8.1 também ilustra as probabilidades de resposta em um lançamento de dois dados. Por meio dele pode-se observar que a probabilidade de se obter somas medianas é maior do que se obter resultados mais extremos (apenas números mais baixos ou apenas números mais altos).

### 8.2.3 Fórmula de cálculo

Bernoulli utilizou a análise combinatória<sup>2</sup> entre as possibilidades de sucessos e de fracassos para calcular a probabilidade de ocorrência dos experimentos. Com base nesses princípios, ele desenvolveu a fórmula abaixo, para se determinar a probabilidade de certo número de sucessos que é dada por:

$$P(x) = nCp^x(1-p)^{n-x} \quad (8.1)$$

onde:

$P(x)$  = probabilidade de  $x$

$n$  = número de experimentos

$x$  = números de sucessos (êxitos) que se deseja obter

$1 - p$  = probabilidade de não ocorrer o sucesso do número de êxitos (ou o fracasso)

$nC$  = determina o número de maneiras de se obter  $x$  sucessos em  $n$  tentativas, independente da ordem.

<sup>2</sup> A análise combinatória é um ramo da Matemática que estuda, em especial, a formação de agrupamentos de elementos, numa abordagem quantitativa, a partir de um determinado conjunto, sendo esses elementos formados sob certas condições previamente estabelecidas.

Substituindo os valores de  $nC$  por:

$$nC = \frac{n!}{x!(n-x)!}$$

tem-se:

$$P(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \quad (8.2)$$

A probabilidade de sucesso é atribuída a “p”. Logo, a probabilidade de fracasso ou insucesso será  $(1 - p)$ , que será chamado de “q”. Substituindo na fórmula, tem-se:

$$P(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x} \quad (8.3)$$

onde:

$n$  = nº de tentativas

$q$  = probabilidade de fracasso

$x$  = nº de sucessos

$p$  = probabilidade de sucesso

! = símbolo de fatorial, o que representa o produto de fatores decrescentes

#### 8.2.4 Tabela de Probabilidade Binomial

A fórmula da probabilidade binomial envolve uma quantidade considerável de cálculos quando a amostra é relativamente grande, e, particularmente, quando se deseja determinar a probabilidade de que o resultado ocorra em um intervalo de valores. Por esse motivo, para a sua resolução, é frequentemente, utilizada a Tabela de Probabilidade Binomial.

A Tabela de Probabilidade Binomial (APÊNDICE I) apresenta a solução prévia de resultados com determinadas probabilidades de sucesso associadas ao número de sucessos e de tentativas, sendo calculada a partir da fórmula 8.3. Assim, caso seja possível encontrar a solução de um problema pela tabela, não há a necessidade de realizar o respectivo cálculo. No entanto, quando a probabilidade de sucesso não constar na tabela, ou o número de tentativas envolvidas for superior a 30, somente será encontrado o resultado mediante a fórmula de cálculo.

A tabela do Apêndice I apresenta as informações conforme modelo abaixo:

Tabela 8.2 - Modelo de Tabela de Probabilidade Binomial

Número de tentativas		Probabilidade de sucessos									
n	x	P									
		0,01	0,05	0,10	0,20	0,25	0,30	0,35	0,40	0,45	0,50
1	0										
	1										
2	0										
	1										
	2										
...	...										

Fonte: Dos autores.

A tabela, portanto, deve ser interpretada da seguinte forma:

- na linha superior encontra-se a probabilidade de sucesso dada por  $p$ ;
- na primeira coluna encontra-se o número de tentativas ( $n$ );
- na segunda coluna encontra-se o número de sucessos desejados ( $x$ );
- para encontrar o resultado, deve-se procurar no corpo da tabela a probabilidade que corresponde à linha que contém o número de tentativas e o número de sucessos desejados com a coluna que contém a probabilidade de sucesso.

Quando a probabilidade de sucesso que se está usando no cálculo se encontrar na tabela, pode-se, para resolver qualquer problema de probabilidade binomial, usar a tabela pré-calculada, sem haver a necessidade de desenvolver o cálculo.

Por outro lado, quando a probabilidade de sucesso for um número diferente do encontrado na tabela, deve-se usar a fórmula de cálculo.

Exemplos:

1. Qual é a probabilidade de se obter três caras em um experimento de quatro lançamentos de uma moeda?

Dado o problema, inicialmente, precisa-se retirar as informações básicas. Neste caso, o sucesso está relacionado à obtenção de três caras no lançamento de quatro moedas, e qualquer resultado diferente desse será considerado um fracasso. Além disso, a probabilidade de sair cara no lançamento de cada moeda permanece sempre a mesma, independente do número de vezes que o experimento for realizado. Com base nessas informações têm-se os seguintes dados:

$$x = 3$$

$$P(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

$$n = 4$$

$$P(3) = \frac{4!}{3!(4-3)!} (1/2)^3 (1/2)^{4-3}$$

$$p = 1/2$$

$$q = 1/2$$

$$P(3) = \frac{4 \times 3 \times 2 \times 1}{3 \times 2 \times 1 \times 1} \times (1/8) \times (1/2) = 4 \times (1/8) \times (1/2) = 4/16 = 1/4$$

Resposta:

A probabilidade de obter três caras no lançamento de quatro moedas é igual a  $1/4$  ou 25%.

2. Em certo lugar a proporção de filhos homens é de  $2/5$ . Em uma família de seis filhos qual seria a probabilidade de pelo menos a metade deles serem homens?

Nesse exemplo, procura-se encontrar a probabilidade de que pelo menos metade dos filhos, sejam homens, ou seja, o sucesso será encontrar no mínimo três filhos homens, em um total de seis filhos. Isso significa que se quer encontrar a probabilidade de três filhos serem homens, ou quatro filhos serem homens, ou cinco filhos serem homens, ou, ainda, que os seis filhos sejam homens. Assim, para resolver o problema, precisa-se calcular a probabilidade de cada um desses quatro eventos ocorrerem isoladamente e depois somá-las.

Dados:

$$n = 6$$

$$x \geq 3$$

$$p = 2/5$$

$$q = 3/5$$

Logo, deve-se calcular  $P(3) + P(4) + P(5) + P(6) = P(x \geq 3)$

$$P(3) = \frac{6!}{3!(6-3)!} (2/5)^3 (3/5)^{6-3}$$

$$P(3) = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1 \times 3 \times 2 \times 1} \times (8/125) \times (27/125) = 20 \times (216/15625) = 4320/15625$$

$$P(4) = \frac{6!}{4!(6-4)!} (2/5)^4 (3/5)^{6-4}$$

$$P(4) = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{4 \times 3 \times 2 \times 1 \times 2 \times 1} \times (16/625) \times (9/25) = 15 \times (144/15625) = 2160/15625$$

$$P(5) = \frac{6!}{5!(6-5)!} (2/5)^5 (3/5)^{6-5}$$

$$P(5) = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{5 \times 4 \times 3 \times 2 \times 1 \times 1} \times (32/3125) \times (3/5) = 6 \times (96/15625) = 576/15625$$

$$P(6) = \frac{6!}{6!(6-6)!} (2/5)^6 (3/5)^{6-6}$$

$$P(6) = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{6 \times 5 \times 4 \times 3 \times 2 \times 1 \times 1} \times (64/15625) \times 1 = 1 \times (64/15625) = 64/15625$$

$$P(x \geq 3) = \frac{4320 + 2160 + 576 + 64}{15625} = \frac{7120}{15625} = 0,45568$$

Resposta:

A probabilidade de encontrar pelo menos três filhos homens é igual a 45,57 %. Esse valor também pode ser obtido por meio do Apêndice I.

3. Os elos de uma corrente têm a probabilidade de 1/3 de se romper sob uma certa carga. Aplicando essa carga a uma corrente constituída de três elos, qual é a probabilidade de ela romper?

Dados:

$$n = 3$$

$$p = 1/3$$

$$q = 2/3$$

$$x = 1 \text{ ou } 2 \text{ ou } 3$$

$$P(1) = \frac{3!}{1!(3-1)!} (1/3)^1 (2/3)^{3-1}$$

$$P(1) = \frac{3 \times 2 \times 1}{1 \times 2 \times 1} \times (1/3) \times (4/9) = 3 \times (1/3) \times (4/9) = 12/27$$

$$P(2) = \frac{3!}{2!(3-2)!} (1/3)^2 (2/3)^{3-2}$$

$$P(2) = \frac{3 \times 2 \times 1}{2 \times 1 \times 1} \times (1/3)^2 \times (2/3)^{3-2} = 3 \times (1/9) \times (2/3) = 6/27$$

$$P(3) = \frac{3!}{3!(3-3)!} (1/3)^3 (2/3)^{3-3}$$

$$P(3) = \frac{3 \times 2 \times 1}{3 \times 2 \times 1 \times 1} \times (1/3)^3 \times (2/3)^{3-3} = 1 \times (1/27) \times 1 = 1/27$$

$$P(1 \leq x \leq 3) = \frac{12 + 6 + 1}{27} = \frac{19}{27} = 0,7037 \text{ ou } 70,37\%$$

Resposta:

A probabilidade de um dos três elos se romper é igual a 70,37%.

4. Se forem jogadas cinco moedas simultaneamente, qual é a probabilidade de se obter três caras e duas coroas?

Os resultados do lançamento de uma moeda caracterizam-se como eventos mutuamente exclusivos, ou seja, a obtenção de uma cara impede a obtenção de uma coroa. Assim, nesse caso, deve-se considerar apenas um dos dois casos, ou calcula-se a probabilidade de obter três caras ou a probabilidade de obter duas coroas, pois o restante do resultado está intrínseco.

Dados:

$$n = 5$$

$$p = 1/2$$

$$q = 1/2$$

$$x = 3 \text{ caras}$$

$$P(3) = \frac{5!}{3!(5-3)!} (1/2)^3 (1/2)^{5-3}$$

$$P(3) = \frac{5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1 \times 2 \times 1} \times (1/8) \times (1/4) = 10 \times (1/8) \times (1/4) = 5/16 = 0,3125 \text{ ou } 31,25$$

$$\text{Ou, } x = 2 \text{ coroas}$$

$$P(2) = \frac{5!}{2!(5-2)!} (1/2)^2 (1/2)^{5-2}$$

$$P(2) = \frac{5 \times 4 \times 3 \times 2 \times 1}{2 \times 1 \times 3 \times 2 \times 1} \times (1/4) \times (1/8) = 10 \times (1/4) \times (1/8) = 5/16 = 0,3125 \text{ ou } 31,25$$

Resposta:

A probabilidade de obter três caras e duas coroas é igual a 31,25%.

5. Considerando a distribuição binomial, nove ferramentas são escolhidas ao acaso, que apresentam 5% de peças defeituosas. Qual é a probabilidade de pelo menos duas peças serem defeituosas?

Dados:

$$n = 9$$

$$x \geq 2 \quad (P(2) + P(3) + P(4) + \dots + P(9))$$

$$p = 0,05$$

$$q = 0,95$$

Esse problema pode ser solucionado de duas formas. Como se quer encontrar a probabilidade de  $x \geq 2$ , pode-se calcular a probabilidade de:

$$P(x \geq 2) = P(2) + P(3) + P(4) + \dots + P(9), \text{ onde,}$$

$$P(x \geq 2) = 0,0629 + 0,0077 + 0,0006 + 0,0000 + 0,0000 + 0,0000 + 0,0000 + 0,0000 = 0,0713$$

Ou

$P(x \geq 2) = 1 - P(0) - P(1)$ , ou seja, nesse caso, desconta-se de 100% das possibilidades a probabilidade dos eventos que não se incluem no problema.

Assim:

$$P(0) = \frac{9!}{0!(9-0)!} (0,05)^0 (0,95)^9 = 1 \times 1 \times 0,6302 = 0,6302$$

$$P(1) = \frac{9!}{1!(9-1)!} (0,05)^1 (0,95)^{9-1} = 9 \times 0,05 \times 0,6634 = 0,2985$$

$$P(x \geq 2) = 1 - 0,6302 - 0,2985 = 0,0713$$

Resposta:

A probabilidade de encontrar pelo menos duas peças com defeito é de 7,13%.

### 8.2.5 Propriedades da Distribuição Binomial

Com base nas características da distribuição binomial é possível encontrar a média, a variância e o desvio-padrão de um conjunto de observações de forma simplificada, conforme aparece abaixo.

$$\text{Média: } \mu = Np \text{ ou } \bar{X} = np \quad (8.4)$$

$$\text{Variância: } \sigma^2 = Npq \text{ ou } s^2 = npq \quad (8.5)$$

$$\text{Desvio-padrão: } \sigma = \sqrt{Npq} \text{ ou } s = \sqrt{npq} \quad (8.6)$$

Exemplo:

Em 100 lances de uma moeda honesta, quais são a média de caras e o seu desvio-padrão?

Dados:

$$n = 100$$

$$p = 1/2$$

$$q = 1/2$$

$$\bar{X} = 100 \times 1/2 = 50 \text{ caras}$$

$$s = \sqrt{100 \times 1/2 \times 1/2} = \sqrt{25} = 5$$

### 8.3 Distribuição de Poisson

A distribuição de *Poisson* encontra aplicações principalmente no planejamento da capacidade de postos de serviço e no Controle Estatístico de Qualidade. Ela consegue descrever o comportamento de eventos com pequenas probabilidades de ocorrência.

Para melhor entender a aplicação dessa distribuição, será usado o exemplo de um grande prédio destinado a escritórios de empresas em que existe um balcão de informações logo na entrada. Depois de muito tempo de funcionamento, chega-se à conclusão de que em média a cada minuto ocorrem dois pedidos de informações mais detalhadas. A partir desta informação, quer se saber qual é a probabilidade de ocorrer um número qualquer dado (inteiro e positivo) de pedidos de informação no intervalo de um minuto. Por exemplo, qual é a probabilidade de nenhum pedido? De dois pedidos? De 15 pedidos? A resposta pode ser dada se for admitido que a distribuição de probabilidade do número de pedidos de informação por minuto obedece à Distribuição de Poisson.

Assim, a Distribuição de Poisson é utilizada quando a variável aleatória está referida a uma certa unidade de exposição, de tempo, de superfície ou de volume que possa assumir qualquer valor inteiro e positivo, ou seja, uma variável discreta, admitindo-se ser conhecido o valor médio da variável por unidade de exposição. Essa distribuição consegue medir eventos dinâmicos não estáticos, baseando-se em seu valor médio e não em uma amostra exata (não considera N).

#### 8.3.1 Características da Distribuição de Poisson

A distribuição de *Poisson* é uma distribuição discreta de probabilidade de uma variável aleatória que satisfaz as seguintes condições:

- o experimento consiste na contagem do número de vezes,  $x$ , que um evento ocorre em um determinado intervalo. O intervalo pode ser de tempo, de área ou de volume;

- a probabilidade de que o evento ocorra é a mesma em cada intervalo;
- o número de ocorrências em um intervalo independe do número de ocorrências em outros intervalos.

### 8.3.2 Fórmula de cálculo

A fórmula para se determinar a probabilidade de que haja exatamente  $x$  ocorrências em um intervalo é dada por:

$$P(x) = \frac{\mu^x e^{-\mu}}{x!} \quad (8.7)$$

onde:

$\mu$  = número médio da variável por unidade de exposição

$e$  = é um número irracional que representa a base do sistema de logaritmos naturais (é uma constante), com valor aproximadamente igual a 2,71828.

$x!$  = fatorial do número de ocorrência do evento

### 8.3.3 Tabela da Probabilidade de Poisson

Da mesma forma como no caso da probabilidade binomial, também a solução dessa distribuição envolve uma quantidade considerável de cálculos, principalmente quando o número de ocorrências é muito grande, e, particularmente, quando se deseja determinar a probabilidade de que o resultado ocorra em um intervalo de valores. Assim, para a sua resolução, também é utilizada uma tabela previamente calculada, conhecida pela Tabela de Probabilidade Poisson (APÊNDICE II), que apresenta as seguintes informações:

Tabela 8.3 - Modelo de Tabela de Probabilidade Poisson

x	$\mu$									
	3,1	3,2	3,3	3,4	3,5	3,6	3,7	3,8	3,9	4,0
0										
1										
2										
3										
4										
...										

Fonte: Dos autores.

A Tabela do Apêndice II tem a seguinte interpretação:

- na linha superior encontra-se o número médio da variável por unidade de exposição;
- na primeira coluna encontra-se o número de ocorrências em um determinado intervalo;
- para encontrar o resultado, deve-se procurar no corpo da tabela a probabilidade que corresponde à coluna que contém o número médio da variável por unidade de exposição e a linha que contém o número de ocorrências em um determinado intervalo desejado.

Quando na tabela encontrar-se o número médio de ocorrências, pode-se resolver qualquer problema de probabilidade de Poisson, ao invés de usar a fórmula.

Quando na tabela não for encontrado o número médio de ocorrência, deve-se usar a fórmula de cálculo.

**Exemplos:**

1. Uma contagem populacional revela a existência de uma média de 3,6 coelhos por ha. em uma determinada região rural (dados fictícios). Use uma tabela para determinar a probabilidade de que sejam encontrados dois coelhos em qualquer ha. determinado da área rural.

**Dados:**

$$x = 2$$

$$\mu = 3,6$$

**Solução:**

Para a solução, foi reproduzida uma parte da Tabela do Apêndice II a seguir. Se  $\mu = 3,6$  e  $x = 2$ , pode-se obter a probabilidade de Poisson conforme mostrado pelas áreas realçadas na tabela:

**Apêndice II - Distribuição de Poisson**

$\mu$							
x	3,1	3,2	3,3	3,4	3,5	3,6	3,7
0	0,0450	0,0408	0,0369	0,0334	0,0302	0,0273	0,0247
1	0,1397	0,1304	0,1217	0,1135	0,1057	0,0984	0,0915
2	0,2165	0,2087	0,2008	0,1929	0,1850	0,1771	0,1692
3	0,2237	0,2226	0,2209	0,2186	0,2158	0,2125	0,2087
4	0,1733	0,1781	0,1823	0,1858	0,1888	0,1912	0,1931
5	0,1075	0,1140	0,1203	0,1264	0,1322	0,1377	0,1429
6	0,0555	0,0608	0,0662	0,0716	0,0771	0,0826	0,0881
7	0,0246	0,0278	0,0312	0,0348	0,0385	0,0425	0,0466
8	0,0095	0,0111	0,0129	0,0148	0,0169	0,0191	0,0215
9	0,0033	0,0040	0,0047	0,0056	0,0066	0,0076	0,0089
10	0,0010	0,0013	0,0016	0,0019	0,0023	0,0028	0,0033

Fonte: Tabela de Poisson.

**Resposta:**

A probabilidade de que sejam encontrados dois coelhos em um determinado ha. é de 0,1771, ou 17,71%.

2. O número médio de acidentes mensais em um determinado cruzamento é três. Qual é a probabilidade de que em um determinado mês ocorram quatro acidentes no cruzamento?

**Dados:**

$$x = 4$$

$$\mu = 3$$

$$P(x) = \frac{\mu^x e^{-\mu}}{x!}$$

$$P(4) = \frac{3^4 (2,71828)^{-3}}{4!} = \frac{81 \times 0,04978717}{4 \times 3 \times 2 \times 1} = \frac{4,03276068}{24} = 0,1680$$

**Resposta:**

A probabilidade de ocorrerem quatro acidentes em determinado mês é igual a 16,8%.



3. O número de chamadas telefônicas recebidas em uma empresa segue uma distribuição de Poisson, com média de 20 chamadas por hora. Qual é a probabilidade de, nos próximos 15 minutos, o telefone tocar no máximo duas vezes?

Nesse exemplo, há duas unidades de medida, ou seja, a informação média é dada em horas, mas na pergunta usa-se minutos. Por isso, inicialmente, é necessário transformar a média para a mesma unidade de medida. Assim, se em uma hora a empresa recebe 20 chamadas, em 15 minutos a empresa recebe em média cinco chamadas.

Além disso, o telefone tocar no máximo duas vezes significa que o telefone pode não tocar; tocar uma vez ou ainda tocar duas vezes nos próximos 15 minutos. Dessa forma, para solucionar o problema, precisa-se calcular a probabilidade individual de cada um desses eventos ocorrer e, ao final, somá-las.

Dados:

$$x \leq 2 \quad (P(0) + P(1) + P(2))$$

$$\mu = 5$$

$$P(x) = \frac{\mu^x e^{-\mu}}{x!}$$

$$P(0) = \frac{5^0 (2,71828)^{-5}}{0!} = \frac{1 \times 0,006737946}{1} = 0,0067$$

$$P(1) = \frac{5^1 (2,71828)^{-5}}{1!} = \frac{5 \times 0,006737946}{1} = 0,0337$$

$$P(2) = \frac{5^2 (2,71828)^{-5}}{2!} = \frac{25 \times 0,006737946}{2 \times 1} = 0,0842$$

$$P(x \leq 2) = 0,0067 + 0,0337 + 0,0842 = 0,1246$$

Resposta:

A probabilidade de o telefone tocar no máximo duas vezes em 15 minutos é igual a 12,46%.

## 8.4 Resumo para fixação

No sentido de auxiliar na memorização dos principais conceitos abordados no capítulo, é apresentado um resumo dos pontos mais relevantes:

- Uma distribuição de probabilidade de uma variável aleatória define os valores de probabilidade de todas as possibilidades de ocorrência dessa variável aleatória  $X$ .
- Quando a variável é aleatória, os valores de probabilidade são determinados por processos acidentais, ao acaso, que não estão sob o controle de nenhum observador.
- Para a distribuição binomial, em cada tentativa existem apenas dois resultados possíveis mutuamente exclusivos, chamados de sucesso e fracasso. Cada tentativa caracteriza-se por ser um evento independente, permanecendo a probabilidade de sucesso de ocorrência do evento constante de tentativa em tentativa, assim como a probabilidade de fracasso.
- A probabilidade de *sucesso* é atribuída a " $p$ ", e a probabilidade de *fracasso* ou insucesso, representada por um " $q$ ", é igual a  $(1 - p)$ .
- A probabilidade binomial determina a probabilidade de ocorrência de " $x$ " número de sucessos em " $n$ " número de experimentos do evento, com probabilidade de sucesso igual a " $p$ " e uma probabilidade de ocorrência de fracasso igual a " $q$ ".
- Nos cálculos de probabilidade é comum o uso de expressões como: "no mínimo", "no máximo", "pelo menos", "menos de", "mais de". Como essas expressões indicam que mais de um resultado é possível, faz-se necessário o cálculo da probabilidade de cada evento separadamente e depois se somam esses valores. Por exemplo, no máximo dois significa que o evento pode não ocorrer ( $P=0$ ), ocorrer uma única vez ( $P=1$ ) ou ainda ocorrer duas vezes ( $P=2$ ); pelo menos cinco significa que o evento pode ocorrer cinco vezes, seis vezes, ... e tantos quantos o problema permitir.
- Para poder aplicar a probabilidade binomial em situações de eventos dependentes, há a necessidade de haver reposição.
- A distribuição de Poisson é utilizada quando a variável aleatória está referida a uma unidade de exposição de tempo, de superfície ou de volume que possa assumir qualquer valor inteiro e positivo, ou seja, uma variável discreta, admitindo-se ser conhecido o valor médio da variável por unidade de exposição.
- Ela é utilizada para descrever o comportamento de eventos com pequenas probabilidades de ocorrência.
- A distribuição de Poisson é uma distribuição discreta de probabilidade de uma variável aleatória, cujo experimento consiste na contagem do número de vezes,  $x$ , que um evento ocorre em determinado intervalo, com igual probabilidade de ocorrência em cada intervalo, independentemente o número de ocorrências em um intervalo do número de ocorrências em outros intervalos.
- Aplica-se no fluxo de veículos, na previsão de acidentes, no uso de estoques por unidade de tempo, fluxo interno de pessoas em um ambiente, em qual não se tem o número total de elementos, pois o fluxo é contínuo, não sendo possível determinar o limite populacional.
- Em um cálculo de estoque, costuma-se utilizar Poisson, pois não há necessidade de haver a reposição de dados.

## 8.5 Exercícios propostos

8.1 Dez parafusos são escolhidos ao acaso da produção de certa máquina, que apresenta 5% de peças defeituosas.

- a) Qual é a probabilidade de serem defeituosos dois deles?
- b) Qual é a probabilidade de nenhuma peça ser defeituosa?
- c) Qual é a probabilidade de no máximo uma ser defeituosa?

8.2 Numa caixa com 10 lâmpadas, duas estão fundidas. Se forem retiradas dessa caixa uma amostra de quatro lâmpadas, determine a probabilidade de:

- a) nenhuma ser defeituosa.
- b) haver uma defeituosa.
- c) no máximo uma ser defeituosa.

8.3 Uma empresa garante, na embalagem de seu produto, que apenas 2% das peças produzidas por ela são defeituosas. Se forem adquiridas 12 peças produzidas por essa empresa, qual é a probabilidade de que:

- a) pelo menos duas sejam defeituosas?
- b) menos de duas sejam defeituosas?
- c) nenhuma seja defeituosa?
- d) no máximo duas sejam defeituosas?

8.4 Foram escolhidas aleatoriamente 20 unidades de um processo produtivo que apresenta 15% de unidades defeituosas.

- a) Determine a probabilidade do número de defeituosas ser menor que oito;
- b) Determine a probabilidade de o número de defeituosas ser maior que dois;
- c) Determine a probabilidade de o número de defeituosas ser igual a cinco.

8.5 Após a realização da primeira avaliação da disciplina de Estatística, a probabilidade de um aluno escolhido ao acaso ser reprovado na disciplina era igual a 15%. Em uma amostra formada por cinco alunos, calcule a probabilidade de:

- a) apenas um aluno ser reprovado;
- b) pelo menos um aluno ser reprovado.
- c) qual a média esperada de alunos reprovados?
- d) qual o desvio-padrão do número de alunos reprovados?

8.6 Num dado posto de vendas apurou-se que a procura de um certo artigo é uma variável aleatória de Poisson de valor médio igual a 5 unidades diárias. Qual a probabilidade de num dia a procura ser de:

- a) até duas unidades;
- b) nula.

8.7 Suponha que os clientes entrem num armazém à média de 18 por hora. Usando adequadamente a distribuição de Poisson determine a probabilidade de que num intervalo de cinco minutos não entre ninguém no armazém.

8.8 Suponha que os defeitos em fios para tear possam ser aproximados por um processo de Poisson com média de 0,2 defeitos por metro. Inspeccionando-se pedaços de fio de seis metros de comprimento, determine a probabilidade de se encontrarem menos de dois defeitos.

8.9 O número de automóveis que atravessam uma ponte durante um determinado período de tempo é uma v.a. (variável aleatória) com distribuição de Poisson. Em média passam cinco carros em 10 minutos.

- a) Qual a probabilidade de não passar nenhum carro em uma hora?
- b) Qual a probabilidade de, em 20 minutos, atravessarem a ponte menos de quatro automóveis?

## 9 DISTRIBUIÇÃO BÁSICA DE PROBABILIDADE DE VARIÁVEIS CONTÍNUAS

Partindo da ideia de organização de intervalos de classe de variáveis contínuas, com determinada amplitude e com um número muito grande de dados, com a redução dessa amplitude e aumento do número de intervalos progressivamente, pode-se chegar a uma situação teórica de um número infinitamente grande de intervalos e amplitudes infinitamente pequenas. Dessa maneira, forma-se uma linha contínua com a forma aproximada de um sino.

Essa curva está ligada à história da descoberta das probabilidades em matemática, no século XVII. O responsável mais direto dessa curva, chamada normal, foi Abraham de Moivre, que a definiu em 1730, a partir dos trabalhos de Jacob e Nicolaus Bernoulli, da Lei dos Grandes Números. Moivre publicou os seus trabalhos em 1773 na obra *The Doctrine of chances*. Assim, Laplace em 1783 a utilizou para descrever a distribuição dos erros e Gauss, 1809, a empregou para analisar dados astronômicos.

Posteriormente Karl Pearson propôs o termo Distribuição Normal para a curva normal desenvolvida por Moivre, na área da Estatística, e o nome de Curva de Gauss foi dado em homenagem a K. F. Gauss, que é aplicada em Matemática.

Atualmente a curva normal representa um ganho para a ciência, pois a normalidade ocorre naturalmente em muitas medidas de situações físicas, biológicas e sociais e é utilizada para fazer inferências estatísticas.

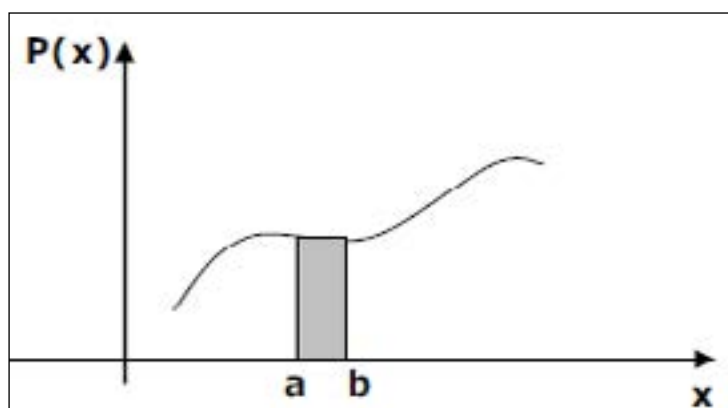
Assim, os modelos contínuos encontram importantes aplicações na engenharia, nas ciências físicas, em finanças e ciências sociais. Alguns exemplos de fenômenos aleatórios contínuos são a altura, peso, tempo, como, por exemplo, tempo necessário de atendimento a clientes, ou na execução de tarefas etc.

### 9.1 Função densidade de probabilidade

Como uma variável aleatória contínua pode assumir qualquer valor fracionário dentro de um intervalo definido de valores, não é possível enumerar todos os possíveis valores de uma variável aleatória contínua com os valores de probabilidades correspondentes, como acontece na variável discreta. Em lugar disso, a abordagem mais conveniente é construir uma função densidade de probabilidade, ou curva de probabilidade, baseada na função matemática correspondente. Assim, tendo-se disponível uma expressão matemática para representar algum fenômeno contínuo, pode-se calcular a probabilidade de que vários valores da variável aleatória ocorram em certos limites ou intervalos.

Nesse caso, a proporção da área incluída entre dois pontos quaisquer, debaixo da curva de probabilidade, identifica a probabilidade de que a variável aleatória contínua selecionada assumira um valor entre tais pontos (FIGURA 9.1).

Figura 9.1 – Proporção da área entre os pontos a e b, debaixo de uma curva de probabilidade

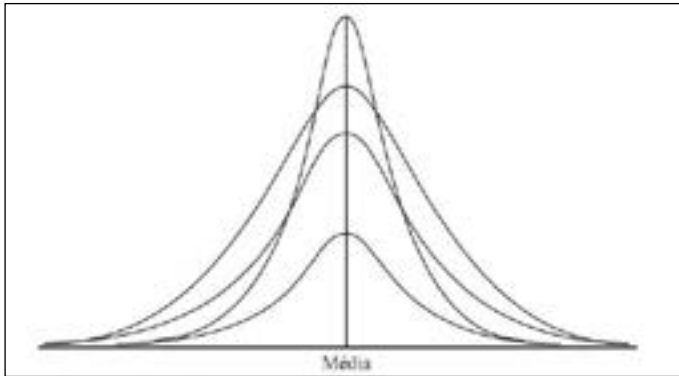


Fonte: Dos autores.

## 9.2 Características da Distribuição Normal

A curva normal é definida como sendo simétrica e unimodal. Essa característica encontra-se na natureza quando o número de dados do universo analisado é relativamente grande e principalmente com uma variável contínua. Nesse caso, a distribuição dos valores acontece em uma curva em forma de um sino, com um ponto máximo no centro, em que as áreas, em ambos os lados da média, são idênticas (FIGURA 9.2).

Figura 9.2 – Curvas simétricas unimodais, com diversas características de curtose



Fonte: Dos autores.

Essa situação simétrica é estabelecida porque os valores da média, mediana e moda são iguais. Como nem sempre essa situação acontece exatamente dessa forma, comumente usa-se também a expressão de distribuição aproximadamente normal, que se caracteriza por pequenas deformações, em que as medidas da média, moda e mediana não são mais iguais, mas com valores muito próximos.

Para facilitar o uso da distribuição Normal, trabalha-se com uma curva normal padronizada, que, além das características já apresentadas, ainda assume que a curtose seja mesocúrtica, ou seja, com um pico mediano, nem muito empinado nem achatado.

Uma curva normal padronizada baseia-se em parâmetros automaticamente definidos para qualquer escala de medida que for utilizada, sendo a média sempre zero e a variância igual a um. Para isso existem tabelas construídas para essa curva que mostram o percentual da população que se encontra em determinada faixa.

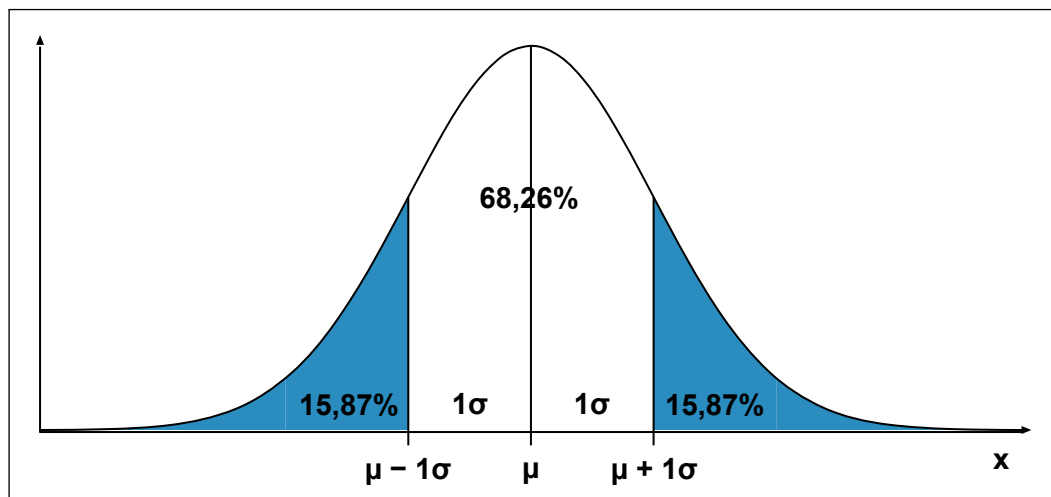
Considerando essas características, a distribuição Normal é importante na estatística pelos seguintes aspectos:

- inúmeros fenômenos contínuos parecem segui-la ou poderem ser aproximados por meio dela;
- pode também ser usada como aproximação para distribuições de probabilidade discreta, em situações específicas;
- oferece a base para a inferência estatística clássica devido à sua afinidade com o teorema do limite central<sup>3</sup>.
- em termos gerais, as características de uma Distribuição Normal padronizada podem ser resumidas da seguinte maneira:
- em termos de aparência, a curva é simétrica e tem o formato de um sino;
- a reta que passa pelo ponto máximo divide a área da curva em duas partes exatamente iguais, em forma e tamanho;

3 O Teorema do limite central indica que a distribuição das médias oriundas de amostras tende para uma distribuição normal, cuja média é igual à média da população. A diferença entre a média da amostra e a verdadeira média da população é conhecida como erro padrão, que se reduz à medida que se aumenta o tamanho da amostra e chega a zero quando o tamanho da amostra coincide com o tamanho da população.

- as medidas de tendência central (média, mediana e moda) são idênticas;
- a dispersão média é igual a 1,33 desvio-padrão. Isto significa que o intervalo interquartilístico está contido dentro de um intervalo de dois terços de um desvio-padrão, abaixo da média aritmética e dois terços de um desvio-padrão, acima da média;
- a área debaixo da curva é igual a 1 ou 100%;
- os valores encontrados abaixo da curva da distribuição Normal padronizada podem ser medidos em desvios-padrão e suas probabilidades por intervalo são as seguintes (FIGURA 9.3):
  - +  $1\sigma$  a  $-1\sigma$  = 68,27% da área da curva;
  - +  $2\sigma$  a  $-2\sigma$  = 95,44% da área da curva;
  - +  $3\sigma$  a  $-3\sigma$  = 99,73% da área da curva.
- a curva apresenta em ambos os lados um ponto de inflexão que está a uma distância de  $(-1)$  desvio-padrão (s) e  $(+1)$  desvio-padrão (s) da origem do centro da curva, onde a média, a mediana e a moda são coincidentes;
- na curva, ambos os lados acercam-se cada vez mais ao eixo das abcissas, porém sem jamais tocá-lo;
- sua variável aleatória associada possui um intervalo infinito  $(-\infty < X < +\infty)$ .

Figura 9.3 – Curva Normal com o primeiro ponto de inflexão e as suas respectivas probabilidades



Fonte: Dos autores.

Já que na prática, às vezes, a distribuição das variáveis não chega a essa perfeição, aproximando-se somente a essas propriedades teóricas, pode-se usá-la na condição de ser uma Distribuição Aproximadamente Normal, desde que apresente os seguintes aspectos:

- a curva deve ter o formato de um sino e com aparência simétrica;
- suas medidas de tendência central (média, moda e mediana) podem divergir ligeiramente;
- o valor do intervalo interquartilístico pode diferir ligeiramente de 1,33 desvio-padrão;
- o intervalo da variável na prática não ser infinito, mas deve estar entre três desvios padrões acima e abaixo da média aritmética.

### 9.3 Fórmula de cálculo da Função densidade

A função densidade de probabilidade representa um modelo ou expressão matemática pelo símbolo  $f(x)$ . Essa função foi desenvolvida pelo matemático Gauss que expressa uma curva simétrica da Distribuição Normal, também conhecida como curva de Gauss.

Para definir os diferentes pontos que permitem traçar a curva, a função é definida pela seguinte expressão matemática:

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X_i - \mu}{\sigma}\right)^2} \quad (9.1)$$

onde:

$f(X)$  = diferentes valores da curva determinados pela expressão

$X_i$  = qualquer valor da variável aleatória contínua, em que  $-\infty < X_i < +\infty$

$\sigma$  = desvio-padrão da população

$\mu$  = média da população

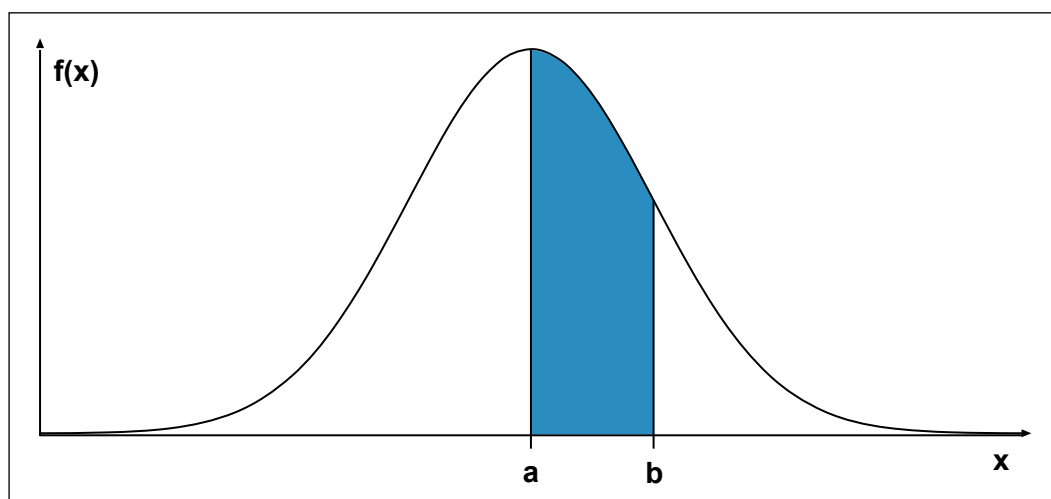
$\pi$  = constante matemática com valor aproximado de 3,14159

$e$  = constante matemática com valor aproximado de 2,71828

Considerando que os símbolos  $\pi$  e  $e$  são constantes matemáticas, para cada valor da variável aleatória  $X_i$ , existem probabilidades diferentes, pois os valores resultantes da função são dependentes dos parâmetros da distribuição normal: a média aritmética e o desvio-padrão da população.

Por meio da função, podem-se determinar as áreas sob qualquer curva normal, entre dois pontos dados na abscissa, por meio de cálculo integral, mas seria trabalhoso (FIGURA 9.4).

Figura 9.4 – Área sob a curva normal entre os pontos a e b



Fonte: Dos autores.

Além disso, o tabelamento dos valores das áreas para qualquer curva normal não seria viável, pois há um número infinito de tais curvas, pelas infinitas possibilidades de combinação de valores da variável  $X$  com diferentes parâmetros  $\mu$  e  $\sigma$ .

Dessa forma, torna-se difícil fazer o cálculo de probabilidades com o modelo original acima apresentado.

Resolveu-se então construir uma curva normal padronizada e tabelar os valores de suas áreas de modo que as áreas entre dois pontos na abscissa definidos, sob qualquer curva, possam ser calculadas.

Assim, foi feita a padronização dos dados, substituindo a expressão  $\frac{X_i - \mu}{\sigma}$  por  $z$ , ou seja:

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



Em que  $z$  representa uma variável reduzida padronizada que tem (média)  $m = 0$  e (desvio-padrão)  $s = 1$ .

Assim, o valor de  $z$  é igual à diferença entre o valor de  $X_i$  e a média aritmética  $m$ , dividida pelo desvio-padrão  $s$ .

Dessa forma, introduzindo a variável  $z$  em substituição à expressão e considerando que o desvio-padrão da variável reduzida  $z$  é igual a 1, a função densidade de probabilidade de uma variável normal padronizada pode ser definida assim:

$$f(X) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad (9.2)$$

onde:

$f(X)$  = diferentes valores da curva determinados pela equação

$z$  = variável reduzida padronizada

$\pi$  = constante matemática com valor aproximado de 3,14159

$e$  = constante matemática com valor aproximado de 2,71828

Utilizando a expressão, pode-se converter qualquer conjunto de dados distribuídos de maneira normal para sua forma padronizada e determinar quaisquer probabilidades desejadas a partir de uma tabela da distribuição normal padronizada.

#### 9.4 Determinação da probabilidade da Distribuição Normal padronizada

Definida a função normal padronizada, que apresenta como característica uma curva simétrica e mesocúrtica, podem-se determinar valores de probabilidade e apresentá-los em tabelas que expressam os valores da função densidade ou curva de probabilidade, com base na variável padronizada  $z$ .

Assim, há apenas a necessidade de converter valores de  $X_i$ , oriundos de uma população normalmente distribuída, em valores normais padronizados  $z$  pelo uso da fórmula:

$$z = \frac{X_i - \mu}{\sigma} \quad (9.3)$$

onde:

$X_i$  = qualquer valor da variável aleatória contínua.

$\sigma$  = desvio-padrão da população

$\mu$  = média da população

$z$  = variável reduzida padronizada

Depois de calculado o valor de  $z$ , procura-se na tabela normal padronizada o respectivo valor da probabilidade.

Considerando que a tabela baseia-se em valores de  $z$ , o seu valor deve ser calculado com dois dígitos após a parte inteira e o resultado deve ser dividido em duas partes: a parte inteira, juntamente com o primeiro dígito após a vírgula, deve ser procurada na primeira coluna e o segundo dígito calculado deve ser procurado na linha superior da tabela do Apêndice III. No corpo da tabela, no encontro da linha e da coluna selecionadas, encontra-se o resultado da probabilidade.

Assim, considerando, como exemplo, o valor de  $z = 1,03$ , ele está formado por 1,0, que deve ser selecionado na primeira coluna, e pelo valor 0,03, que se encontra na linha superior. O valor da respectiva probabilidade para esse valor de  $z$  será encontrado na conjugação da linha que contém o valor de 1,0 com a coluna que contém o valor de 0,03, cujo valor é 0,3485, ou 34,85% (TABELA 9.1a).



Tabela 9.1a - APÊNDICE III - Distribuição Normal Padronizada - z

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
4,5	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Fonte: Tabela construída com as funções estatísticas do EXCEL pelos autores.

O uso da fórmula padronizada  $z = \frac{X_i - \mu}{\sigma}$  permite encontrar várias soluções distintas, entre as quais as principais são as seguintes:

- determinação da **probabilidade** em função de  $X_i$ , sendo a média e o desvio-padrão conhecidos;
- determinação do valor de  $X_i$  em função da **probabilidade desejada**, conhecendo a média e o desvio-padrão.

#### 9.4.1 Determinação da probabilidade em função de $X_i$ , sendo a média e o desvio-padrão conhecidos

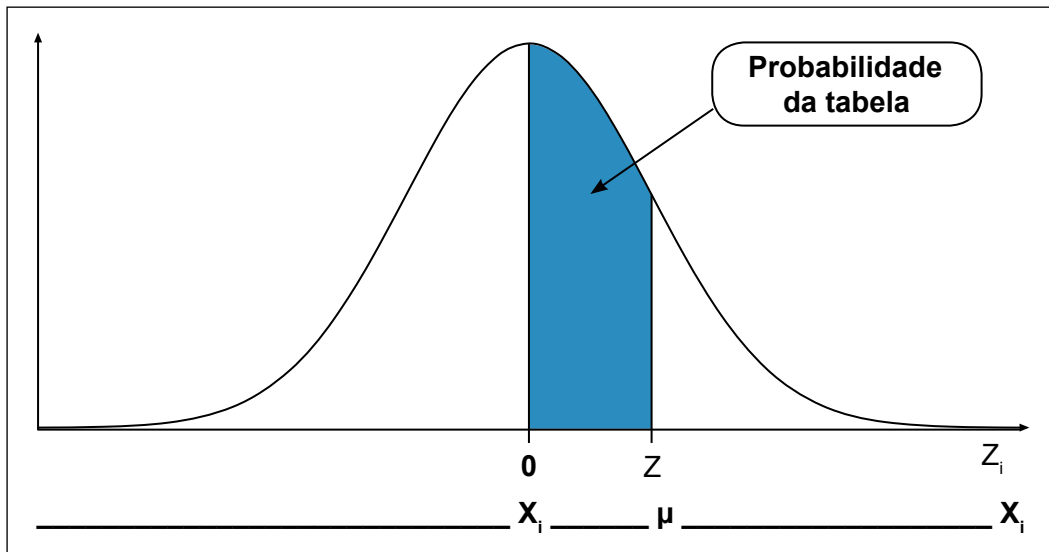
Para encontrar o valor de probabilidade entre dois pontos, é necessário conhecer os valores da média e do desvio-padrão da população da qual foi extraído o valor da variável aleatória para a qual se quer determinar a probabilidade. Com base nesses valores, determina-se o cálculo do valor de z, com dois dígitos além da parte inteira, calculados pela fórmula 9.3:

$$z = \frac{X_i - \mu}{\sigma}$$

Em seguida, procuram-se na tabela do Apêndice III os respectivos valores da probabilidade.

Em sendo o valor do z positivo, os valores de probabilidade encontrados na tabela referem-se ao lado positivo da curva (lado direito), contados a partir do centro, conforme Figura 9.5.

Figura 9.5 – Valores de probabilidade entre dois pontos, a partir do centro, no lado direito da curva normal

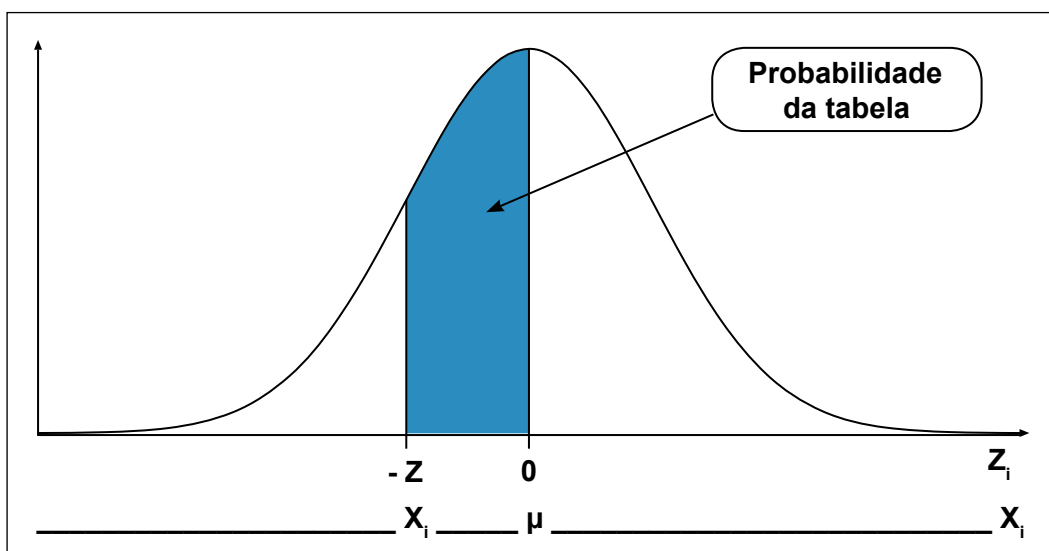


Fonte: Dos autores.

As probabilidades encontradas na tabela sempre se referem ao intervalo formado pelo valor da média populacional e o valor de  $X_i$  aleatório, usado no cálculo, isto é, entre os valores usados no numerador da fórmula de cálculo ( $X_i - \mu$ ).

Quando o valor de  $z$  calculado for um valor negativo, significa que o resultado da probabilidade procurada localiza-se no lado esquerdo da curva. A obtenção do valor na tabela não difere da situação anterior (valor positivo), uma vez que a curva é simétrica, sendo apenas ignorado o sinal negativo do  $z$  para tal. A probabilidade encontrada na tabela refere-se à área da curva da direita para a esquerda, a partir do seu centro, conforme Figura 9.6.

Figura 9.6 – Valores de probabilidade entre dois pontos, a partir do centro, no lado esquerdo da curva normal



Fonte: Dos autores.

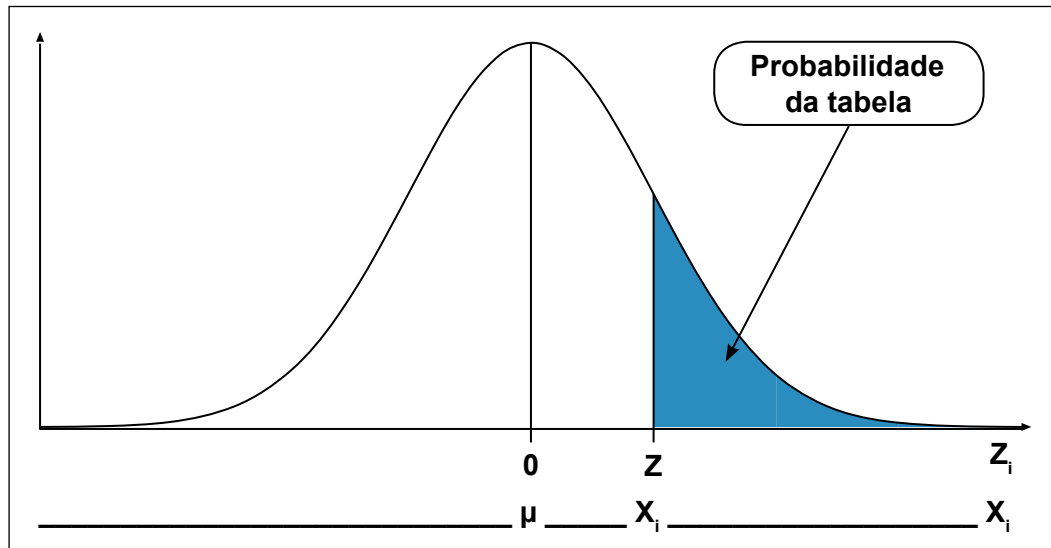
Na solução de problemas de probabilidade, existem diversas alternativas que não se resolvem diretamente com os valores encontrados na tabela.

Assim, quando se quer determinar o valor das probabilidades localizadas nos extremos da curva, tanto do lado direito como do lado esquerdo, conforme Figuras 9.7 e 9.8, respectivamente,

deve-se subtrair o resultado encontrado na tabela do valor de 0,5 que representa a metade simétrica da curva.

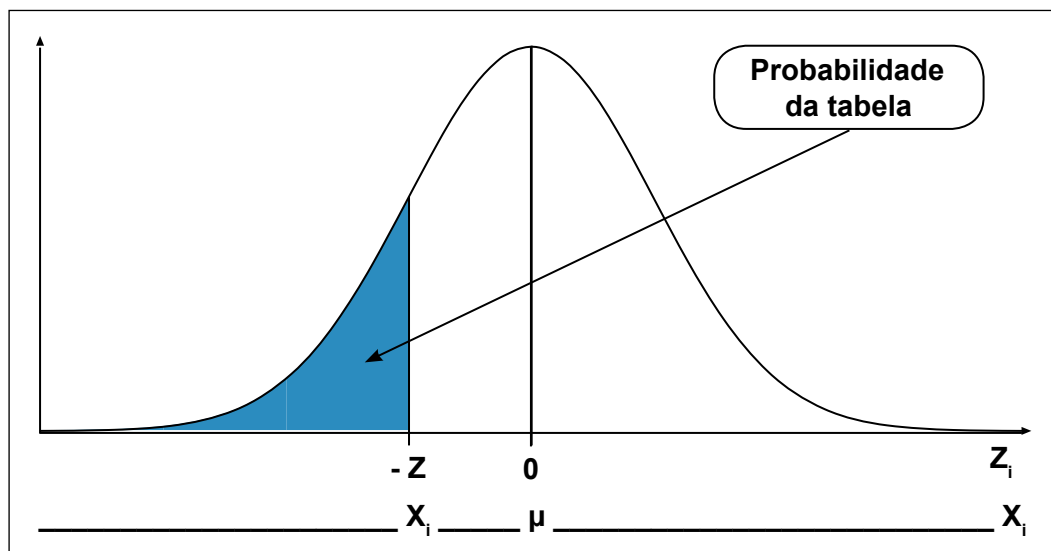
Nesse caso, é sempre importante que se faça o esboço da curva, destacando a área da probabilidade desejada, para determinar a resposta adequada, a partir da probabilidade encontrada na tabela.

Figura 9.7 – Valores de probabilidade entre dois pontos, localizados nos extremos da curva, no lado direito



Fonte: Dos autores.

Figura 9.8 – Valores de probabilidade entre dois pontos, localizados nos extremos da curva, no lado esquerdo

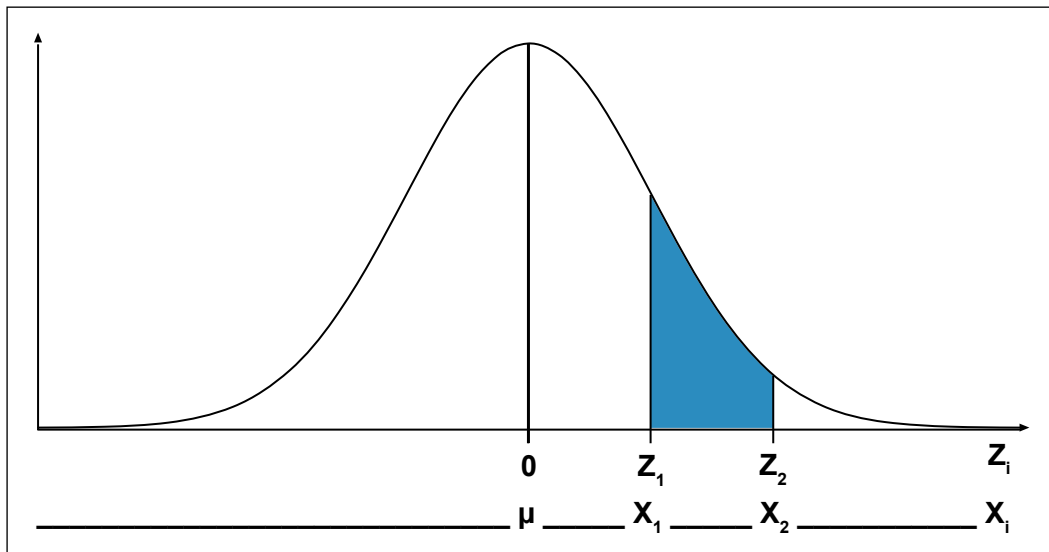


Fonte: Dos autores.

Além das quatro alternativas apresentadas, existem ainda duas outras possibilidades: a) probabilidade entre dois pontos quaisquer situados em um dos lados da curva, que não incluem o centro e também não o extremo da curva e, b) a probabilidade que reúne pontos situados à direita e à esquerda do centro da curva.

Para a determinação da probabilidade entre dois pontos quaisquer, que não incluem o elemento central nem o extremo, em um lado da curva, deve-se fazer o cálculo do valor de  $z$ , em duas etapas (FIGURA 9.9).

Figura 9.9 – Valores de probabilidade entre dois pontos quaisquer, que não incluem o elemento central, localizados em um dos lados da curva



Fonte: Dos autores.

Inicialmente usa-se o valor  $X_i$  do primeiro ponto para o cálculo do valor de  $Z_1$ , com dois dígitos além da parte inteira, calculado pela fórmula 9.3:

$$z = \frac{X_i - \mu}{\sigma}$$

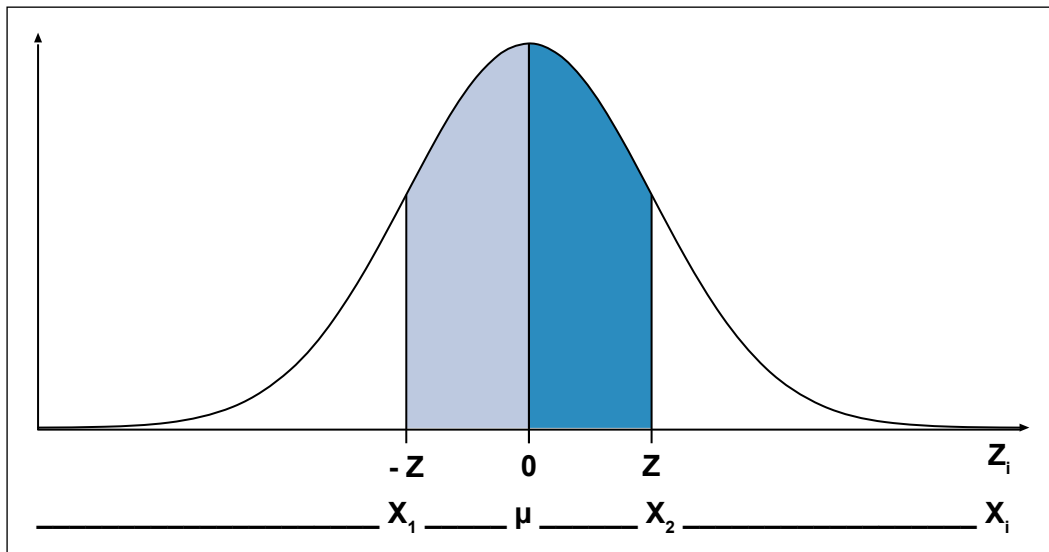
Calculado o valor de  $Z_1$ , procura-se na tabela do Apêndice III, respectivo valor da probabilidade.

Após, calcula-se o valor de  $Z_2$ , usando para isso o valor de  $X_2$  do segundo ponto, e procuram-se o respectivo valor da probabilidade também na tabela.

Como as probabilidades encontradas na tabela sempre se referem ao intervalo formado pelo valor da média populacional e o valor do respectivo  $X_i$  usado no cálculo, ou seja, entre os valores usados no numerador da fórmula de cálculo ( $X_i - \mu$ ), para encontrar o resultado da probabilidade entre dois pontos quaisquer, que não incluem o elemento central, subtrai-se a probabilidade menor da maior. Assim  $P = \text{probabilidade } Z_2 - \text{probabilidade } Z_1$

Da mesma forma, também para a determinação da probabilidade entre dois pontos situados à direita e à esquerda do centro da curva, deve-se fazer o cálculo do valor de  $z$ , em duas etapas (FIGURA 9.10).

Figura 9.10 – Valores de probabilidade entre dois pontos situados à direita e à esquerda do centro da curva



Fonte: Dos autores.

Assim, inicialmente usa-se o valor  $X_i$  do primeiro ponto para o cálculo do valor de  $Z_1$ , com dois dígitos além da parte inteira, calculados pela fórmula 9.3:

$$z = \frac{X_i - \mu}{\sigma}$$

Calculado o valor de  $Z_1$ , procura-se na tabela do Apêndice III, respectivo valor da probabilidade.

Após, calcula-se o valor de  $Z_2$ , usando para isso o valor de  $X_i$  do segundo ponto, e procuram-se na tabela do Apêndice III o respectivo valor da probabilidade.

Como as probabilidades encontradas na tabela sempre se referem ao intervalo formado pelo valor da média populacional e o valor do respectivo  $X_i$  usado no cálculo, ou seja, entre os valores usados no numerador da fórmula de cálculo ( $X_i - \mu$ ), para encontrar o resultado da probabilidade entre dois pontos situados um à direita e outro à esquerda do centro da curva, somam-se as duas probabilidades.

Para elucidar dúvidas e propiciar melhor entendimento de como são solucionados problemas de determinação de probabilidade são apresentados vários exemplos a seguir.

Exemplos:

1. Uma estante tem pacotes de diferentes pesos. O peso médio desses pacotes é de 100 gramas e o desvio-padrão é de 10 gramas. Sabe-se que o peso dos pacotes segue uma distribuição normal. Qual é a probabilidade de se obter “ao azar” um pacote que tenha mais de 115 gramas?

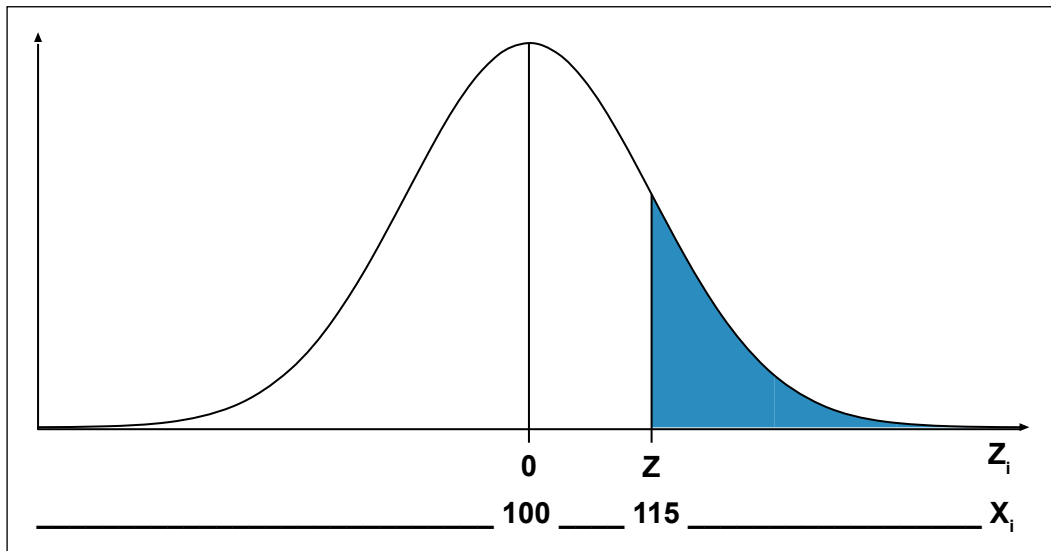
Dados:

$$\mu = 100 \text{ gramas}$$

$$\sigma = 10 \text{ gramas}$$

$$P(x \geq 115 \text{ gramas})?$$

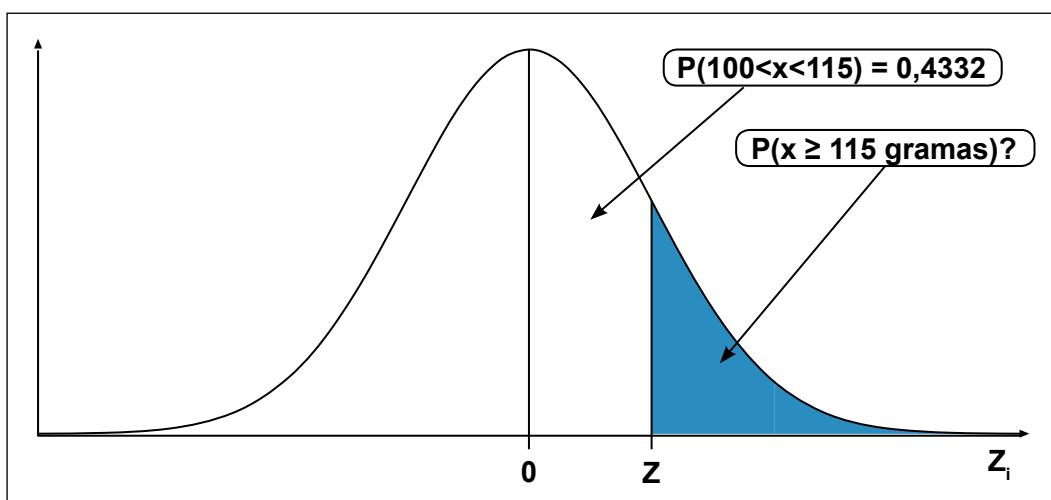
Representação gráfica do problema da probabilidade procurada:



Cálculo do valor de  $z$ :

$$z = \frac{X_i - \mu}{\sigma} = \frac{115 - 100}{10} = 1,50$$

Probabilidade encontrada na tabela com o valor de  $z = 1,50$ :



Solução para encontrar a probabilidade desejada no problema:

$$P(x \geq 115 \text{ gramas}) = 0,5 - 0,4332 = 0,0668$$

Resposta:

A probabilidade de encontrar um pacote com peso acima de 115 gramas é de 6,68%.

2. O peso médio de 500 estudantes do sexo masculino é 75 Kg. Admite-se que essa medida tem distribuição normal, com desvio-padrão igual a 6 Kg. Determine a probabilidade de encontrar um estudante com:

- peso menor de 69 kg;
- peso entre 69 e 81 Kg;
- número de estudantes com peso menor de 69 Kg;
- número de estudantes com peso entre 69 e 81 Kg.

Solução de a:

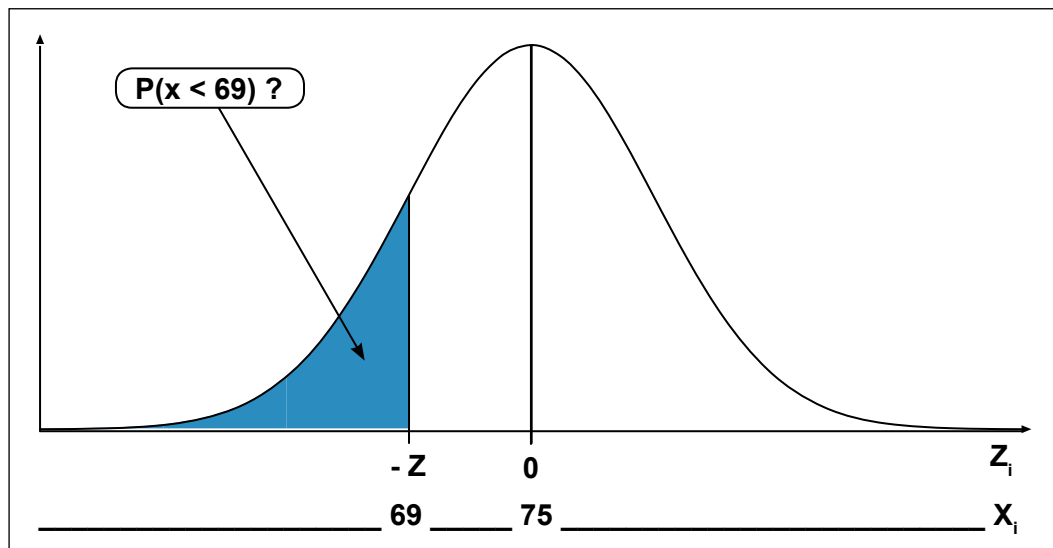
Dados:

$$\mu = 75 \text{ Kg}$$

$$\sigma = 6 \text{ Kg}$$

$$P(x < 69)?$$

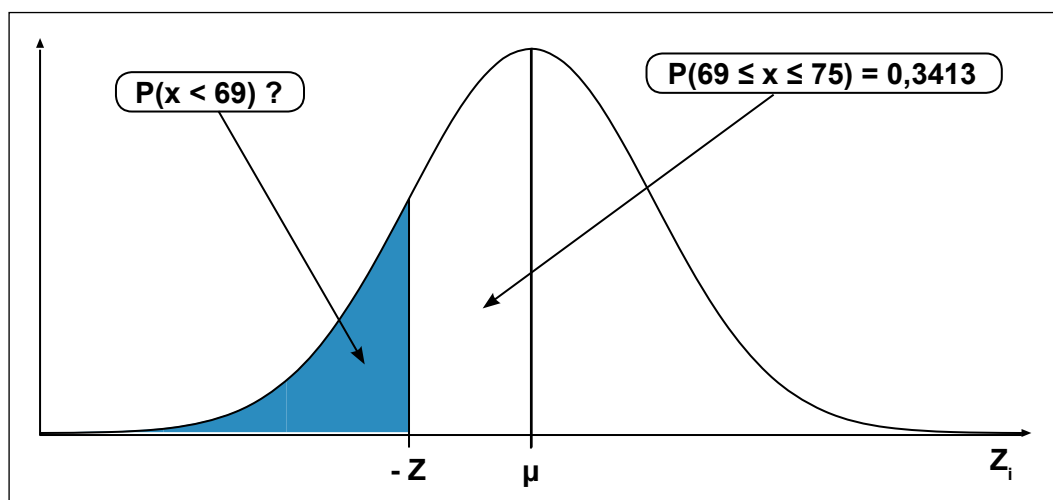
Representação gráfica do problema da probabilidade procurada:



Cálculo do valor de z:

$$z = \frac{X_i - \mu}{\sigma} = \frac{69 - 75}{6} = -1,00$$

Probabilidade encontrada na tabela com o valor de  $z = 1,00$ :



Solução para encontrar a probabilidade desejada no problema:

$$P(x < 69 \text{ Kg}) = 0,5 - 0,3413 = 0,1587$$

Resposta:

A probabilidade de encontrar um estudante com peso menor a 69 Kg é de 15,87%.

Solução de b:

Dados:

$$\mu = 75 \text{ Kg}$$

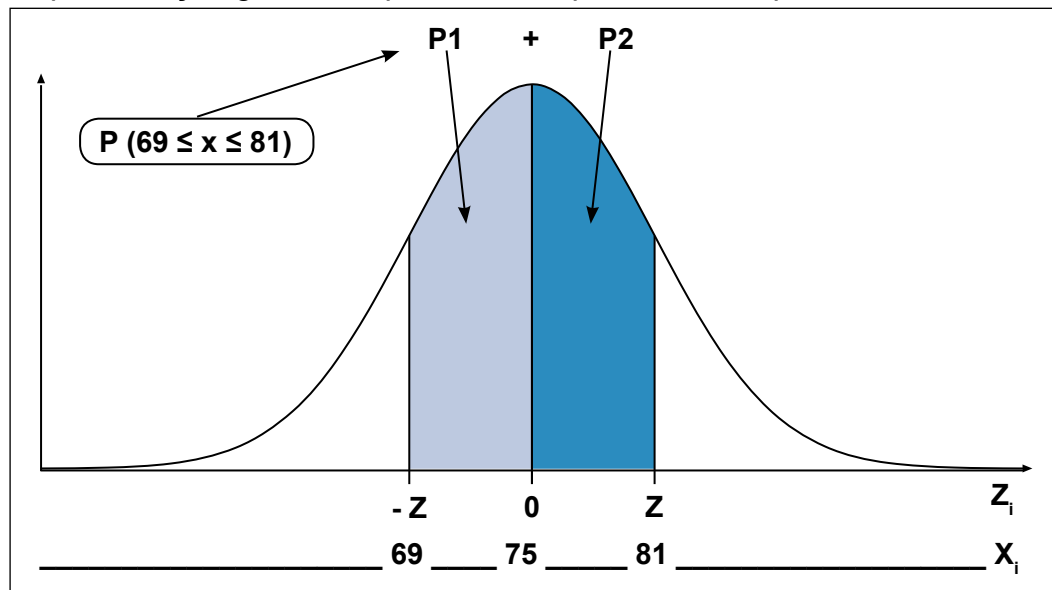
$$\sigma = 6 \text{ Kg}$$

$$X_1 = 69$$

$$X_2 = 81$$

$$P(69 \leq x \leq 81)$$

Representação gráfica do problema da probabilidade procurada:

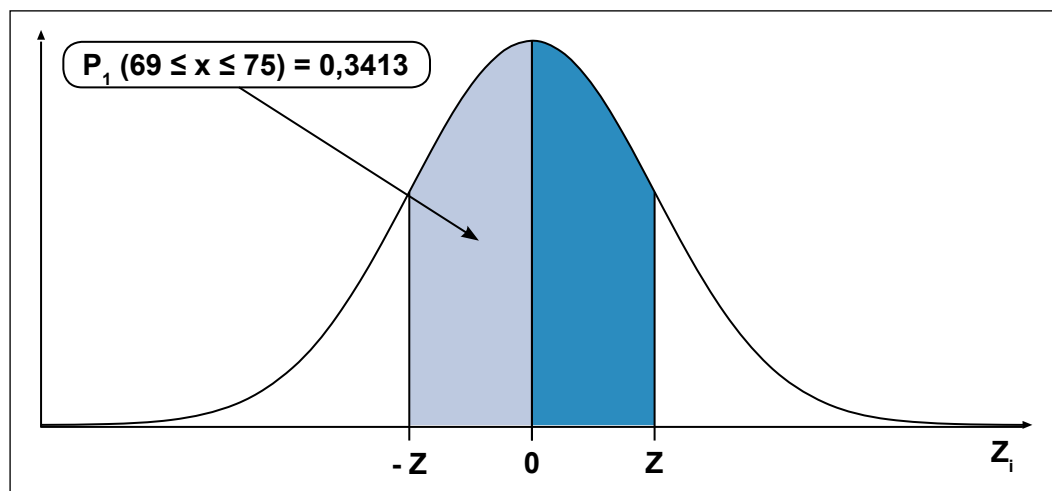


Para encontrar a solução do problema, deve-se fazer o cálculo do valor de  $z$ , em duas etapas. Primeiro calcular o valor de  $Z_1$  e encontrar a respectiva probabilidade e depois calcular o valor de  $Z_2$  e também encontrar a sua probabilidade.

Cálculo do valor de  $z_1$

$$z = \frac{X_i - \mu}{\sigma} = \frac{69 - 75}{6} = -1,00$$

Probabilidade encontrada na tabela com o valor de  $Z_1 = 0,3413$ :

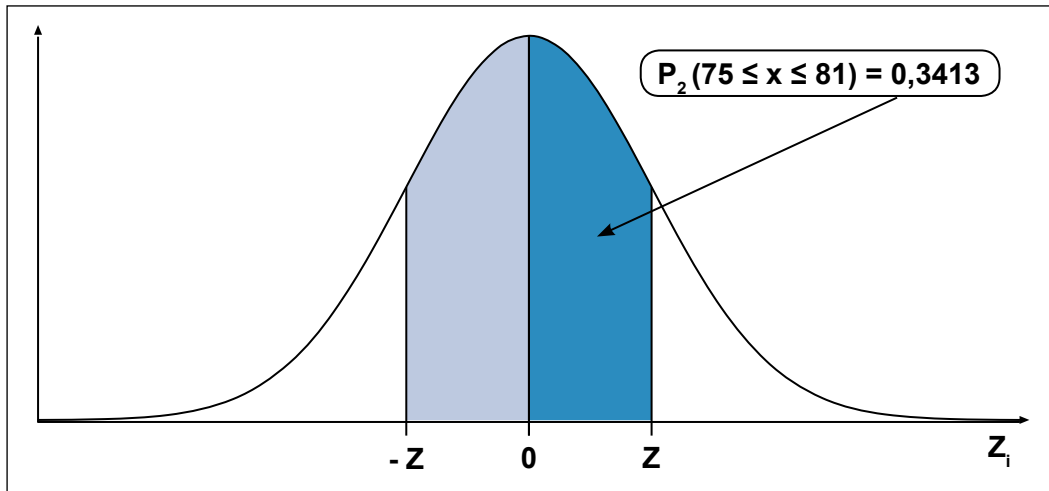




Cálculo do valor de  $z_2$ :

$$z = \frac{X_i - \mu}{\sigma} = \frac{81 - 75}{6} = 1,00$$

Probabilidade encontrada na tabela com o valor de  $Z_2 = 0,3413$ :



Solução para encontrar a probabilidade desejada no problema:

Para encontrar a resposta do problema proposto, devem-se somar as duas probabilidades:

Assim,

$$P(69 \leq x \leq 81) = 0,3413 + 0,3413 = 0,6826$$

Resposta:

A probabilidade de encontrar um estudante com peso maior ou igual a 69 Kg e menor ou igual a 81 Kg é de 68,26%.

Solução de c:

Dados:

$N = 500$  (número de estudantes da população)

$P(x < 69) = 15,87\%$  (resposta da questão a)

$n = ?$  (Número esperado de estudantes com peso igual ou menor de 69 Kg.)

Solução:

Para transformar o percentual da probabilidade em número de unidades, multiplica-se o total de unidades da população pela probabilidade calculada, ou seja,  $N \times P$ .

Assim,

$$n = 500 \times 0,1587 = 79$$

Resposta:

O número esperado de estudantes com peso menor ou igual a 69 Kg é de 79 estudantes.

Solução de d:

Dados:

$N = 500$  (número de estudantes da população)

$P(69 \leq x \leq 81) = 68,26\%$  (resposta da questão b)

$n = ?$  (Número esperado de estudantes com peso maior ou igual a 69 Kg e menor ou igual a 81 Kg.)

Solução:

Para transformar o percentual da probabilidade em número de unidades, multiplica-se o total de unidades da população pela probabilidade calculada, ou seja,  $N \times P$ .

Assim,  $n = 500 \times 0,6826 = 341$

Resposta:

O número esperado de estudantes com peso maior ou igual a 69 Kg e menor ou igual a 81 Kg é de 341 alunos.

3. Os pesos dos estudantes de uma universidade são distribuídos normalmente com média de 65,3 e desvio-padrão de 5,5 Kg.

- Determine o percentual de estudantes que pesam entre 70 e 75 Kg.
- Determine o percentual de estudantes que pesam até 70 Kg.

Solução da questão a:

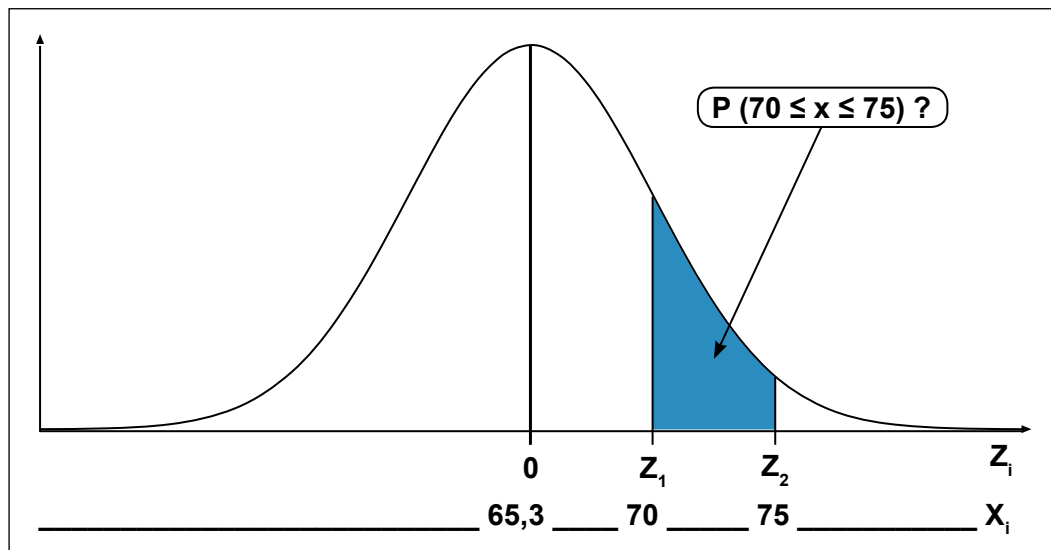
Dados:

$$\mu = 65,3 \text{ Kg}$$

$$\sigma = 5,5 \text{ Kg}$$

$$P(70 \leq x \leq 75)?$$

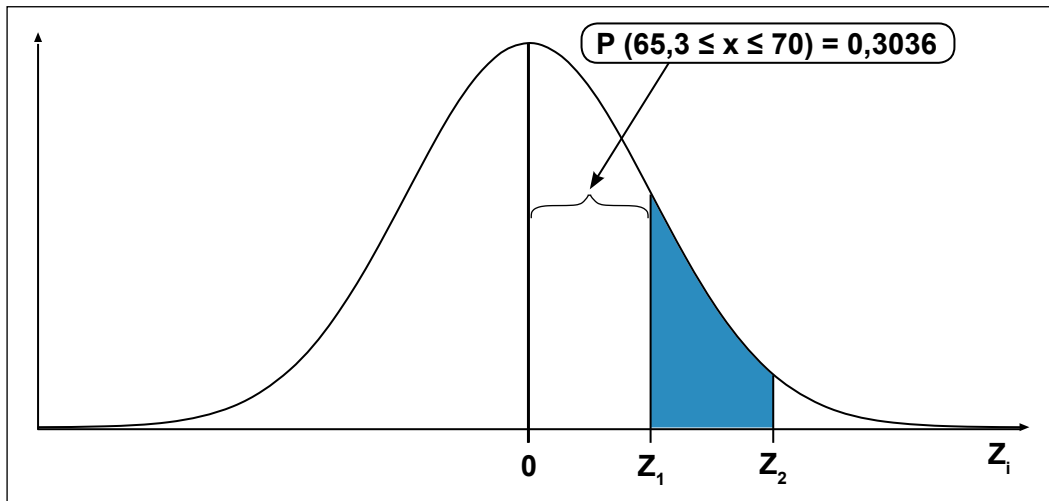
Representação gráfica do problema da probabilidade procurada:



Cálculo do valor de  $z_1$ :

$$z_1 = \frac{X_i - \mu}{\sigma} = \frac{70 - 65,3}{5,5} = 0,85$$

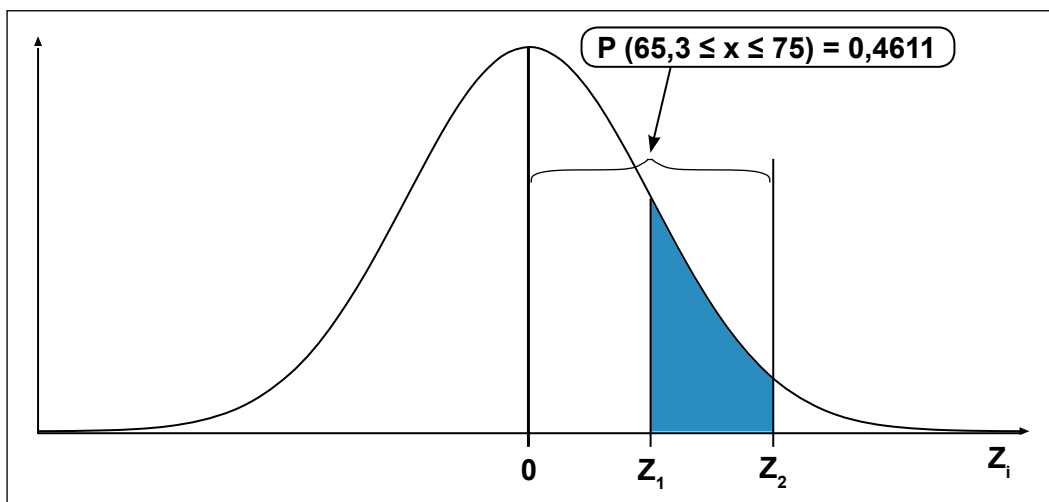
Probabilidade encontrada na tabela com o valor de  $Z_1 = 0,3036$ :



Cálculo do valor de  $Z_2$ :

$$z_1 = \frac{X_i - \mu}{\sigma} = \frac{75 - 65,3}{5,5} = 1,76$$

Probabilidade encontrada na tabela com o valor de  $Z_2 = 0,4611$ :



Solução para encontrar a probabilidade desejada no problema:

$$P(70 \leq x \leq 75) = 0,4611 - 0,3036 = 0,1575$$

Resposta:

A probabilidade de encontrar estudantes com peso entre 70 e 75 Kg é de 15,75%.

Solução da questão b:

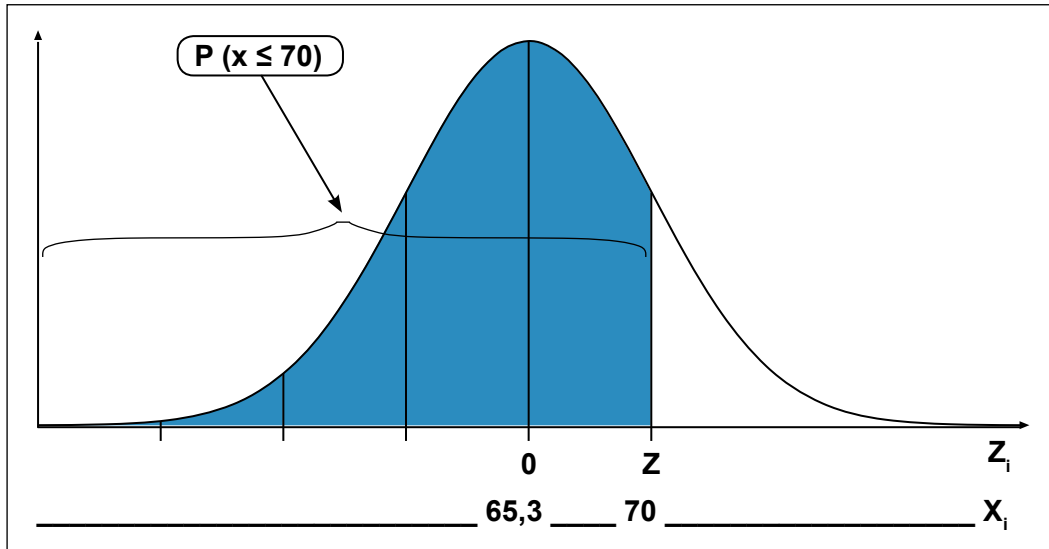
Dados:

$$\mu = 65,3 \text{ Kg}$$

$$\sigma = 5,5 \text{ Kg}$$

$$P(x \leq 70)?$$

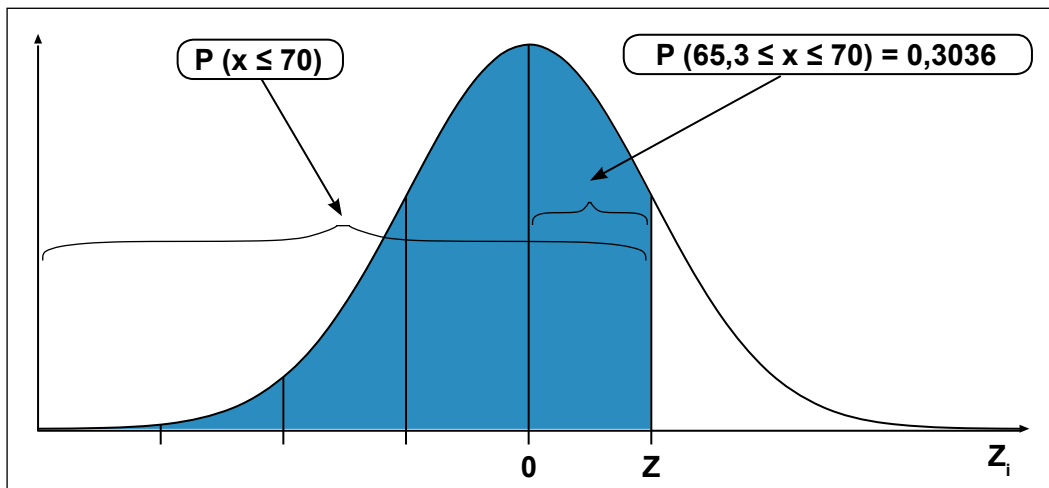
Representação gráfica do problema da probabilidade procurada:



Cálculo do valor de z:

$$z_1 = \frac{X_i - \mu}{\sigma} = \frac{70 - 65,3}{5,5} = 0,85$$

Probabilidade encontrada na tabela com o valor de  $z_1 = 0,85$ :



Solução para encontrar a probabilidade desejada no problema:

Como a questão considera o peso até 70 Kg, devem ser considerados todos os estudantes que tenham entre o menor peso até os com 70 Kg. Assim, na solução do problema devem ser consideradas a probabilidade dos estudantes com peso abaixo da média, ou seja, menos de 65,3 Kg, que é a metade da curva à esquerda, que representa 50%, e a parte dos estudantes que têm peso acima da média até 70 Kg.

Dessa forma,

$$P(x \leq 70) = 0,50 + 0,3036 = 0,8036$$

Resposta:

A probabilidade de encontrar estudantes com peso até 70 Kg é de 80,36%.

### 9.4.2 Determinação de $X_i$ , em função da probabilidade desejada, conhecendo a média e o desvio-padrão

Dados a probabilidade desejada, a média e o desvio-padrão, pode-se calcular o respectivo valor de  $X_i$ .

Nesse tipo de cálculo, parte-se da probabilidade definida e procura-se o valor de  $z$  na tabela. Se a probabilidade dada for acima de 50%, deve-se subtrair **do valor dado** 50%, que é o valor da metade da curva, localizada à direita.

Se a probabilidade dada for inferior a 50%, deve-se subtrair **o valor dado** de 50%, que representa a metade da curva, localizada à esquerda. As probabilidades assim encontradas devem ser expressas em número e não em percentual. Por exemplo, 48,3% deve ser expresso como 0,4830.

O valor da probabilidade deve ser procurado no corpo da tabela e o valor do respectivo  $z$  será encontrado na conjugação da linha que apresentará o valor principal e o primeiro dígito após a vírgula e a coluna que contém o último dígito.

Assim, usando o exemplo da probabilidade de 0,4830, encontra-se na Tabela 9.1b - Apêndice III o valor de  $z$ , como sendo 2,12.

Tabela 9.1b - APÊNDICE III - Distribuição Normal Padronizada -  $z$

$z$	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
4,5	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Fonte: Tabela construída com as funções estatísticas do EXCEL pelos autores.

Quando o valor procurado não coincidir com algum valor da tabela, usa-se a probabilidade que está mais próxima do valor procurado. Quando os valores da tabela tiverem a mesma distância em relação ao valor procurado, usa-se aquele que ocorrer primeiro, como, por exemplo, se a probabilidade a ser procurada é 0,3780, os valores mais próximos na tabela são 0,3770 e 0,3790, usa-se 0,3770. Outra forma é fazer uma interpolação entre o valor procurado e o valor existente na tabela, criando um terceiro dígito para o valor de  $z$ .

Encontrado o valor de  $z$ , este é substituído como dado na fórmula 9.3:

$$z = \frac{X_i - \mu}{\sigma} \quad \text{onde a incógnita será o valor de } X_i.$$

O valor do  $z$  encontrado na tabela sempre será positivo. Assim, deve-se atribuir ao seu valor o sinal negativo quando a probabilidade procurada no problema é de valores abaixo de 50%, localizados no lado esquerdo da curva.

Assim, desdobrando a fórmula de cálculo, o valor de  $X_i$  será obtido pela equação:

$X_i = \mu + z \sigma$ , quando a probabilidade fornecida localiza-se à direita da curva, ou

$X_i = \mu - z \sigma$ , quando a probabilidade fornecida localiza-se à esquerda da curva.

Para elucidar dúvidas e propiciar melhor entendimento de como são solucionados problemas de determinação de valores da variável aleatória, a partir de probabilidades definidas, são apresentados vários exemplos a seguir.

Exemplos:

1. Em uma prova aplicada aos alunos de estatística, a média foi 7 e o desvio-padrão foi 2. Qual foi a nota máxima obtida por 91% dos alunos?

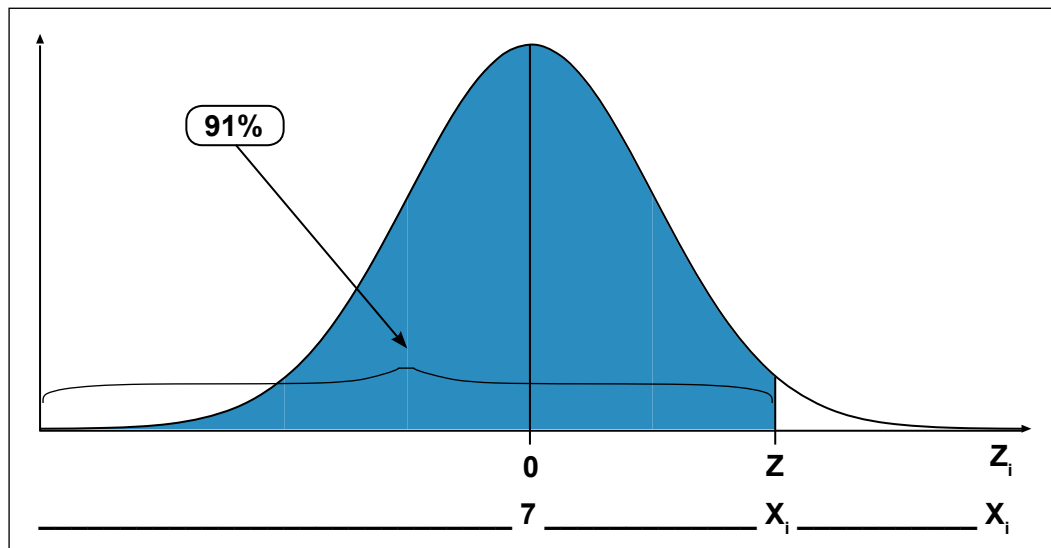
Dados:

$$\mu = 7$$

$$\sigma = 2$$

$$X_i \text{ de } 91\% = ? \text{ (nota máxima de 91\% dos alunos)}$$

Representação gráfica do problema da probabilidade procurada:



Valor da probabilidade a ser usada para procurar o valor de  $z$  na tabela:

Para determinar o valor de  $z$ , deve-se definir a probabilidade a ser usada na tabela:

Como a probabilidade de 91% está acima de 50%, deve-se subtrair de 0,91 o valor de 0,50, ou seja, o percentual ocupado pela curva normal à esquerda. Assim, o valor a procurar na tabela será:  $0,91 - 0,50 = 0,4100$

Procurando-se no corpo da Tabela 9.1c o valor que mais se aproxima da probabilidade de 0,4100, encontrar-se-á na linha correspondente a parte inteira e o primeiro dígito, com o valor de 1,3, e, no topo da coluna, o segundo dígito, que é 0,04. Assim, o valor da probabilidade mais aproximada de 0,4100 tem o valor de  $z$  de 1,34.

Tabela 9.1c - APÊNDICE III - Distribuição Normal Padronizada - z

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
4,5	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Fonte: Tabela construída com as funções estatísticas do EXCEL pelos autores.

Valor de  $X_i$  calculado:

Para calcular o valor  $X_i$ , transforma-se a fórmula:

$$z = \frac{X_i - \mu}{\sigma} \quad \text{para } X_i = \mu + z \sigma$$

O sinal de z é positivo, pois a probabilidade refere-se a um valor acima de 50% ( $0,91 > 0,50$ ).

Como  $\mu = 7$  e  $\sigma = 2$ , logo

$$X_i = \mu + z \sigma,$$

$$X_i = 7 + 1,34 \times 2$$

$$X_i = 7 + 2,68$$

$$X_i = 9,68$$

Resposta:

A nota máxima obtida por 91% dos alunos é 9,68.

Como opção, pode-se, também, afirmar que a nota mínima dos 9% dos alunos com maiores notas (parte superior que completa os 100%) é 9,68.

2. Uma fábrica produz motores cujo tempo de vida é uma variável aleatória com distribuição normal, de parâmetros  $\mu = 10$  anos e  $\sigma = 2$  anos. A fábrica quer criar um período de garantia com previsão de substituição gratuita no máximo de 1,5% dos motores. Qual deverá ser o período de garantia máximo de anos a ser oferecido pela fábrica?

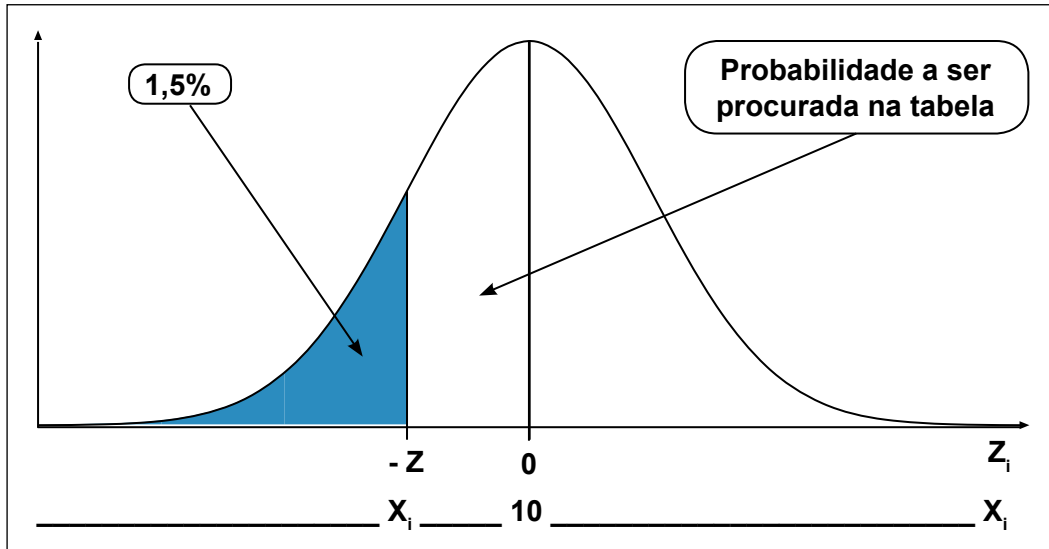
Dados:

$$\mu = 10$$

$$\sigma = 2$$

$$X_i \text{ de } 1,5\% = ? \text{ (número máximo de anos de garantia para } 1,5\% \text{ de trocas)}$$

Representação gráfica do problema da probabilidade procurada:



Valor da probabilidade a ser usado para procurar o valor de  $z$  na tabela:

Para determinar o valor de  $z$ , deve-se definir a probabilidade a ser usada na tabela:

Como a probabilidade de 1,5% está abaixo de 50%, deve-se subtrair do total da proporção da curva normal à esquerda, que é de 50%, o valor 1,5%, que é a probabilidade dada, para encontrar a probabilidade a ser procurada na tabela, que sempre se refere àquela existente no intervalo entre  $X_i$  e  $m$ . Assim, o valor a procurar na tabela será:  $0,50 - 0,015 = 0,4850$ .

Procurando-se no corpo da Tabela 9.1d o valor que mais se aproxima da probabilidade de 0,4850 encontrar-se-á na linha correspondente à parte inteira o primeiro dígito, com o valor de 2,1, e, no topo da coluna, o segundo dígito, que é 0,07. Assim, o valor da probabilidade mais aproximada de 0,4850 tem o valor de  $z$  de 2,17.

Tabela 9.1d - APÊNDICE III - Distribuição Normal Padronizada -  $z$

$z$	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
4,5	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Fonte: Tabela construída com as funções estatísticas do EXCEL pelos autores



Valor de  $X_i$  calculado:

Para calcular o valor  $X_i$ , transforma-se a fórmula:

$$z = \frac{X_i - \mu}{\sigma} \quad \text{para } X_i = \mu - z \sigma$$

O sinal de  $z$  é negativo, pois a probabilidade refere-se a um valor abaixo de 50% ( $0,015 < 0,50$ ).

Como  $\mu = 10$  e  $\sigma = 2$ , logo

$$X_i = \mu - z \sigma$$

$$X_i = 10 - 2,17 \times 2$$

$$X_i = 10 - 4,34$$

$$X_i = 5,66 \text{ anos}$$

Resposta:

O tempo de garantia máxima a ser dada é de 5,66 anos.

## 9.5 Relação entre as Distribuições Binomial e Normal

Com base na definição na Lei dos Grandes Números, pode-se considerar como verdadeira a aproximação da distribuição de probabilidade de variáveis discretas à distribuição de variáveis contínuas. Essa aproximação torna-se mais concreta à medida que aumenta o número de observações da variável.

Assim, aceita-se que as duas distribuições aproximam-se quando:

- o número de observações for grande, ou seja,  $n \geq 30$ ;
- as probabilidades de sucesso ( $p$ ) e de fracasso ( $q$ ) não forem muito próximas a zero;
- as médias de sucessos e de fracassos forem maiores do que cinco ( $n.p > 5$  ou  $n.q > 5$ );
- a aproximação melhora com o crescimento do número de observações e no limite (infinito) as duas distribuições coincidem.

### 9.5.1 Cálculo da aproximação do valor de $z$

Com base nas propriedades da Distribuição Binomial, desenvolvidas no capítulo 8, Tópico 8.2.5, o cálculo de  $z$  da aproximação será dado por:

$$z = \frac{X_i - \mu}{\sigma} \quad \text{substituindo os valores de } \mu \text{ e } \sigma, \text{ temos:}$$

$$z = \frac{X_i - Np}{\sqrt{Npq}} \quad \text{onde o número médio de sucessos é dado por:}$$

$\mu = Np$  e o desvio-padrão do número de sucessos é dado por:

$$\sigma = \sqrt{Npq}$$

Quando se usa a aproximação da Normal para a Binomial, aproximando-se uma variável discreta (que só assume valores inteiros) a uma variável contínua (que pode assumir quaisquer valores dentro de um intervalo de número reais). Nesse caso, deve-se fazer um ajuste que é denominado de correção de continuidade.

Essa correção é necessária, pois entre dois números consecutivos da variável discreta há um espaço vazio de uma unidade, como, por exemplo, entre os valores 2 e 3. Esse espaço na variável contínua não existe, pois é ocupado por infinitos valores não inteiros. Assim, para dar uma forma contínua aos dados é necessário distribuir o espaço vazio entre os dois números consecutivos de forma equitativa. Dessa forma, a correção é feita simplesmente pela soma ou subtração de 0,5 ao valor inteiro da variável discreta  $X$  (número de sucessos).

Assim, no cálculo de valores de  $z$ , precisa-se ajustar o valor de  $X_i$  discreto à situação da variável contínua. Já na análise de probabilidades de uma variável contínua a inclusão ou não do valor limite do intervalo definido pelo seu  $X_i$  não interfere no resultado, pois a probabilidade do valor exato de  $X_i$  é zero. No caso da variável discreta, o valor exato tem um valor de probabilidade.

Dessa forma, ao calcular o valor de  $z$  na aproximação de uma variável discreta à distribuição Normal, deve-se usar a correção de continuidade das seguintes formas:

- 1) Soma de probabilidades  $\geq X$  de números de sucessos. Para definir o valor de  $X_i$ , na fórmula do cálculo de  $z$ , subtrai-se de  $X$  o valor de 0,5. Exemplo:  $P(X \geq 5) \rightarrow X_i = 5 - 0,5 = 4,5$ .
- 2) Soma de probabilidades  $> X$  de números de sucessos. Para definir o valor de  $X_i$ , na fórmula do cálculo de  $z$ , soma-se a  $X$  o valor de 0,5. Exemplo:  $P(X > 5) \rightarrow X_i = 5 + 0,5 = 5,5$ .
- 3) Soma de probabilidades  $\leq X$  de números de sucessos. Para definir o valor de  $X_i$ , na fórmula do cálculo de  $z$ , soma-se a  $X$  o valor de 0,5. Exemplo:  $P(X \leq 5) \rightarrow X_i = 5 + 0,5 = 5,5$ .
- 4) Soma de probabilidades  $< X$  de números de sucessos. Para definir o valor de  $X_i$ , na fórmula do cálculo de  $z$ , subtrai-se de  $X$  o valor de 0,5. Exemplo:  $P(X < 5) \rightarrow X_i = 5 - 0,5 = 4,5$ .
- 5) Probabilidade de um determinado número  $X$  de sucessos. Nesse caso, como o valor de  $X$  tem um valor de probabilidade, ele ocupa um espaço na curva localizado entre dois pontos. Assim, o cálculo deve ser feito em duas etapas:  $z_1$  e  $z_2$ , com os respectivos  $X_1$  e  $X_2$ . Para definir o valor de  $X_1$ , na fórmula do cálculo de  $z_1$ , subtrai-se de  $X$  o valor de 0,5 e para definir o valor de  $X_2$ , na fórmula do cálculo de  $z_2$ , soma-se a  $X$  o valor de 0,5. Exemplo:  $P(X = 5) \rightarrow X_1 = 5 - 0,5 = 4,5$  e  $X_2 = 5 + 0,5 = 5,5$ .

#### Exemplo

1. Fazendo o lançamento de 100 moedas simultaneamente, qual é a probabilidade de obter?
  - a) 62 caras ou  $P(x = 62)$ ;
  - b) menos de 62 caras ou  $P(x < 62)$ .

Solução da questão a:

Dados:

$$n = 100$$

$$p = 1/2$$

$$q = 1/2$$

$$P(x = 62)?$$

A solução da questão a pode ser encontrada com o uso da fórmula da Distribuição Binomial, que é a seguinte:

$$P(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

Utilizando os dados do problema, a solução pela fórmula será a seguinte:

$$P(62) = \frac{100!}{62!(100-62)!} 1/2^{62} 1/2^{38}$$

O desenvolvimento dos cálculos resulta no seguinte resultado:

$$P(62) = 0,00447 \text{ ou } 0,447\%$$

Assim, a probabilidade de obter 62 caras em 100 lançamentos de moedas é de 0,447% ou, arredondado para dois dígitos, 0,45%.

Em se tratando de uma probabilidade específica de sucessos, o problema, mesmo com um grande número de repetições, pode ser resolvido pela fórmula da probabilidade Binomial.

Mas, como o problema se enquadra nas regras exigidas para a aproximação à Normal, o problema também será resolvido por essa distribuição.

Dados:

$$n = 100$$

$$p = \frac{1}{2} \text{ ou } 0,5$$

$$q = \frac{1}{2} \text{ ou } 0,5$$

$$P(x = 62)?$$

Valores de  $\mu$  e  $\sigma$ :

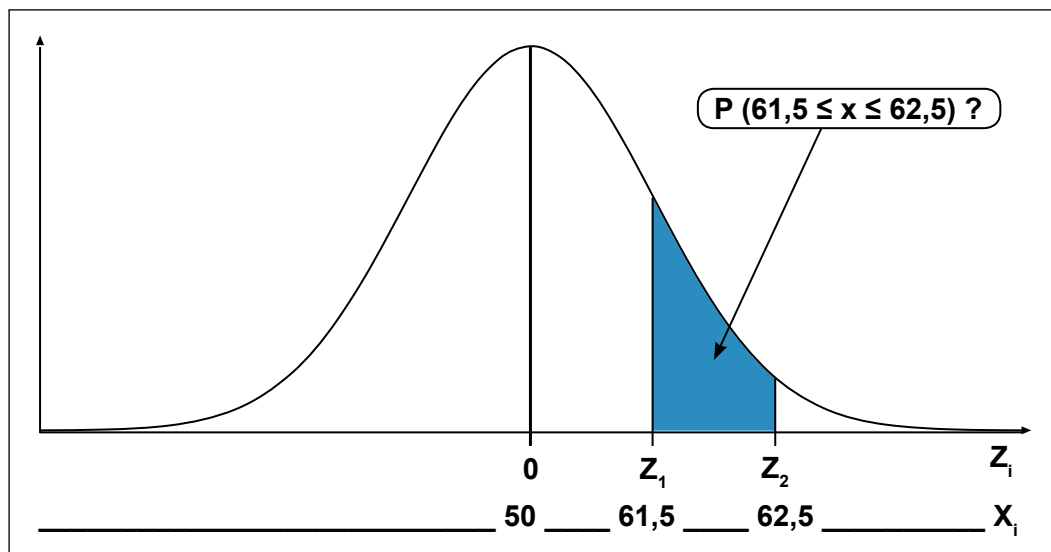
$$\mu = np = 100 \times \frac{1}{2} = 50$$

$$\sigma = \sqrt{npq} = \sqrt{100 \times 0,5 \times 0,5} = 5$$

Como se está usando uma variável discreta em que cada valor dessa variável ocupa uma determinada área da curva, diferente do dado contínuo, em função de haver um espaço vazio entre uma unidade e outra, cuja amplitude é igual a 1, esse espaço deve ser distribuído proporcionalmente para cada unidade envolvida, subtraindo-se 0,5 e somando-se 0,5 ao valor de  $X$  que está sendo analisado. Assim, o número 62 ocupa o espaço entre 61,5 até 62,5. Portanto, o problema pode ser apresentado da seguinte forma:

$$P(61,5 \leq x \leq 62,5)?$$

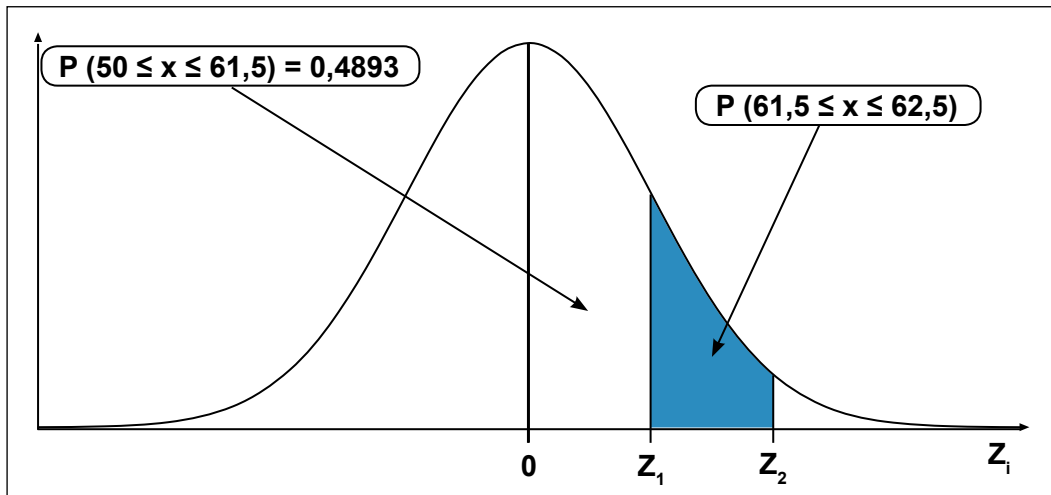
Representação gráfica do problema da probabilidade procurada:



Cálculo do valor de  $z_1$ :

$$z_1 = \frac{X_i - \mu}{\sigma} = \frac{61,5 - 50}{5} = 2,30$$

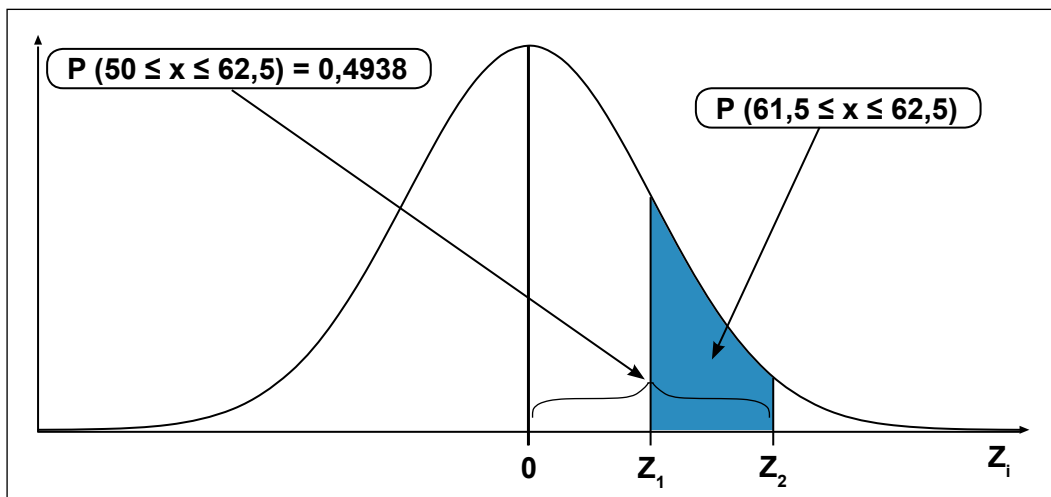
Probabilidade encontrada na tabela com o valor de  $Z_1 = 2,30$ :



Cálculo do valor de  $z_2$ :

$$z_2 = \frac{X_i - \mu}{\sigma} = \frac{62,5 - 50}{5} = 2,50$$

Probabilidade encontrada na tabela com o valor de  $Z_2 = 2,50$ :



Solução para encontrar a probabilidade desejada no problema:

$$P(61,5 \leq x \leq 62,5) = 0,4938 - 0,4893 = 0,0045 \text{ ou } 0,45\%$$

Resposta:

A probabilidade de obter 62 caras em 100 lançamentos de moedas é de 0,45%. Portanto, o resultado coincide com o valor calculado pela fórmula da distribuição Binomial.

Solução da questão b:

Dados:

$$n = 100$$

$$p = 1/2$$

$$q = 1/2$$

$$P(x < 62)?$$

A solução da questão **b** pela distribuição binomial seria muito demorada, pois há a necessidade de fazer 62 cálculos pela fórmula, ou seja, determinar as probabilidades de 0 caras até 61 caras, que no final devem ser somadas.

Para facilitar essa tarefa, usa-se uma aproximação à Distribuição Normal.

Assim,

Dados:

$$n = 100$$

$$p = \frac{1}{2} \text{ ou } 0,5$$

$$q = \frac{1}{2} \text{ ou } 0,5$$

$$P(x < 62)?$$

Valores de  $\mu$  e  $\sigma$ :

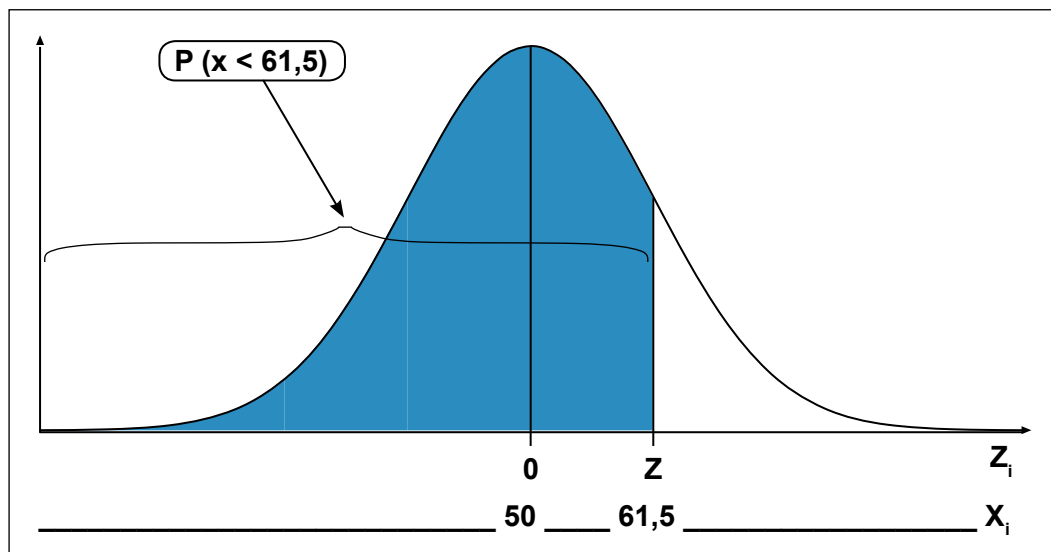
$$\mu = np = 100 \times \frac{1}{2} = 50$$

$$\sigma = \sqrt{npq} = \sqrt{100 \times 0,5 \times 0,5} = 5$$

Como se está usando uma variável discreta em que cada valor dessa variável ocupa uma determinada área da curva, diferente do dado contínuo, em função de haver um espaço vazio entre uma unidade e outra, cuja amplitude é igual a 1, deve-se subtrair 0,5 de 62, que é o limite inferior do número 62. Portanto, o problema pode ser apresentado da seguinte forma:

$$P(x \leq 61,5)?$$

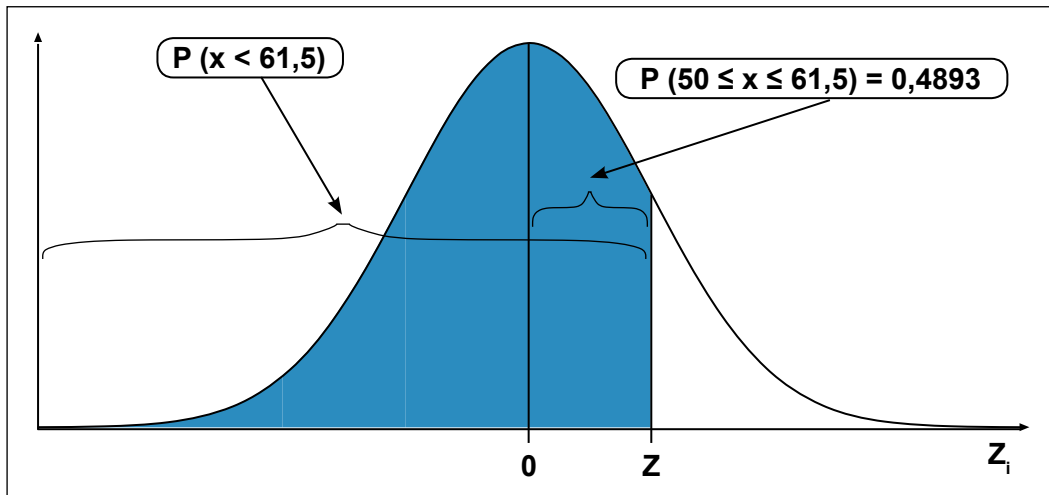
Representação gráfica do problema da probabilidade procurada:



Cálculo do valor de  $Z$ :

$$z_1 = \frac{X_i - \mu}{\sigma} = \frac{61,5 - 50}{5} = 2,30$$

Probabilidade encontrada na tabela com o valor de  $Z_i = 2,30$ :



Solução para encontrar a probabilidade desejada no problema:

Como a questão considera o número de sucessos até 61, devem ser consideradas todas as possibilidades que existem entre 0 até 61. Assim, na solução do problema, deve ser considerada a probabilidade do número de sucessos abaixo da média, ou seja, até 50, que é a metade da curva à esquerda e que representa 50%, e a parte dos sucessos acima da média, até 61.

Dessa forma,  $P(x < 61,5) = 0,50 + 0,4893 = 0,9893$  ou 98,93%

Resposta:

A probabilidade de encontrar até 61 caras em 100 lançamentos de moedas é de 98,93%.

2. Uma máquina produz parafusos dos quais 15% são defeituosos. Tomou-se uma amostra de 500 parafusos. Qual é a probabilidade de que dentre esses saiam até 65 parafusos defeituosos?

Dados:

$$n = 500$$

$$p = 0,15$$

$$q = 0,85$$

$$P(x \leq 65)?$$

Valores de  $\mu$  e  $\sigma$ :

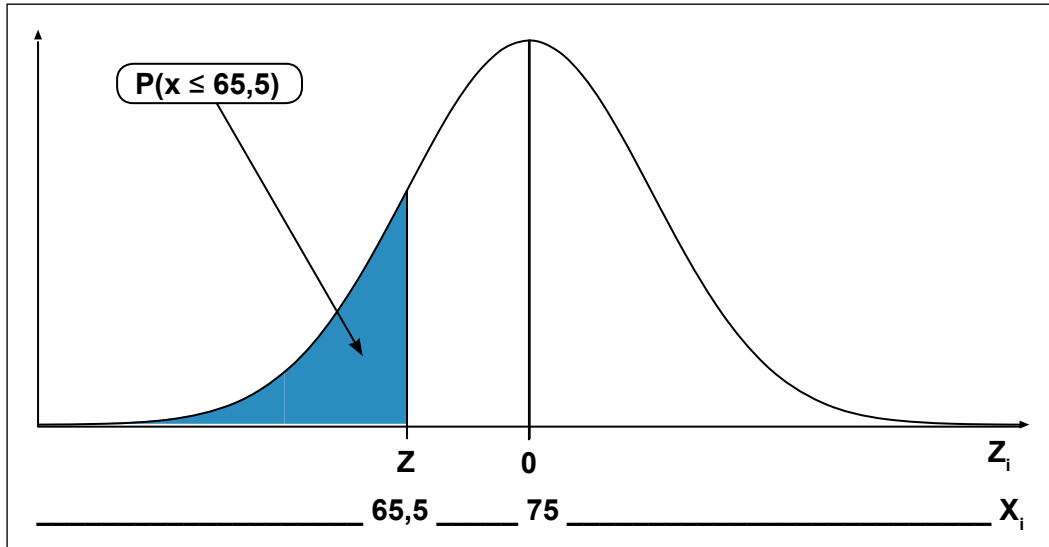
$$\mu = np = 500 \times 0,15 = 75$$

$$\sigma = \sqrt{npq} = \sqrt{500 \times 0,15 \times 0,85} = 7,98 \simeq 8$$

Como se está usando uma variável discreta em que cada valor desta variável ocupa uma determinada área da curva, diferente do dado contínuo, em função de haver um espaço vazio entre uma unidade e outra, cuja amplitude é igual a 1, deve-se somar 0,5 ao valor de 65, que é o limite superior do número 65. Portanto, o problema pode ser apresentado da seguinte forma:

$$P(x \leq 65,5)?$$

Representação gráfica do problema da probabilidade procurada:

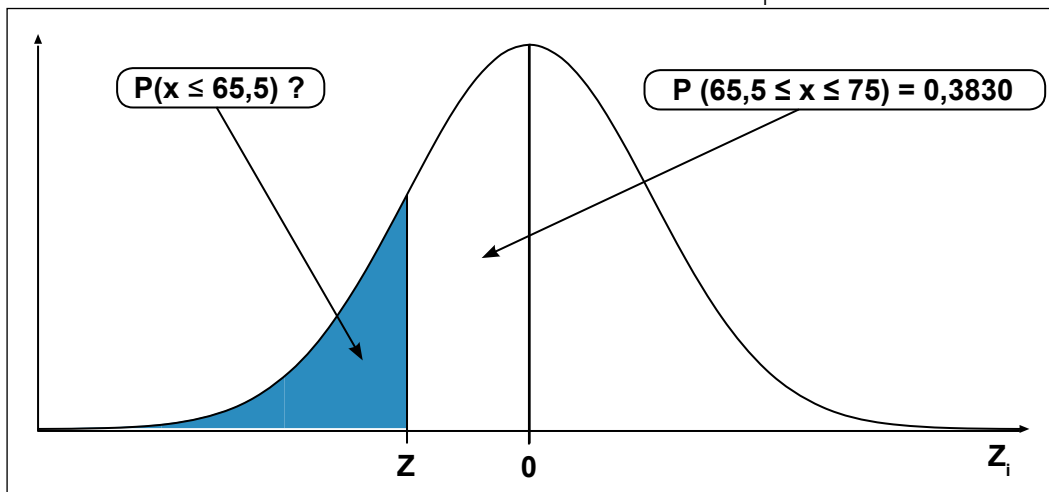


Cálculo do valor de z:

O valor de z será calculado pela fórmula abaixo:

$$z = \frac{X_i - \mu}{\sigma} = \frac{65,5 - 75}{7,98} = -1,19$$

Probabilidade encontrada na tabela com o valor de  $Z_1 = 1,19$ :



Solução para encontrar a probabilidade desejada no problema:

Como a questão considera o número de sucessos até 65, devem ser consideradas todas as possibilidades que existem entre 0 até 65. Assim, na solução do problema, deve ser considerada a probabilidade do número de sucessos ser abaixo da média de 75.

Dessa forma,  $P(x \leq 65,5) = 0,50 - 0,3830 = 0,1170$  ou 11,70%

Resposta:

A probabilidade de encontrar até 65 parafusos com defeitos de uma amostra de 500 é de 11,70%.

## 9.6 Resumo para fixação

Para facilitar a memorização dos principais conceitos abordados neste capítulo, é apresentado um resumo dos pontos mais relevantes:

- Na natureza, quando o número de dados do universo analisado é relativamente grande e principalmente quando for de uma variável contínua, a distribuição dos dados apresenta uma curva com formato de um sino, com um ponto máximo no centro, em que as áreas, em ambos os lados da média, são idênticas.
- A curva formada chama-se de curva normal, que é definida como sendo simétrica e unimodal, tendo como característica a igualdade entre as medidas: média, moda e mediana.
- Quando as medidas, média, moda e mediana não são iguais, mas semelhantes, chama-se a distribuição de aproximadamente normal.
- Pela distribuição normal padronizada, que apresenta como características uma curva simétrica e mesocúrtica, podem-se determinar valores de probabilidade e apresentá-los em tabelas que expressam os valores da função densidade ou curva de probabilidade, com base na variável padronizada  $z$ .
- Os valores da função densidade localizam-se abaixo da curva da distribuição e seu valor total é de 100%, localizando-se 50% à direita da média e 50% à sua esquerda.
- Para determinar os valores de  $z$ , precisa-se dos valores da média, do desvio-padrão e do valor de  $X_i$  de referência.
- Os valores de  $z$ , à direita da média, são positivos e os localizados à esquerda, negativos.
- Uma das alternativas é a determinação da probabilidade em função de  $X_i$ , sendo a média e o desvio-padrão conhecidos.
- No cálculo do valor de  $z$  devem ser usados dois dígitos após a vírgula. A parte inteira e o primeiro dígito localizam-se na primeira coluna e o segundo dígito, no cabeçalho do Apêndice III.
- Os valores da probabilidade encontrados no Apêndice III referem-se sempre ao intervalo entre o valor da média e o valor de  $X_i$ .
- Para a solução do problema, existem várias possibilidades: o próprio valor encontrado na tabela, a soma ou subtração do valor da tabela de 0,5, a soma de 0,5 ao valor da tabela e subtração do valor menor do valor maior, encontrados na tabela.
- Para facilitar a solução de problemas de probabilidade Normal, recomenda-se sempre fazer o desenho da curva e indicar a área a ser conhecida.
- Outra alternativa é a determinação do valor de  $X_i$  em função da probabilidade desejada, conhecendo a média e o desvio-padrão.
- Deve-se ter o cuidado para atribuir o sinal correto de  $z$ , pois os seus valores de  $z$ , à direita da média, são positivos e os localizados à esquerda, negativos.
- Valores de variáveis discretas podem ser adaptados à distribuição normal, nas seguintes condições: número de observações  $\geq 30$ ; probabilidades de sucesso ( $p$ ) e de fracasso ( $q$ ) não serem muito próximas a zero e as médias de sucesso e de fracasso serem maiores do que cinco.
- Na aproximação da Binomial à Normal deve-se fazer a correção de continuidade, pela soma ou subtração de 0,5 ao valor inteiro da variável discreta  $X$  (número de sucessos). Essa correção é necessária e deve ser feita de forma equitativa, pois entre dois números consecutivos da variável discreta há um espaço vazio de uma unidade.
- Os valores da média e do desvio-padrão são obtidos pelas fórmulas de cálculo das propriedades das variáveis discretas.



## 9.7 Exercícios propostos

9.1 Sendo  $z$  uma variável aleatória contínua com distribuição  $N(0, 1)$ , determine:

- |                                  |                                   |
|----------------------------------|-----------------------------------|
| a) $P(0 \leq z \leq 1)$ ;        | d) $P(-2,33 \leq z \leq -0,65)$ ; |
| b) $P(0 \leq z \leq 1,35)$ ;     | e) à esquerda de $z = -1,38$ ;    |
| c) $P(-2,33 \leq z \leq 1,96)$ ; | f) à direita de $z = -1,38$ ;     |

9.2 Supondo que o tempo de duração, dado em horas, de certo dispositivo eletrônico, esteja distribuído segundo uma  $N(500;20)$ , determine a probabilidade de o tempo de duração de uma unidade selecionada ao acaso ser:

- |                            |                           |
|----------------------------|---------------------------|
| a) menor que 498 horas;    | c) entre 480 e 510 horas. |
| b) maior do que 503 horas; |                           |

9.3 O tempo de vida de luzes de jardim é normalmente distribuído com média de 240 minutos e desvio-padrão de 24 minutos. Defina:

- a) a proporção de luzes que se pode esperar que necessitam ser repostas com até 200 minutos de funcionamento;
- b) a proporção de luzes que se pode esperar que necessitam ser repostas com mais de 260 minutos de funcionamento;
- c) a proporção de luzes que se pode esperar que necessitam ser repostas entre 200 e 260 minutos de funcionamento;
- d) em uma remessa de 650 luzes de jardim, quantas seriam de se esperar que queimassem em menos de 200 minutos de funcionamento?

9.4 O faturamento diário, excluindo sábados, em um restaurante tem distribuição normal com média de R\$ 5.300,00 e desvio-padrão de R\$320,00. Com base nesses dados defina o seguinte:

- a) Qual a probabilidade de que o faturamento exceda, em um dado dia, a R\$ 6.000,00?
- b) O restaurante necessita faturar pelo menos R\$ 4.500,00 diariamente para cobrir seus custos. Qual a probabilidade de que, em um dado dia, o estabelecimento não consiga cobrir os custos?

9.5 Uma fábrica de carros sabe que os motores de sua fabricação têm duração normal, com média de 150.000 km e desvio-padrão de 5.000 km. Qual a probabilidade de que um carro, escolhido ao acaso, dos fabricados por essa firma tenha um motor que dure:

- a) menos de 160.000 km?
- b) entre 140.000 km e 165.000 km?
- c) se a fábrica substitui o motor que apresenta duração inferior à garantia, qual deve ser essa garantia para que a porcentagem de motores substituídos seja inferior a 0,2%?

9.6 Supondo que as notas em certa disciplina estão normalmente distribuídas com média 5,0 e desvio-padrão 1,5. Determine a probabilidade de que:

- a) um estudante tire nota superior a 8,0;
- b) se a nota mínima para obter aprovação é 3,0, qual o percentual de estudantes reprovados?
- c) determine a nota mínima dos 10% melhores estudantes;
- d) explique por que a probabilidade de um estudante obter nota acima de 9,8 é praticamente zero.

9.7 O tempo para realizar uma entrevista de uma pesquisa sócioeconômica tem distribuição normal com média de 60 minutos e desvio-padrão de 12 minutos. Determine:

- a) Qual o percentual de entrevistadores que levarão até 87 minutos?
- b) Se o entrevistador que levar mais de 90 minutos para realizar uma entrevista for excluído nas pesquisas futuras, qual o percentual de excluídos?

9.8 Considerando que o peso de determinado artigo produzido por uma fábrica é normalmente distribuído com média de 20 gramas e desvio-padrão de 4 gramas, determine a probabilidade de que uma unidade, selecionada ao acaso, tenha peso:

- a) entre 16 e 22 gramas;
- b) entre 22 e 25 gramas;
- c) maior que 23 gramas;
- d) inferior a 16 gramas;
- e) o peso abaixo do qual se espera encontrar 25,78% das unidades;
- f) o peso acima do qual se espera encontrar 34,46% das unidades.

9.9 Em uma praça se colocam 10.000 lâmpadas em serviço, entre 15 de dezembro e 15 de março, do ano seguinte, durante à noite. A vida útil pode ser considerada com vida média de 50 dias e um coeficiente de variabilidade de 20%.  $\mu = 50$  dia  $\sigma = 10$  dia  $CV = 0,20$ . Defina:

- a) Qual será o número de lâmpadas que deverão ser repostas entre as 12 horas do dia 15 de dezembro até as 12 horas do dia 31 de dezembro?
- b) Qual será o número de lâmpadas a serem repostas entre às 12 horas do dia 15 de dezembro até as 12 horas do dia 31 de janeiro?

9.10 O tempo de vida útil de uma máquina de lavar automática tem distribuição aproximadamente normal, com média de 3,1 e desvio-padrão de 1,2 anos. Se esse tipo de lavadoras tiver garantia de 1 ano, que fração das vendas originais exigirá substituição?

9.11 Por meio de documentação e observação cuidadosa, constatou-se que o tempo médio para se fazer um teste padrão de estatística é aproximadamente normal com média de 70 minutos e desvio-padrão de 25 minutos. Defina:

- a) Que percentagem de candidatos levará menos de 80 minutos para fazer o teste?
- b) Que percentagem de candidatos não terminará o teste no tempo máximo de 2 horas?
- c) Que percentagem terminará o teste na primeira hora?
- d) Se 100 pessoas fazem o teste, quantos terminam na primeira hora?

9.12 O grau médio de um exame final foi 72 e o desvio-padrão 9. Dez por cento dos melhores estudantes receberam a classificação A. Qual o grau mínimo que um estudante deve obter para classificar-se em A?

9.13 O salário médio mensal dos funcionários de uma empresa é de R\$ 800,00 com um desvio-padrão de R\$ 150,00. Supondo que os salários estejam normalmente distribuídos, determine:

- a) a probabilidade de que um funcionário dessa empresa selecionado ao acaso tenha salário entre R\$ 640,00 e R\$ 925,00;
- b) a probabilidade de que um funcionário tenha salário inferior a R\$ 624,00;
- c) o percentual de funcionários que tenham salário superior a R\$ 1.000,00;
- d) se for selecionada uma amostra aleatória de 500 funcionários, quanto deles terão salário abaixo de R\$ 1.000,00 ?
- e) os limites de salário, centrados na média, em que se encontram 68,26% dos funcionários.

9.14 Sabendo que, em uma prova constituída de 100 questões de escolha simples, cada questão apresenta quatro respostas possíveis das quais uma só é correta, determine a probabilidade de acertar ao acaso, pela aproximação da Binomial à Normal:

- a) 20 questões;
- b) mais de 20 questões;
- c) menos de 20 questões;
- d) 30 questões;
- e) de 20 a 30 questões.

9.15 O número de pessoas que almoçam num restaurante suburbano é aproximadamente normal, com média 250 e desvio-padrão de 20 por dia.

- a) Qual é a probabilidade de haver ao menos 200 clientes em determinado dia?
- b) Determine a probabilidade de comparecerem entre 225 e 275 clientes?
- c) Se o preço médio pago por dia por cada cliente é R\$ 4,00, qual é a receita diária esperada?

9.16 Uma máquina produz parafusos, dos quais 10% são defeituosos. Usando a aproximação da distribuição binomial pela normal, determine a probabilidade de numa amostra formada ao acaso de 400 parafusos produzidos pela máquina serem defeituosos:

- a) no máximo 30;
- b) entre 30 e 50 (inclusive os extremos);
- c) mais de 35 e menos de 45;
- d) mais de 55.

9.17 Sabe-se que 16 pessoas a cada 80 consultadas são consumidoras de certo produto. Entrevistando-se 400 pessoas qual a probabilidade de encontrar-se:

- a) no máximo 90 pessoas consumidoras de tal produto;
- b) entre 300 e 330 pessoas (extremos incluídos) não consumidoras de tal produto.

9.18 Sabendo que certo processo industrial produz 4% de unidades defeituosas e considerando uma amostra ocasional de 150 unidades, determine a probabilidade do número de unidades defeituosas:

- a) ser igual a 8;
- b) ser maior que 8;
- c) ser menor que 8;
- d) ser igual ou maior que 8;
- e) estar contido no intervalo  $[5;8]$ .

## REFERÊNCIAS

- DOWNING, Douglas; CLARK, Jeffrey. **Estatística Aplicada**, São Paulo: Editora Saraiva, 1998.
- FEE. Fundação de Economia e Estatística. Disponível em: <[www.fee.rs.gov.br](http://www.fee.rs.gov.br)>.
- FONSECA, Jairo Simon da; MARTINS, Gilberto de Andrade. **Curso de Estatística**. 5. ed. São Paulo: Editora Atlas S.A, 1994.
- IBGE. Instituto Brasileiro de Geografia e Estatística. Disponível em: <[www.ibge.gov.br](http://www.ibge.gov.br)>.
- MATTAR, F. N. **Pesquisa de marketing**. São Paulo: Atlas, 1996.
- MORETTIN, Luiz Gonzaga. **Estatística Básica: Probabilidade**. Volume 1. São Paulo: Makron Books Ltda.
- LARSON, Ron; FARBER, Betsi. **Estatística aplicada**. 2. de. São Paulo: Pearson Prentice Hall, 2006.
- LEVINE, David M.; BERENSON, Mark L.; STEPHAN, David. **Estatística: teoria e aplicações usando Microsoft Excel e português**. Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora S.A, 2000.
- SPIEGEL, Murray R. **Probabilidade e Estatística**. São Paulo: McGraw- Hill do Brasil, 1978.
- STEVENSON, William J. **Estatística Aplicada à Administração**. São Paulo: Editora Harbra Ltda. 1986.
- TOLEDO Geraldo L.; OVALLE Ivo I. **Estatística Básica**. São Paulo: Editora Atlas SA, 1992.
- TRIOLA, Mário F. **Introdução à Estatística**. Rio de Janeiro: Livros Técnicos e Científicos Editora S.A., 1999.

## LISTA DE APÊNDICES

APÊNDICE I - Distribuição das Probabilidades Binomiais

APÊNDICE II - Distribuição das Probabilidades de Poisson

APÊNDICE III - Distribuição Normal Padronizada -  $z$

APÊNDICE IV - Respostas dos exercícios propostos

**APÊNDICE I – Distribuição das Probabilidades Binomiais (1)**

## Probabilidade

n	x	0,01	0,02	0,03	0,04	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
1	0	0,9900	0,9800	0,9700	0,9600	0,9500	0,9000	0,8500	0,8000	0,7500	0,7000	0,6500	0,6000	0,5500	0,5000
	1	0,0100	0,0200	0,0300	0,0400	0,0500	0,1000	0,1500	0,2000	0,2500	0,3000	0,3500	0,4000	0,4500	0,5000
2	0	0,9801	0,9604	0,9409	0,9216	0,9025	0,8100	0,7225	0,6400	0,5625	0,4900	0,4225	0,3600	0,3025	0,2500
	1	0,0198	0,0392	0,0582	0,0768	0,0950	0,1800	0,2550	0,3200	0,3750	0,4200	0,4550	0,4800	0,4950	0,5000
	2	0,0001	0,0004	0,0009	0,0016	0,0025	0,0100	0,0225	0,0400	0,0625	0,0900	0,1225	0,1600	0,2025	0,2500
3	0	0,9703	0,9412	0,9127	0,8847	0,8574	0,7290	0,6141	0,5120	0,4219	0,3430	0,2746	0,2160	0,1664	0,1250
	1	0,0294	0,0576	0,0847	0,1106	0,1354	0,2430	0,3251	0,3840	0,4219	0,4410	0,4436	0,4320	0,4084	0,3750
	2	0,0003	0,0012	0,0026	0,0046	0,0071	0,0270	0,0574	0,0960	0,1406	0,1890	0,2389	0,2880	0,3341	0,3750
	3	0,0000	0,0000	0,0000	0,0001	0,0001	0,0010	0,0034	0,0080	0,0156	0,0270	0,0429	0,0640	0,0911	0,1250
4	0	0,9606	0,9224	0,8853	0,8493	0,8145	0,6561	0,5220	0,4096	0,3164	0,2401	0,1785	0,1296	0,0915	0,0625
	1	0,0388	0,0753	0,1095	0,1416	0,1715	0,2916	0,3685	0,4096	0,4219	0,4116	0,3845	0,3456	0,2995	0,2500
	2	0,0006	0,0023	0,0051	0,0088	0,0135	0,0486	0,0975	0,1536	0,2109	0,2646	0,3105	0,3456	0,3675	0,3750
	3	0,0000	0,0000	0,0001	0,0002	0,0005	0,0036	0,0115	0,0256	0,0469	0,0756	0,1115	0,1536	0,2005	0,2500
	4	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0016	0,0039	0,0081	0,0150	0,0256	0,0410	0,0625
5	0	0,9510	0,9039	0,8587	0,8154	0,7738	0,5905	0,4437	0,3277	0,2373	0,1681	0,1160	0,0778	0,0503	0,0313
	1	0,0480	0,0922	0,1328	0,1699	0,2036	0,3281	0,3915	0,4096	0,3955	0,3602	0,3124	0,2592	0,2059	0,1563
	2	0,0010	0,0038	0,0082	0,0142	0,0214	0,0729	0,1382	0,2048	0,2637	0,3087	0,3364	0,3456	0,3369	0,3125
	3	0,0000	0,0001	0,0003	0,0006	0,0011	0,0081	0,0244	0,0512	0,0879	0,1323	0,1811	0,2304	0,2757	0,3125
	4	0,0000	0,0000	0,0000	0,0000	0,0000	0,0005	0,0022	0,0064	0,0146	0,0284	0,0488	0,0768	0,1128	0,1563
	5	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0010	0,0024	0,0053	0,0102	0,0185	0,0313
6	0	0,9415	0,8858	0,8330	0,7828	0,7351	0,5314	0,3771	0,2621	0,1780	0,1176	0,0754	0,0467	0,0277	0,0156
	1	0,0571	0,1085	0,1546	0,1957	0,2321	0,3543	0,3993	0,3932	0,3560	0,3025	0,2437	0,1866	0,1359	0,0938
	2	0,0014	0,0055	0,0120	0,0204	0,0305	0,0984	0,1762	0,2458	0,2966	0,3241	0,3280	0,3110	0,2780	0,2344
	3	0,0000	0,0002	0,0005	0,0011	0,0021	0,0146	0,0415	0,0819	0,1318	0,1852	0,2355	0,2765	0,3032	0,3125
	4	0,0000	0,0000	0,0000	0,0000	0,0001	0,0012	0,0055	0,0154	0,0330	0,0595	0,0951	0,1382	0,1861	0,2344
	5	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0015	0,0044	0,0102	0,0205	0,0369	0,0609	0,0938
	6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002	0,0007	0,0018	0,0041	0,0083	0,0156
7	0	0,9321	0,8681	0,8080	0,7514	0,6983	0,4783	0,3206	0,2097	0,1335	0,0824	0,0490	0,0280	0,0152	0,0078
	1	0,0659	0,1240	0,1749	0,2192	0,2573	0,3720	0,3960	0,3670	0,3115	0,2471	0,1848	0,1306	0,0872	0,0547
	2	0,0020	0,0076	0,0162	0,0274	0,0406	0,1240	0,2097	0,2753	0,3115	0,3177	0,2985	0,2613	0,2140	0,1641
	3	0,0000	0,0003	0,0008	0,0019	0,0036	0,0230	0,0617	0,1147	0,1730	0,2269	0,2679	0,2903	0,2918	0,2734
	4	0,0000	0,0000	0,0000	0,0001	0,0002	0,0026	0,0109	0,0287	0,0577	0,0972	0,1442	0,1935	0,2388	0,2734
	5	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0012	0,0043	0,0115	0,0250	0,0466	0,0774	0,1172	0,1641
	6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0013	0,0036	0,0084	0,0172	0,0320	0,0547
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002	0,0006	0,0016	0,0037	0,0078

Fonte: Tabela construída com as funções estatísticas do EXCEL

Continua...

**APÊNDICE I – Distribuição das Probabilidades Binomiais (2)****Probabilidade**

n	x	0,01	0,02	0,03	0,04	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
8	0	0,9227	0,8508	0,7837	0,7214	0,6634	0,4305	0,2725	0,1678	0,1001	0,0576	0,0319	0,0168	0,0084	0,0039
	1	0,0746	0,1389	0,1939	0,2405	0,2793	0,3826	0,3847	0,3355	0,2670	0,1977	0,1373	0,0896	0,0548	0,0313
	2	0,0026	0,0099	0,0210	0,0351	0,0515	0,1488	0,2376	0,2936	0,3115	0,2965	0,2587	0,2090	0,1569	0,1094
	3	0,0001	0,0004	0,0013	0,0029	0,0054	0,0331	0,0839	0,1468	0,2076	0,2541	0,2786	0,2787	0,2568	0,2188
	4	0,0000	0,0000	0,0001	0,0002	0,0004	0,0046	0,0185	0,0459	0,0865	0,1361	0,1875	0,2322	0,2627	0,2734
	5	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0026	0,0092	0,0231	0,0467	0,0808	0,1239	0,1719	0,2188
	6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0011	0,0038	0,0100	0,0217	0,0413	0,0703	0,1094
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0012	0,0033	0,0079	0,0164	0,0313
9	0	0,9135	0,8337	0,7602	0,6925	0,6302	0,3874	0,2316	0,1342	0,0751	0,0404	0,0207	0,0101	0,0046	0,0020
	1	0,0830	0,1531	0,2116	0,2597	0,2985	0,3874	0,3679	0,3020	0,2253	0,1556	0,1004	0,0605	0,0339	0,0176
	2	0,0034	0,0125	0,0262	0,0433	0,0629	0,1722	0,2597	0,3020	0,3003	0,2668	0,2162	0,1612	0,1110	0,0703
	3	0,0001	0,0006	0,0019	0,0042	0,0077	0,0446	0,1069	0,1762	0,2336	0,2668	0,2716	0,2508	0,2119	0,1641
	4	0,0000	0,0000	0,0001	0,0003	0,0006	0,0074	0,0283	0,0661	0,1168	0,1715	0,2194	0,2508	0,2600	0,2461
	5	0,0000	0,0000	0,0000	0,0000	0,0000	0,0008	0,0050	0,0165	0,0389	0,0735	0,1181	0,1672	0,2128	0,2461
	6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0028	0,0087	0,0210	0,0424	0,0743	0,1160	0,1641
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0012	0,0039	0,0098	0,0212	0,0407	0,0703
10	0	0,9044	0,8171	0,7374	0,6648	0,5987	0,3487	0,1969	0,1074	0,0563	0,0282	0,0135	0,0060	0,0025	0,0010
	1	0,0914	0,1667	0,2281	0,2770	0,3151	0,3874	0,3474	0,2684	0,1877	0,1211	0,0725	0,0403	0,0207	0,0098
	2	0,0042	0,0153	0,0317	0,0519	0,0746	0,1937	0,2759	0,3020	0,2816	0,2335	0,1757	0,1209	0,0763	0,0439
	3	0,0001	0,0008	0,0026	0,0058	0,0105	0,0574	0,1298	0,2013	0,2503	0,2668	0,2522	0,2150	0,1665	0,1172
	4	0,0000	0,0000	0,0001	0,0004	0,0010	0,0112	0,0401	0,0881	0,1460	0,2001	0,2377	0,2508	0,2384	0,2051
	5	0,0000	0,0000	0,0000	0,0000	0,0001	0,0015	0,0085	0,0264	0,0584	0,1029	0,1536	0,2007	0,2340	0,2461
	6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0012	0,0055	0,0162	0,0368	0,0689	0,1115	0,1596	0,2051
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0008	0,0031	0,0090	0,0212	0,0425	0,0746	0,1172
11	0	0,8953	0,8007	0,7153	0,6382	0,5688	0,3138	0,1673	0,0859	0,0422	0,0198	0,0088	0,0036	0,0014	0,0005
	1	0,0995	0,1798	0,2433	0,2925	0,3293	0,3835	0,3248	0,2362	0,1549	0,0932	0,0518	0,0266	0,0125	0,0054
	2	0,0050	0,0183	0,0376	0,0609	0,0867	0,2131	0,2866	0,2953	0,2581	0,1998	0,1395	0,0887	0,0513	0,0269
	3	0,0002	0,0011	0,0035	0,0076	0,0137	0,0710	0,1517	0,2215	0,2581	0,2568	0,2254	0,1774	0,1259	0,0806
	4	0,0000	0,0000	0,0002	0,0006	0,0014	0,0158	0,0536	0,1107	0,1721	0,2201	0,2428	0,2365	0,2060	0,1611
	5	0,0000	0,0000	0,0000	0,0000	0,0001	0,0025	0,0132	0,0388	0,0803	0,1321	0,1830	0,2207	0,2360	0,2256
	6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0023	0,0097	0,0268	0,0566	0,0985	0,1471	0,1931	0,2256
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0017	0,0064	0,0173	0,0379	0,0701	0,1128	0,1611

Fonte: Tabela construída com as funções estatísticas do EXCEL

Continua...

**APÊNDICE I – Distribuição das Probabilidades Binomiais (3)****Probabilidade**

n	x	0,01	0,02	0,03	0,04	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
11	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0011	0,0037	0,0102	0,0234	0,0462	0,0806
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0018	0,0052	0,0126	0,0269
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0007	0,0021	0,0054
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0005
12	0	0,8864	0,7847	0,6938	0,6127	0,5404	0,2824	0,1422	0,0687	0,0317	0,0138	0,0057	0,0022	0,0008	0,0002
	1	0,1074	0,1922	0,2575	0,3064	0,3413	0,3766	0,3012	0,2062	0,1267	0,0712	0,0368	0,0174	0,0075	0,0029
	2	0,0060	0,0216	0,0438	0,0702	0,0988	0,2301	0,2924	0,2835	0,2323	0,1678	0,1088	0,0639	0,0339	0,0161
	3	0,0002	0,0015	0,0045	0,0098	0,0173	0,0852	0,1720	0,2362	0,2581	0,2397	0,1954	0,1419	0,0923	0,0537
	4	0,0000	0,0001	0,0003	0,0009	0,0021	0,0213	0,0683	0,1329	0,1936	0,2311	0,2367	0,2128	0,1700	0,1208
	5	0,0000	0,0000	0,0000	0,0001	0,0002	0,0038	0,0193	0,0532	0,1032	0,1585	0,2039	0,2270	0,2225	0,1934
	6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0005	0,0040	0,0155	0,0401	0,0792	0,1281	0,1766	0,2124	0,2256
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0006	0,0033	0,0115	0,0291	0,0591	0,1009	0,1489	0,1934
	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0024	0,0078	0,0199	0,0420	0,0762	0,1208
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0015	0,0048	0,0125	0,0277	0,0537
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0008	0,0025	0,0068	0,0161
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0010	0,0029
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002
13	0	0,8775	0,7690	0,6730	0,5882	0,5133	0,2542	0,1209	0,0550	0,0238	0,0097	0,0037	0,0013	0,0004	0,0001
	1	0,1152	0,2040	0,2706	0,3186	0,3512	0,3672	0,2774	0,1787	0,1029	0,0540	0,0259	0,0113	0,0045	0,0016
	2	0,0070	0,0250	0,0502	0,0797	0,1109	0,2448	0,2937	0,2680	0,2059	0,1388	0,0836	0,0453	0,0220	0,0095
	3	0,0003	0,0019	0,0057	0,0122	0,0214	0,0997	0,1900	0,2457	0,2517	0,2181	0,1651	0,1107	0,0660	0,0349
	4	0,0000	0,0001	0,0004	0,0013	0,0028	0,0277	0,0838	0,1535	0,2097	0,2337	0,2222	0,1845	0,1350	0,0873
	5	0,0000	0,0000	0,0000	0,0001	0,0003	0,0055	0,0266	0,0691	0,1258	0,1803	0,2154	0,2214	0,1989	0,1571
	6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0008	0,0063	0,0230	0,0559	0,1030	0,1546	0,1968	0,2169	0,2095
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0011	0,0058	0,0186	0,0442	0,0833	0,1312	0,1775	0,2095
	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0011	0,0047	0,0142	0,0336	0,0656	0,1089	0,1571
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0009	0,0034	0,0101	0,0243	0,0495	0,0873
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0022	0,0065	0,0162	0,0349
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0012	0,0036	0,0095
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0016
	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
14	0	0,8687	0,7536	0,6528	0,5647	0,4877	0,2288	0,1028	0,0440	0,0178	0,0068	0,0024	0,0008	0,0002	0,0001
	1	0,1229	0,2153	0,2827	0,3294	0,3593	0,3559	0,2539	0,1539	0,0832	0,0407	0,0181	0,0073	0,0027	0,0009
	2	0,0081	0,0286	0,0568	0,0892	0,1229	0,2570	0,2912	0,2501	0,1802	0,1134	0,0634	0,0317	0,0141	0,0056
	3	0,0003	0,0023	0,0070	0,0149	0,0259	0,1142	0,2056	0,2501	0,2402	0,1943	0,1366	0,0845	0,0462	0,0222
	4	0,0000	0,0001	0,0006	0,0017	0,0037	0,0349	0,0998	0,1720	0,2202	0,2290	0,2022	0,1549	0,1040	0,0611
	5	0,0000	0,0000	0,0000	0,0001	0,0004	0,0078	0,0352	0,0860	0,1468	0,1963	0,2178	0,2066	0,1701	0,1222

Fonte: Tabela construída com as funções estatísticas do EXCEL

Continua...



**APÊNDICE I – Distribuição das Probabilidades Binomiais (4)****Probabilidade**

n	x	0,01	0,02	0,03	0,04	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
14	6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0013	0,0093	0,0322	0,0734	0,1262	0,1759	0,2066	0,2088	0,1833
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0019	0,0092	0,0280	0,0618	0,1082	0,1574	0,1952	0,2095
	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0020	0,0082	0,0232	0,0510	0,0918	0,1398	0,1833
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0018	0,0066	0,0183	0,0408	0,0762	0,1222
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0014	0,0049	0,0136	0,0312	0,0611
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0010	0,0033	0,0093	0,0222
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0019	0,0056
	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002	0,0009
	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
15	0	0,8601	0,7386	0,6333	0,5421	0,4633	0,2059	0,0874	0,0352	0,0134	0,0047	0,0016	0,0005	0,0001	0,0000
	1	0,1303	0,2261	0,2938	0,3388	0,3658	0,3432	0,2312	0,1319	0,0668	0,0305	0,0126	0,0047	0,0016	0,0005
	2	0,0092	0,0323	0,0636	0,0988	0,1348	0,2669	0,2856	0,2309	0,1559	0,0916	0,0476	0,0219	0,0090	0,0032
	3	0,0004	0,0029	0,0085	0,0178	0,0307	0,1285	0,2184	0,2501	0,2252	0,1700	0,1110	0,0634	0,0318	0,0139
	4	0,0000	0,0002	0,0008	0,0022	0,0049	0,0428	0,1156	0,1876	0,2252	0,2186	0,1792	0,1268	0,0780	0,0417
	5	0,0000	0,0000	0,0001	0,0002	0,0006	0,0105	0,0449	0,1032	0,1651	0,2061	0,2123	0,1859	0,1404	0,0916
	6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0019	0,0132	0,0430	0,0917	0,1472	0,1906	0,2066	0,1914	0,1527
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0030	0,0138	0,0393	0,0811	0,1319	0,1771	0,2013	0,1964
	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0005	0,0035	0,0131	0,0348	0,0710	0,1181	0,1647	0,1964
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0007	0,0034	0,0116	0,0298	0,0612	0,1048	0,1527
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0007	0,0030	0,0096	0,0245	0,0515	0,0916
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0024	0,0074	0,0191	0,0417
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0016	0,0052	0,0139
	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0010	0,0032
	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005
16	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	0	0,8515	0,7238	0,6143	0,5204	0,4401	0,1853	0,0743	0,0281	0,0100	0,0033	0,0010	0,0003	0,0001	0,0000
	1	0,1376	0,2363	0,3040	0,3469	0,3706	0,3294	0,2097	0,1126	0,0535	0,0228	0,0087	0,0030	0,0009	0,0002
	2	0,0104	0,0362	0,0705	0,1084	0,1463	0,2745	0,2775	0,2111	0,1336	0,0732	0,0353	0,0150	0,0056	0,0018
	3	0,0005	0,0034	0,0102	0,0211	0,0359	0,1423	0,2285	0,2463	0,2079	0,1465	0,0888	0,0468	0,0215	0,0085
	4	0,0000	0,0002	0,0010	0,0029	0,0061	0,0514	0,1311	0,2001	0,2252	0,2040	0,1553	0,1014	0,0572	0,0278
	5	0,0000	0,0000	0,0001	0,0003	0,0008	0,0137	0,0555	0,1201	0,1802	0,2099	0,2008	0,1623	0,1123	0,0667
	6	0,0000	0,0000	0,0000	0,0000	0,0001	0,0028	0,0180	0,0550	0,1101	0,1649	0,1982	0,1983	0,1684	0,1222
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0045	0,0197	0,0524	0,1010	0,1524	0,1889	0,1969	0,1746
	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0009	0,0055	0,0197	0,0487	0,0923	0,1417	0,1812	0,1964
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0012	0,0058	0,0185	0,0442	0,0840	0,1318	0,1746
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0014	0,0056	0,0167	0,0392	0,0755	0,1222
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0013	0,0049	0,0142	0,0337	0,0667
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0011	0,0040	0,0115	0,0278
	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0008	0,0029	0,0085

Fonte: Tabela construída com as funções estatísticas do EXCEL

Continua...

**APÊNDICE I – Distribuição das Probabilidades Binomiais (5)****Probabilidade**

n	x	0,01	0,02	0,03	0,04	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
16	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0018
	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002
	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
17	0	0,8429	0,7093	0,5958	0,4996	0,4181	0,1668	0,0631	0,0225	0,0075	0,0023	0,0007	0,0002	0,0000	0,0000
	1	0,1447	0,2461	0,3133	0,3539	0,3741	0,3150	0,1893	0,0957	0,0426	0,0169	0,0060	0,0019	0,0005	0,0001
	2	0,0117	0,0402	0,0775	0,1180	0,1575	0,2800	0,2673	0,1914	0,1136	0,0581	0,0260	0,0102	0,0035	0,0010
	3	0,0006	0,0041	0,0120	0,0246	0,0415	0,1556	0,2359	0,2393	0,1893	0,1245	0,0701	0,0341	0,0144	0,0052
	4	0,0000	0,0003	0,0013	0,0036	0,0076	0,0605	0,1457	0,2093	0,2209	0,1868	0,1320	0,0796	0,0411	0,0182
	5	0,0000	0,0000	0,0001	0,0004	0,0010	0,0175	0,0668	0,1361	0,1914	0,2081	0,1849	0,1379	0,0875	0,0472
	6	0,0000	0,0000	0,0000	0,0000	0,0001	0,0039	0,0236	0,0680	0,1276	0,1784	0,1991	0,1839	0,1432	0,0944
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0007	0,0065	0,0267	0,0668	0,1201	0,1685	0,1927	0,1841	0,1484
	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0014	0,0084	0,0279	0,0644	0,1134	0,1606	0,1883	0,1855
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0021	0,0093	0,0276	0,0611	0,1070	0,1540	0,1855
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0025	0,0095	0,0263	0,0571	0,1008	0,1484
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0026	0,0090	0,0242	0,0525	0,0944
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0024	0,0081	0,0215	0,0472
	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0021	0,0068	0,0182
	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0016	0,0052
	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0010
	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
	17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
18	0	0,8345	0,6951	0,5780	0,4796	0,3972	0,1501	0,0536	0,0180	0,0056	0,0016	0,0004	0,0001	0,0000	0,0000
	1	0,1517	0,2554	0,3217	0,3597	0,3763	0,3002	0,1704	0,0811	0,0338	0,0126	0,0042	0,0012	0,0003	0,0001
	2	0,0130	0,0443	0,0846	0,1274	0,1683	0,2835	0,2556	0,1723	0,0958	0,0458	0,0190	0,0069	0,0022	0,0006
	3	0,0007	0,0048	0,0140	0,0283	0,0473	0,1680	0,2406	0,2297	0,1704	0,1046	0,0547	0,0246	0,0095	0,0031
	4	0,0000	0,0004	0,0016	0,0044	0,0093	0,0700	0,1592	0,2153	0,2130	0,1681	0,1104	0,0614	0,0291	0,0117
	5	0,0000	0,0000	0,0001	0,0005	0,0014	0,0218	0,0787	0,1507	0,1988	0,2017	0,1664	0,1146	0,0666	0,0327
	6	0,0000	0,0000	0,0000	0,0000	0,0002	0,0052	0,0301	0,0816	0,1436	0,1873	0,1941	0,1655	0,1181	0,0708
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0010	0,0091	0,0350	0,0820	0,1376	0,1792	0,1892	0,1657	0,1214
	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0022	0,0120	0,0376	0,0811	0,1327	0,1734	0,1864	0,1669
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0033	0,0139	0,0386	0,0794	0,1284	0,1694	0,1855
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0008	0,0042	0,0149	0,0385	0,0771	0,1248	0,1669
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0010	0,0046	0,0151	0,0374	0,0742	0,1214
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0012	0,0047	0,0145	0,0354	0,0708
	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0012	0,0045	0,0134	0,0327
	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0011	0,0039	0,0117
	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0009	0,0031
	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006
	17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
	18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Fonte: Tabela construída com as funções estatísticas do EXCEL

Continua...

**APÊNDICE I – Distribuição das Probabilidades Binomiais (6)****Probabilidade**

n	x	0,01	0,02	0,03	0,04	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
19	0	0,8262	0,6812	0,5606	0,4604	0,3774	0,1351	0,0456	0,0144	0,0042	0,0011	0,0003	0,0001	0,0000	0,0000
	1	0,1586	0,2642	0,3294	0,3645	0,3774	0,2852	0,1529	0,0685	0,0268	0,0093	0,0029	0,0008	0,0002	0,0000
	2	0,0144	0,0485	0,0917	0,1367	0,1787	0,2852	0,2428	0,1540	0,0803	0,0358	0,0138	0,0046	0,0013	0,0003
	3	0,0008	0,0056	0,0161	0,0323	0,0533	0,1796	0,2428	0,2182	0,1517	0,0869	0,0422	0,0175	0,0062	0,0018
	4	0,0000	0,0005	0,0020	0,0054	0,0112	0,0798	0,1714	0,2182	0,2023	0,1491	0,0909	0,0467	0,0203	0,0074
	5	0,0000	0,0000	0,0002	0,0007	0,0018	0,0266	0,0907	0,1636	0,2023	0,1916	0,1468	0,0933	0,0497	0,0222
	6	0,0000	0,0000	0,0000	0,0001	0,0002	0,0069	0,0374	0,0955	0,1574	0,1916	0,1844	0,1451	0,0949	0,0518
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0014	0,0122	0,0443	0,0974	0,1525	0,1844	0,1797	0,1443	0,0961
	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0032	0,0166	0,0487	0,0981	0,1489	0,1797	0,1771	0,1442
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0007	0,0051	0,0198	0,0514	0,0980	0,1464	0,1771	0,1762
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0013	0,0066	0,0220	0,0528	0,0976	0,1449	0,1762
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0018	0,0077	0,0233	0,0532	0,0970	0,1442
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0022	0,0083	0,0237	0,0529	0,0961
	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0024	0,0085	0,0233	0,0518
	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0024	0,0082	0,0222
	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0022	0,0074
	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0018
	17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003
	18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
20	0	0,8179	0,6676	0,5438	0,4420	0,3585	0,1216	0,0388	0,0115	0,0032	0,0008	0,0002	0,0000	0,0000	0,0000
	1	0,1652	0,2725	0,3364	0,3683	0,3774	0,2702	0,1368	0,0576	0,0211	0,0068	0,0020	0,0005	0,0001	0,0000
	2	0,0159	0,0528	0,0988	0,1458	0,1887	0,2852	0,2293	0,1369	0,0669	0,0278	0,0100	0,0031	0,0008	0,0002
	3	0,0010	0,0065	0,0183	0,0364	0,0596	0,1901	0,2428	0,2054	0,1339	0,0716	0,0323	0,0123	0,0040	0,0011
	4	0,0000	0,0006	0,0024	0,0065	0,0133	0,0898	0,1821	0,2182	0,1897	0,1304	0,0738	0,0350	0,0139	0,0046
	5	0,0000	0,0000	0,0002	0,0009	0,0022	0,0319	0,1028	0,1746	0,2023	0,1789	0,1272	0,0746	0,0365	0,0148
	6	0,0000	0,0000	0,0000	0,0001	0,0003	0,0089	0,0454	0,1091	0,1686	0,1916	0,1712	0,1244	0,0746	0,0370
	7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0020	0,0160	0,0545	0,1124	0,1643	0,1844	0,1659	0,1221	0,0739
	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0046	0,0222	0,0609	0,1144	0,1614	0,1797	0,1623	0,1201
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0011	0,0074	0,0271	0,0654	0,1158	0,1597	0,1771	0,1602
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0020	0,0099	0,0308	0,0686	0,1171	0,1593	0,1762
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0005	0,0030	0,0120	0,0336	0,0710	0,1185	0,1602
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0008	0,0039	0,0136	0,0355	0,0727	0,1201
	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0010	0,0045	0,0146	0,0366	0,0739
	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0012	0,0049	0,0150	0,0370
	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0013	0,0049	0,0148
	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0013	0,0046
	17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0011
	18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002
	19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	20	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Fonte: Tabela construída com as funções estatísticas do EXCEL

Continua...

**APÊNDICE I – Distribuição das Probabilidades Binomiais (7)****Probabilidade**

n	x	0,01	0,02	0,03	0,04	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
25	0	0,7778	0,6035	0,4670	0,3604	0,2774	0,0718	0,0172	0,0038	0,0008	0,0001	0,0000	0,0000	0,0000	0,0000
	1	0,1964	0,3079	0,3611	0,3754	0,3650	0,1994	0,0759	0,0236	0,0063	0,0014	0,0003	0,0000	0,0000	0,0000
	2	0,0238	0,0754	0,1340	0,1877	0,2305	0,2659	0,1607	0,0708	0,0251	0,0074	0,0018	0,0004	0,0001	0,0000
	3	0,0018	0,0118	0,0318	0,0600	0,0930	0,2265	0,2174	0,1358	0,0641	0,0243	0,0076	0,0019	0,0004	0,0001
	4	0,0001	0,0013	0,0054	0,0137	0,0269	0,1384	0,2110	0,1867	0,1175	0,0572	0,0224	0,0071	0,0018	0,0004
	5	0,0000	0,0001	0,0007	0,0024	0,0060	0,0646	0,1564	0,1960	0,1645	0,1030	0,0506	0,0199	0,0063	0,0016
	6	0,0000	0,0000	0,0001	0,0003	0,0010	0,0239	0,0920	0,1633	0,1828	0,1472	0,0908	0,0442	0,0172	0,0053
	7	0,0000	0,0000	0,0000	0,0000	0,0001	0,0072	0,0441	0,1108	0,1654	0,1712	0,1327	0,0800	0,0381	0,0143
	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0018	0,0175	0,0623	0,1241	0,1651	0,1607	0,1200	0,0701	0,0322
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0058	0,0294	0,0781	0,1336	0,1635	0,1511	0,1084	0,0609
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0016	0,0118	0,0417	0,0916	0,1409	0,1612	0,1419	0,0974
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0040	0,0189	0,0536	0,1034	0,1465	0,1583	0,1328
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0012	0,0074	0,0268	0,0650	0,1140	0,1511	0,1550
	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0025	0,0115	0,0350	0,0760	0,1236	0,1550
	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0007	0,0042	0,0161	0,0434	0,0867	0,1328
	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0013	0,0064	0,0212	0,0520	0,0974
	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0021	0,0088	0,0266	0,0609
	17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0031	0,0115	0,0322
	18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0009	0,0042	0,0143
	19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0013	0,0053
	20	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0016
	21	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004
	22	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
	23	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	24	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	25	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
28	0	0,7547	0,5680	0,4262	0,3189	0,2378	0,0523	0,0106	0,0019	0,0003	0,0000	0,0000	0,0000	0,0000	0,0000
	1	0,2135	0,3246	0,3691	0,3720	0,3505	0,1628	0,0522	0,0135	0,0030	0,0006	0,0001	0,0000	0,0000	0,0000
	2	0,0291	0,0894	0,1541	0,2092	0,2490	0,2442	0,1243	0,0457	0,0133	0,0032	0,0006	0,0001	0,0000	0,0000
	3	0,0025	0,0158	0,0413	0,0756	0,1136	0,2352	0,1901	0,0990	0,0385	0,0119	0,0030	0,0006	0,0001	0,0000
	4	0,0002	0,0020	0,0080	0,0197	0,0374	0,1633	0,2097	0,1547	0,0803	0,0318	0,0099	0,0025	0,0005	0,0001
	5	0,0000	0,0002	0,0012	0,0039	0,0094	0,0871	0,1776	0,1856	0,1284	0,0654	0,0257	0,0079	0,0019	0,0004
	6	0,0000	0,0000	0,0001	0,0006	0,0019	0,0371	0,1202	0,1779	0,1641	0,1074	0,0530	0,0203	0,0061	0,0014
	7	0,0000	0,0000	0,0000	0,0001	0,0003	0,0130	0,0667	0,1398	0,1719	0,1446	0,0897	0,0426	0,0156	0,0044
	8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0038	0,0309	0,0917	0,1504	0,1627	0,1269	0,0745	0,0335	0,0116
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0009	0,0121	0,0510	0,1114	0,1550	0,1518	0,1103	0,0610	0,0257
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0041	0,0242	0,0706	0,1262	0,1553	0,1398	0,0948	0,0489
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0012	0,0099	0,0385	0,0885	0,1368	0,1525	0,1269	0,0800
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0035	0,0182	0,0537	0,1044	0,1440	0,1471	0,1133
	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0011	0,0075	0,0283	0,0692	0,1181	0,1481	0,1395
	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0027	0,0130	0,0399	0,0844	0,1298	0,1494

Fonte: Tabela construída com as funções estatísticas do EXCEL

Continua...

**APÊNDICE I – Distribuição das Probabilidades Binomiais (8)****Probabilidade**

n	x	0,01	0,02	0,03	0,04	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
28	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0008	0,0052	0,0201	0,0525	0,0991	0,1395
	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0018	0,0088	0,0284	0,0659	0,1133
	17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0005	0,0033	0,0134	0,0381	0,0800
	18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0011	0,0055	0,0190	0,0489
	19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0019	0,0082	0,0257
	20	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0030	0,0116
	21	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0009	0,0044
	22	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0014
	23	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004
	24	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
	25	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	26	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	27	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
29	28	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	5	0,0000	0,0002	0,0014	0,0046	0,0108	0,0947	0,1825	0,1795	0,1164	0,0553	0,0202	0,0058	0,0013	0,0002
	6	0,0000	0,0000	0,0002	0,0008	0,0023	0,0421	0,1288	0,1795	0,1552	0,0948	0,0435	0,0154	0,0042	0,0009
	7	0,0000	0,0000	0,0000	0,0001	0,0004	0,0154	0,0747	0,1474	0,1699	0,1335	0,0769	0,0337	0,0113	0,0029
	8	0,0000	0,0000	0,0000	0,0000	0,0001	0,0047	0,0362	0,1013	0,1558	0,1573	0,1139	0,0617	0,0255	0,0080
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0012	0,0149	0,0591	0,1212	0,1573	0,1431	0,0960	0,0486	0,0187
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0053	0,0296	0,0808	0,1348	0,1541	0,1280	0,0796	0,0373
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0016	0,0128	0,0465	0,0998	0,1433	0,1474	0,1124	0,0644
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0048	0,0233	0,0642	0,1157	0,1474	0,1380	0,0967
	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0016	0,0101	0,0360	0,0815	0,1285	0,1476	0,1264
	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0039	0,0176	0,0502	0,0979	0,1381	0,1445
	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0013	0,0075	0,0270	0,0653	0,1130	0,1445
	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0028	0,0127	0,0381	0,0809	0,1264
	17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0009	0,0052	0,0194	0,0506	0,0967
	18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0019	0,0086	0,0276	0,0644
	19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0033	0,0131	0,0373
	20	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0011	0,0053	0,0187
	21	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0019	0,0080
	22	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0029
	23	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0009
	24	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002
	25	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	26	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	27	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	28	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	29	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Fonte: Tabela construída com as funções estatísticas do EXCEL

Continua...

**APÊNDICE I – Distribuição das Probabilidades Binomiais (9)****Probabilidade**

n	x	0,01	0,02	0,03	0,04	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
30	0	0,7397	0,5455	0,4010	0,2939	0,2146	0,0424	0,0076	0,0012	0,0002	0,0000	0,0000	0,0000	0,0000	0,0000
	1	0,2242	0,3340	0,3721	0,3673	0,3389	0,1413	0,0404	0,0093	0,0018	0,0003	0,0000	0,0000	0,0000	0,0000
	2	0,0328	0,0988	0,1669	0,2219	0,2586	0,2277	0,1034	0,0337	0,0086	0,0018	0,0003	0,0000	0,0000	0,0000
	3	0,0031	0,0188	0,0482	0,0863	0,1270	0,2361	0,1703	0,0785	0,0269	0,0072	0,0015	0,0003	0,0000	0,0000
	4	0,0002	0,0026	0,0101	0,0243	0,0451	0,1771	0,2028	0,1325	0,0604	0,0208	0,0056	0,0012	0,0002	0,0000
	5	0,0000	0,0003	0,0016	0,0053	0,0124	0,1023	0,1861	0,1723	0,1047	0,0464	0,0157	0,0041	0,0008	0,0001
	6	0,0000	0,0000	0,0002	0,0009	0,0027	0,0474	0,1368	0,1795	0,1455	0,0829	0,0353	0,0115	0,0029	0,0006
	7	0,0000	0,0000	0,0000	0,0001	0,0005	0,0180	0,0828	0,1538	0,1662	0,1219	0,0652	0,0263	0,0081	0,0019
	8	0,0000	0,0000	0,0000	0,0000	0,0001	0,0058	0,0420	0,1106	0,1593	0,1501	0,1009	0,0505	0,0191	0,0055
	9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0016	0,0181	0,0676	0,1298	0,1573	0,1328	0,0823	0,0382	0,0133
	10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0004	0,0067	0,0355	0,0909	0,1416	0,1502	0,1152	0,0656	0,0280
	11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0022	0,0161	0,0551	0,1103	0,1471	0,1396	0,0976	0,0509
	12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0006	0,0064	0,0291	0,0749	0,1254	0,1474	0,1265	0,0806
	13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0022	0,0134	0,0444	0,0935	0,1360	0,1433	0,1115
	14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0007	0,0054	0,0231	0,0611	0,1101	0,1424	0,1354
	15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0019	0,0106	0,0351	0,0783	0,1242	0,1445
	16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0006	0,0042	0,0177	0,0489	0,0953	0,1354
	17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0015	0,0079	0,0269	0,0642	0,1115
	18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0005	0,0031	0,0129	0,0379	0,0806
	19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0010	0,0054	0,0196	0,0509
	20	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0020	0,0088	0,0280
	21	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0034	0,0133
	22	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0012	0,0055
	23	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0003	0,0019
	24	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006
	25	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
	26	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	27	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	28	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	29	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
	30	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Fonte: Tabela construída com as funções estatísticas do EXCEL

**APÊNDICE II – Distribuição da probabilidades de Poisson (1)**

μ																				
x	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0	1,1	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	2,0
0	0,9048	0,8187	0,7408	0,6703	0,6065	0,5488	0,4966	0,4493	0,4066	0,3679	0,3329	0,3012	0,2725	0,2466	0,2231	0,2019	0,1827	0,1653	0,1496	0,1353
1	0,0905	0,1637	0,2222	0,2681	0,3033	0,3293	0,3476	0,3595	0,3659	0,3679	0,3662	0,3614	0,3543	0,3452	0,3347	0,3230	0,3106	0,2975	0,2842	0,2707
2	0,0045	0,0164	0,0333	0,0536	0,0758	0,0988	0,1217	0,1438	0,1647	0,1839	0,2014	0,2169	0,2303	0,2417	0,2510	0,2584	0,2640	0,2678	0,2700	0,2707
3	0,0002	0,0011	0,0033	0,0072	0,0126	0,0198	0,0284	0,0383	0,0494	0,0613	0,0738	0,0867	0,0998	0,1128	0,1255	0,1378	0,1496	0,1607	0,1710	0,1804
4	0,0000	0,0001	0,0003	0,0007	0,0016	0,0030	0,0050	0,0077	0,0111	0,0153	0,0203	0,0260	0,0324	0,0395	0,0471	0,0551	0,0636	0,0723	0,0812	0,0902
5	0,0000	0,0000	0,0000	0,0001	0,0002	0,0004	0,0007	0,0012	0,0020	0,0031	0,0045	0,0062	0,0084	0,0111	0,0141	0,0176	0,0216	0,0260	0,0309	0,0361
6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002	0,0003	0,0005	0,0008	0,0012	0,0018	0,0026	0,0035	0,0047	0,0061	0,0078	0,0098	0,0120
7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0002	0,0003	0,0005	0,0008	0,0011	0,0015	0,0020	0,0027	0,0034
8	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0002	0,0003	0,0005	0,0006	0,0009
9	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0002
10	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Fonte: Tabela construída com as funções estatísticas do EXCEL

**APÊNDICE II – Distribuição da probabilidades de Poisson (2)**

μ																				
x	2,1	2,2	2,3	2,4	2,5	2,6	2,7	2,8	2,9	3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,7	3,8	3,9	4,0
0	0,1225	0,1108	0,1003	0,0907	0,0821	0,0743	0,0672	0,0608	0,0550	0,0498	0,0450	0,0408	0,0369	0,0334	0,0302	0,0273	0,0247	0,0224	0,0202	0,0183
1	0,2572	0,2438	0,2306	0,2177	0,2052	0,1931	0,1815	0,1703	0,1596	0,1494	0,1397	0,1304	0,1217	0,1135	0,1057	0,0984	0,0915	0,0850	0,0789	0,0733
2	0,2700	0,2681	0,2652	0,2613	0,2565	0,2510	0,2450	0,2384	0,2314	0,2240	0,2165	0,2087	0,2008	0,1929	0,1850	0,1771	0,1692	0,1615	0,1539	0,1465
3	0,1890	0,1966	0,2033	0,2090	0,2138	0,2176	0,2205	0,2225	0,2237	0,2240	0,2237	0,2226	0,2209	0,2186	0,2158	0,2125	0,2087	0,2046	0,2001	0,1954
4	0,0992	0,1082	0,1169	0,1254	0,1336	0,1414	0,1488	0,1557	0,1622	0,1680	0,1733	0,1781	0,1823	0,1858	0,1888	0,1912	0,1931	0,1944	0,1951	0,1954
5	0,0417	0,0476	0,0538	0,0602	0,0668	0,0735	0,0804	0,0872	0,0940	0,1008	0,1075	0,1140	0,1203	0,1264	0,1322	0,1377	0,1429	0,1477	0,1522	0,1563
6	0,0146	0,0174	0,0206	0,0241	0,0278	0,0319	0,0362	0,0407	0,0455	0,0504	0,0555	0,0608	0,0662	0,0716	0,0771	0,0826	0,0881	0,0936	0,0989	0,1042
7	0,0044	0,0055	0,0068	0,0083	0,0099	0,0118	0,0139	0,0163	0,0188	0,0216	0,0246	0,0278	0,0312	0,0348	0,0385	0,0425	0,0466	0,0508	0,0551	0,0595
8	0,0011	0,0015	0,0019	0,0025	0,0031	0,0038	0,0047	0,0057	0,0068	0,0081	0,0095	0,0111	0,0129	0,0148	0,0169	0,0191	0,0215	0,0241	0,0269	0,0298
9	0,0003	0,0004	0,0005	0,0007	0,0009	0,0011	0,0014	0,0018	0,0022	0,0027	0,0033	0,0040	0,0047	0,0056	0,0066	0,0076	0,0089	0,0102	0,0116	0,0132
10	0,0001	0,0001	0,0001	0,0002	0,0002	0,0003	0,0004	0,0005	0,0006	0,0008	0,0010	0,0013	0,0016	0,0019	0,0023	0,0028	0,0033	0,0039	0,0045	0,0053
11	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0002	0,0002	0,0003	0,0004	0,0005	0,0006	0,0007	0,0009	0,0011	0,0013	0,0016	0,0019
12	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002	0,0003	0,0003	0,0004	0,0005	0,0006
13	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002
14	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
15	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Fonte: Tabela construída com as funções estatísticas do EXCEL

**APÊNDICE II – Distribuição da probabilidades de Poisson (3)**

x	$\mu$																			
	4,1	4,2	4,3	4,4	4,5	4,6	4,7	4,8	4,9	5,0	5,1	5,2	5,3	5,4	5,5	5,6	5,7	5,8	5,9	6,0
0	0,0166	0,0150	0,0136	0,0123	0,0111	0,0101	0,0091	0,0082	0,0074	0,0067	0,0061	0,0055	0,0050	0,0045	0,0041	0,0037	0,0033	0,0030	0,0027	0,0025
1	0,0679	0,0630	0,0583	0,0540	0,0500	0,0462	0,0427	0,0395	0,0365	0,0337	0,0311	0,0287	0,0265	0,0244	0,0225	0,0207	0,0191	0,0176	0,0162	0,0149
2	0,1393	0,1323	0,1254	0,1188	0,1125	0,1063	0,1005	0,0948	0,0894	0,0842	0,0793	0,0746	0,0701	0,0659	0,0618	0,0580	0,0544	0,0509	0,0477	0,0446
3	0,1904	0,1852	0,1798	0,1743	0,1687	0,1631	0,1574	0,1517	0,1460	0,1404	0,1348	0,1293	0,1239	0,1185	0,1133	0,1082	0,1033	0,0985	0,0938	0,0892
4	0,1951	0,1944	0,1933	0,1917	0,1898	0,1875	0,1849	0,1820	0,1789	0,1755	0,1719	0,1681	0,1641	0,1600	0,1558	0,1515	0,1472	0,1428	0,1383	0,1339
5	0,1600	0,1633	0,1662	0,1687	0,1708	0,1725	0,1738	0,1747	0,1753	0,1755	0,1753	0,1748	0,1740	0,1728	0,1714	0,1697	0,1678	0,1656	0,1632	0,1606
6	0,1093	0,1143	0,1191	0,1237	0,1281	0,1323	0,1362	0,1398	0,1432	0,1462	0,1490	0,1515	0,1537	0,1555	0,1571	0,1584	0,1594	0,1601	0,1605	0,1606
7	0,0640	0,0686	0,0732	0,0778	0,0824	0,0869	0,0914	0,0959	0,1002	0,1044	0,1086	0,1125	0,1163	0,1200	0,1234	0,1267	0,1298	0,1326	0,1353	0,1377
8	0,0328	0,0360	0,0393	0,0428	0,0463	0,0500	0,0537	0,0575	0,0614	0,0653	0,0692	0,0731	0,0771	0,0810	0,0849	0,0887	0,0925	0,0962	0,0998	0,1033
9	0,0150	0,0168	0,0188	0,0209	0,0232	0,0255	0,0281	0,0307	0,0334	0,0363	0,0392	0,0423	0,0454	0,0486	0,0519	0,0552	0,0586	0,0620	0,0654	0,0688
10	0,0061	0,0071	0,0081	0,0092	0,0104	0,0118	0,0132	0,0147	0,0164	0,0181	0,0200	0,0220	0,0241	0,0262	0,0285	0,0309	0,0334	0,0359	0,0386	0,0413
11	0,0023	0,0027	0,0032	0,0037	0,0043	0,0049	0,0056	0,0064	0,0073	0,0082	0,0093	0,0104	0,0116	0,0129	0,0143	0,0157	0,0173	0,0190	0,0207	0,0225
12	0,0008	0,0009	0,0011	0,0013	0,0016	0,0019	0,0022	0,0026	0,0030	0,0034	0,0039	0,0045	0,0051	0,0058	0,0065	0,0073	0,0082	0,0092	0,0102	0,0113
13	0,0002	0,0003	0,0004	0,0005	0,0006	0,0007	0,0008	0,0009	0,0011	0,0013	0,0015	0,0018	0,0021	0,0024	0,0028	0,0032	0,0036	0,0041	0,0046	0,0052
14	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002	0,0003	0,0003	0,0004	0,0005	0,0006	0,0007	0,0008	0,0009	0,0011	0,0013	0,0015	0,0017	0,0019	0,0022
15	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002	0,0002	0,0003	0,0003	0,0004	0,0005	0,0006	0,0007	0,0008	0,0009
16	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002	0,0002	0,0003	0,0003
17	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001

Fonte: Tabela construída com as funções estatísticas do EXCEL



**APÊNDICE II – Distribuição da probabilidades de Poisson (4)**

x	$\mu$																			
	6,1	6,2	6,3	6,4	6,5	6,6	6,7	6,8	6,9	7,0	7,1	7,2	7,3	7,4	7,5	7,6	7,7	7,8	7,9	8,0
0	0,0022	0,0020	0,0018	0,0017	0,0015	0,0014	0,0012	0,0011	0,0010	0,0009	0,0008	0,0007	0,0007	0,0006	0,0006	0,0005	0,0005	0,0004	0,0004	0,0003
1	0,0137	0,0126	0,0116	0,0106	0,0098	0,0090	0,0082	0,0076	0,0070	0,0064	0,0059	0,0054	0,0049	0,0045	0,0041	0,0038	0,0035	0,0032	0,0029	0,0027
2	0,0417	0,0390	0,0364	0,0340	0,0318	0,0296	0,0276	0,0258	0,0240	0,0223	0,0208	0,0194	0,0180	0,0167	0,0156	0,0145	0,0134	0,0125	0,0116	0,0107
3	0,0848	0,0806	0,0765	0,0726	0,0688	0,0652	0,0617	0,0584	0,0552	0,0521	0,0492	0,0464	0,0438	0,0413	0,0389	0,0366	0,0345	0,0324	0,0305	0,0286
4	0,1294	0,1249	0,1205	0,1162	0,1118	0,1076	0,1034	0,0992	0,0952	0,0912	0,0874	0,0836	0,0799	0,0764	0,0729	0,0696	0,0663	0,0632	0,0602	0,0573
5	0,1579	0,1549	0,1519	0,1487	0,1454	0,1420	0,1385	0,1349	0,1314	0,1277	0,1241	0,1204	0,1167	0,1130	0,1094	0,1057	0,1021	0,0986	0,0951	0,0916
6	0,1605	0,1601	0,1595	0,1586	0,1575	0,1562	0,1546	0,1529	0,1511	0,1490	0,1468	0,1445	0,1420	0,1394	0,1367	0,1339	0,1311	0,1282	0,1252	0,1221
7	0,1399	0,1418	0,1435	0,1450	0,1462	0,1472	0,1480	0,1486	0,1489	0,1490	0,1489	0,1486	0,1481	0,1474	0,1465	0,1454	0,1442	0,1428	0,1413	0,1396
8	0,1066	0,1099	0,1130	0,1160	0,1188	0,1215	0,1240	0,1263	0,1284	0,1304	0,1321	0,1337	0,1351	0,1363	0,1373	0,1381	0,1388	0,1392	0,1395	0,1396
9	0,0723	0,0757	0,0791	0,0825	0,0858	0,0891	0,0923	0,0954	0,0985	0,1014	0,1042	0,1070	0,1096	0,1121	0,1144	0,1167	0,1187	0,1207	0,1224	0,1241
10	0,0441	0,0469	0,0498	0,0528	0,0558	0,0588	0,0618	0,0649	0,0679	0,0710	0,0740	0,0770	0,0800	0,0829	0,0858	0,0887	0,0914	0,0941	0,0967	0,0993
11	0,0244	0,0265	0,0285	0,0307	0,0330	0,0353	0,0377	0,0401	0,0426	0,0452	0,0478	0,0504	0,0531	0,0558	0,0585	0,0613	0,0640	0,0667	0,0695	0,0722
12	0,0124	0,0137	0,0150	0,0164	0,0179	0,0194	0,0210	0,0227	0,0245	0,0263	0,0283	0,0303	0,0323	0,0344	0,0366	0,0388	0,0411	0,0434	0,0457	0,0481
13	0,0058	0,0065	0,0073	0,0081	0,0089	0,0099	0,0108	0,0119	0,0130	0,0142	0,0154	0,0168	0,0181	0,0196	0,0211	0,0227	0,0243	0,0260	0,0278	0,0296
14	0,0025	0,0029	0,0033	0,0037	0,0041	0,0046	0,0052	0,0058	0,0064	0,0071	0,0078	0,0086	0,0095	0,0104	0,0113	0,0123	0,0134	0,0145	0,0157	0,0169
15	0,0010	0,0012	0,0014	0,0016	0,0018	0,0020	0,0023	0,0026	0,0029	0,0033	0,0037	0,0041	0,0046	0,0051	0,0057	0,0062	0,0069	0,0075	0,0083	0,0090
16	0,0004	0,0005	0,0005	0,0006	0,0007	0,0008	0,0010	0,0011	0,0013	0,0014	0,0016	0,0019	0,0021	0,0024	0,0026	0,0030	0,0033	0,0037	0,0041	0,0045
17	0,0001	0,0002	0,0002	0,0002	0,0003	0,0003	0,0004	0,0004	0,0005	0,0006	0,0007	0,0008	0,0009	0,0010	0,0012	0,0013	0,0015	0,0017	0,0019	0,0021
18	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002	0,0002	0,0003	0,0003	0,0004	0,0004	0,0005	0,0006	0,0006	0,0007	0,0008	0,0009
19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002	0,0002	0,0003	0,0003	0,0004
20	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002
21	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001
22	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Fonte: Tabela construída com as funções estatísticas do EXCEL

**APÊNDICE II – Distribuição da probabilidades de Poisson (5)**

x	$\mu$																			
	8,1	8,2	8,3	8,4	8,5	8,6	8,7	8,8	8,9	9,0	9,1	9,2	9,3	9,4	9,5	9,6	9,7	9,8	9,9	10,0
0	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0000
1	0,0025	0,0023	0,0021	0,0019	0,0017	0,0016	0,0014	0,0013	0,0012	0,0011	0,0010	0,0009	0,0009	0,0008	0,0007	0,0007	0,0006	0,0005	0,0005	0,0005
2	0,0100	0,0092	0,0086	0,0079	0,0074	0,0068	0,0063	0,0058	0,0054	0,0050	0,0046	0,0043	0,0040	0,0037	0,0034	0,0031	0,0029	0,0027	0,0025	0,0023
3	0,0269	0,0252	0,0237	0,0222	0,0208	0,0195	0,0183	0,0171	0,0160	0,0150	0,0140	0,0131	0,0123	0,0115	0,0107	0,0100	0,0093	0,0087	0,0081	0,0076
4	0,0544	0,0517	0,0491	0,0466	0,0443	0,0420	0,0398	0,0377	0,0357	0,0337	0,0319	0,0302	0,0285	0,0269	0,0254	0,0240	0,0226	0,0213	0,0201	0,0189
5	0,0882	0,0849	0,0816	0,0784	0,0752	0,0722	0,0692	0,0663	0,0635	0,0607	0,0581	0,0555	0,0530	0,0506	0,0483	0,0460	0,0439	0,0418	0,0398	0,0378
6	0,1191	0,1160	0,1128	0,1097	0,1066	0,1034	0,1003	0,0972	0,0941	0,0911	0,0881	0,0851	0,0822	0,0793	0,0764	0,0736	0,0709	0,0682	0,0656	0,0631
7	0,1378	0,1358	0,1338	0,1317	0,1294	0,1271	0,1247	0,1222	0,1197	0,1171	0,1145	0,1118	0,1091	0,1064	0,1037	0,1010	0,0982	0,0955	0,0928	0,0901
8	0,1395	0,1392	0,1388	0,1382	0,1375	0,1366	0,1356	0,1344	0,1332	0,1318	0,1302	0,1286	0,1269	0,1251	0,1232	0,1212	0,1191	0,1170	0,1148	0,1126
9	0,1256	0,1269	0,1280	0,1290	0,1299	0,1306	0,1311	0,1315	0,1317	0,1318	0,1317	0,1315	0,1311	0,1306	0,1300	0,1293	0,1284	0,1274	0,1263	0,1251
10	0,1017	0,1040	0,1063	0,1084	0,1104	0,1123	0,1140	0,1157	0,1172	0,1186	0,1198	0,1210	0,1219	0,1228	0,1235	0,1241	0,1245	0,1249	0,1250	0,1251
11	0,0749	0,0776	0,0802	0,0828	0,0853	0,0878	0,0902	0,0925	0,0948	0,0970	0,0991	0,1012	0,1031	0,1049	0,1067	0,1083	0,1098	0,1112	0,1125	0,1137
12	0,0505	0,0530	0,0555	0,0579	0,0604	0,0629	0,0654	0,0679	0,0703	0,0728	0,0752	0,0776	0,0799	0,0822	0,0844	0,0866	0,0888	0,0908	0,0928	0,0948
13	0,0315	0,0334	0,0354	0,0374	0,0395	0,0416	0,0438	0,0459	0,0481	0,0504	0,0526	0,0549	0,0572	0,0594	0,0617	0,0640	0,0662	0,0685	0,0707	0,0729
14	0,0182	0,0196	0,0210	0,0225	0,0240	0,0256	0,0272	0,0289	0,0306	0,0324	0,0342	0,0361	0,0380	0,0399	0,0419	0,0439	0,0459	0,0479	0,0500	0,0521
15	0,0098	0,0107	0,0116	0,0126	0,0136	0,0147	0,0158	0,0169	0,0182	0,0194	0,0208	0,0221	0,0235	0,0250	0,0265	0,0281	0,0297	0,0313	0,0330	0,0347
16	0,0050	0,0055	0,0060	0,0066	0,0072	0,0079	0,0086	0,0093	0,0101	0,0109	0,0118	0,0127	0,0137	0,0147	0,0157	0,0168	0,0180	0,0192	0,0204	0,0217
17	0,0024	0,0026	0,0029	0,0033	0,0036	0,0040	0,0044	0,0048	0,0053	0,0058	0,0063	0,0069	0,0075	0,0081	0,0088	0,0095	0,0103	0,0111	0,0119	0,0128
18	0,0011	0,0012	0,0014	0,0015	0,0017	0,0019	0,0021	0,0024	0,0026	0,0029	0,0032	0,0035	0,0039	0,0042	0,0046	0,0051	0,0055	0,0060	0,0065	0,0071
19	0,0005	0,0005	0,0006	0,0007	0,0008	0,0009	0,0010	0,0011	0,0012	0,0014	0,0015	0,0017	0,0019	0,0021	0,0023	0,0026	0,0028	0,0031	0,0034	0,0037
20	0,0002	0,0002	0,0002	0,0003	0,0003	0,0004	0,0004	0,0005	0,0005	0,0006	0,0007	0,0008	0,0009	0,0010	0,0011	0,0012	0,0014	0,0015	0,0017	0,0019
21	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002	0,0002	0,0002	0,0003	0,0003	0,0003	0,0004	0,0004	0,0005	0,0006	0,0006	0,0007	0,0008	0,0009
22	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002	0,0002	0,0002	0,0003	0,0003	0,0004	0,0004
23	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0002	0,0002
24	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0001	0,0001
25	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Fonte: Tabela construída com as funções estatísticas do EXCEL

**APÊNDICE III – Distribuição Normal Padronizada - z**

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4,0	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4,1	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4,2	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4,3	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4,4	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4,5	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

Fonte: Tabela construída com as funções estatísticas do EXCEL

**APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 1**

1.1

- a) qualitativa e nominal
- b) qualitativa e nominal
- c) quantitativa e razão
- d) quantitativa e razão
- e) qualitativa e nominal
- f) qualitativa e ordinal ou quantitativa e intervalar
- g) quantitativa e intervalar
- h) quantitativa e razão
- i) qualitativa e ordinal
- j) qualitativa e nominal

1.2

- a) discreta
- b) contínua
- c) contínua
- d) discreta
- e) contínua
- f) discreta
- g) contínua
- h) discreta
- i) discreta
- j) contínua

1.3

- a) 68 e 68,5
- b) 323 e 322,6
- c) 30 e 29,7
- d) 204 e 203,9
- e) 218 e 217,5
- f) 13 e 13,5
- g) 70 e 70,1
- h) 129 e 129,0
- i) 154 e 154,0
- j) 92 e 91,7

1.4

- a) R\$ 800,00; R\$ 910,00; R\$ 985,00; R\$ 1.040,00; R\$ 1.340,00; R\$ 1.590,00; R\$ 1.750,00; R\$ 2.100,00; R\$ 2.620,00; R\$ 3.735,00.
- b) R\$ 2.935,00

1.5

- a) 0; 1; 2; 2; 3; 4; 4; 5; 8; 9; 12; 14.
- b) 15

**APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 2**

2.1

- a) V
- b) V
- c) F
- d) F

2.2

- a) V
- b) F
- c) F
- d) F
- e) F

2.3

a) Tabela 2.3 – Consumo mensal de energia elétrica, em kWh, em uma residência do Vale do Taquari, no período de outubro de 2013 a setembro de 2014

Consumo energia (em Kw/h)	$f_i$	$F_i$	$f_r$	$F_r$	$X_i$
210  — 267	5	5	41,67	41,67	238,5
267  — 324	4	9	33,33	75,00	295,5
324  — 381	1	10	8,33	83,33	352,5
381  — 438	1	11	8,33	91,67	409,5
438  — 495	1	12	8,33	100	466,5
<b>Total</b>	<b>12</b>		<b>100</b>		

Fonte: Dados fictícios, apenas para fins ilustrativos.

2.4

a) Tabela 2.4 – Número diário de sapatos produzidos com defeito na empresa SAPATO NOVO durante determinado mês de 2014

Peças rejeitadas	$f_i$	$F_i$	$f_r$	$F_r$	$X_i$
0  — 6	8	8	28,57	28,57	3
7  — 13	6	14	21,43	50,00	10
14  — 20	7	21	25,00	75,00	17
21  — 27	4	25	14,29	89,29	24
28  — 34	2	27	7,14	96,43	31
35  — 41	1	28	3,57	100	38
<b>Total</b>	<b>28</b>		<b>100</b>		

Fonte: Empresa Sapato Novo.

2.5

a) Tabela 2.5 – O gasto mensal com alimentos consumidos na residência de uma família residente em Lajeado, ao longo de 2013

Valor gasto (em R\$)	$f_i$	$F_i$	$f_r$	$F_r$	$X_i$
860,00  — 980,00	5	5	41,67	41,67	920,00
980,00  — 1.100,00	1	6	8,33	50,00	1.040,00
1.100,00  — 1.220,00	1	7	8,33	58,33	1.160,00
1.220,00  — 1.340,00	2	9	16,67	75,00	1.280,00
1.340,00  — 1.460,00	3	12	25,00	100,00	1.400,00
<b>Total</b>	<b>12</b>		<b>100</b>		

Fonte: Dados fictícios, apenas para fins ilustrativos.

## 2.6

a) Tabela 2.6 – Salários pagos aos funcionários da empresa JKL Marketing, no mês de março de 2014

Salário pago (em R\$)	$f_i$	$F_i$	$f_r$	$F_r$	$X_i$
500,00  — 1.000,00	4	4	25,00	25,00	750,00
1.000,00  — 1.500,00	3	7	18,75	43,75	1.250,00
1.500,00  — 2.000,00	4	11	25,00	68,75	1.750,00
2.000,00  — 2.500,00	2	13	12,50	81,25	2.250,00
2.500,00  — 3.000,00	2	15	12,50	93,75	2.750,00
3.000,00  — 3.500,00	1	16	6,25	100,00	3.250,00
<b>Total</b>	<b>16</b>		<b>100</b>		

Fonte: Empresa JKL Marketing.

## 2.7

a); b) Tabela 2.7 – Número de encomendas (em toneladas) num semestre dos clientes de uma empresa exportadora

Encomendas em toneladas	$f_i$	$F_i$	$f_r$	$F_r$	$X_i$
3,0  — 3,5	5	5	25,00	25,00	3,3
3,5  — 4,0	0	5	0,00	25,00	3,8
4,0  — 4,5	3	8	15,00	40,00	4,3
4,5  — 5,0	0	8	0,00	40,00	4,8
5,0  — 5,5	8	16	40,00	80,00	5,3
5,5  — 6,0	4	20	20,00	100	5,8
<b>Total</b>	<b>20</b>		<b>100</b>		

Fonte: Empresa exportadora.

## 2.8

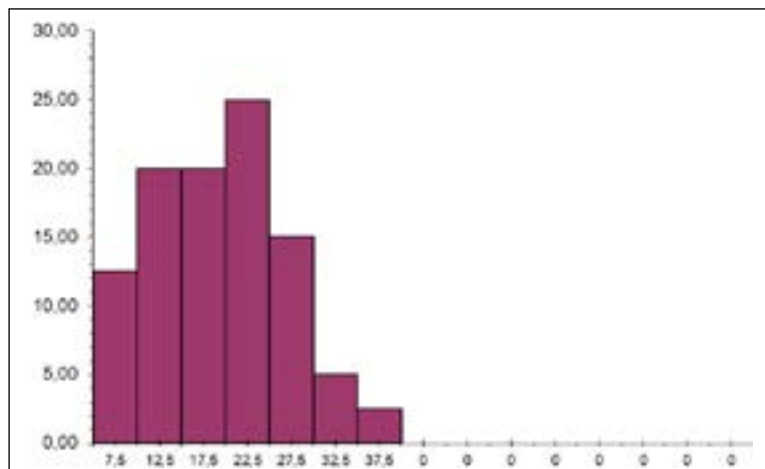
a); b) Tabela 2.8 – Número de empréstimos para habitação, de 50 bancos comerciais, concedidos durante um mês

Número de empréstimos	$f_i$	$F_i$	$f_r$	$F_r$	$X_i$
0  — 4	20	20	40,00	40,00	2,5
5  — 9	17	37	34,00	74,00	7,5
10  — 14	5	42	10,00	84,00	12,5
15  — 19	4	46	8,00	92,00	17,5
20  — 24	2	48	4,00	96,00	22,5
25  — 29	1	49	2,00	98,00	27,5
30  — 34	1	50	2,00	100	32,5
<b>Total</b>	<b>50</b>		<b>100</b>		

Fonte: Bancos comerciais.

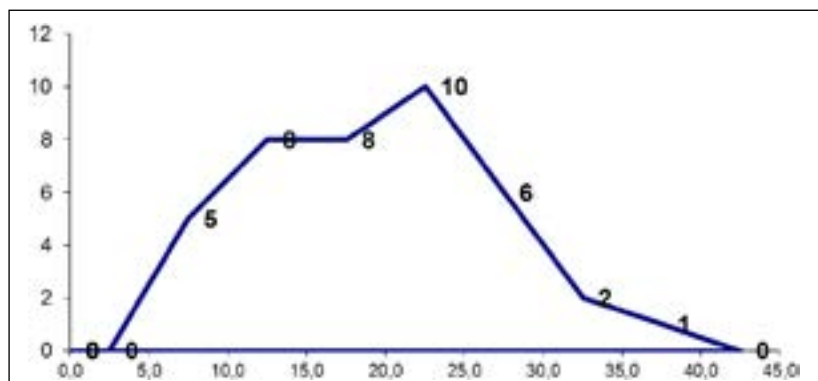
**APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 3****3.1**

a) Gráfico 3.1.a – Histograma de frequências de uma amostra de 40 pessoas que fizeram declaração de renda no ano base de 2014, em Rincão dos Anjos



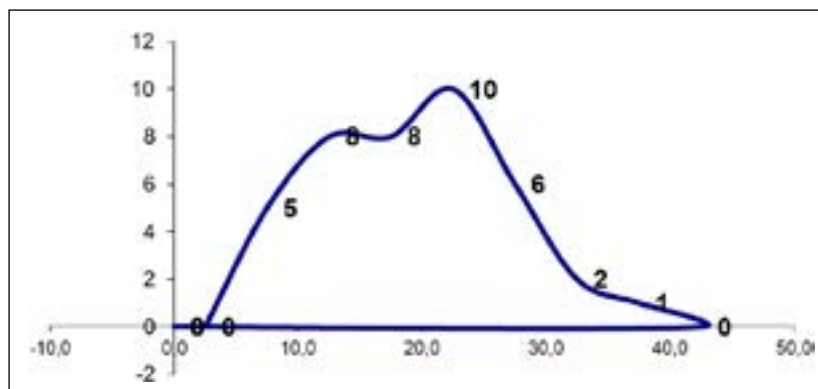
Fonte: Amostra de 40 pessoas que fizeram declaração do Imposto de Renda.

b) Gráfico 3.1.b – Polígono de frequências de uma amostra de 40 pessoas que fizeram declaração de renda no ano base de 2014, em Rincão dos Anjos



Fonte: Amostra de 40 pessoas que fizeram declaração do Imposto de Renda.

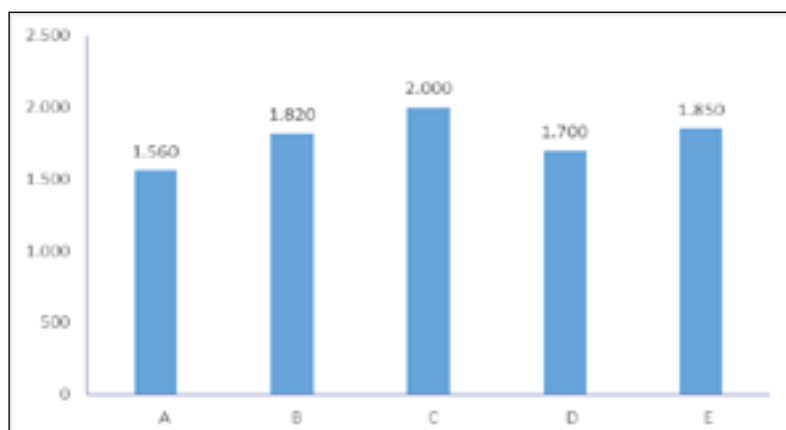
c) Gráfico 3.1.c – Curva de frequências de uma amostra de 40 pessoas que fizeram declaração de renda no ano base de 2014, em Rincão dos Anjos



Fonte: Amostra de 40 pessoas que fizeram declaração do Imposto de Renda.

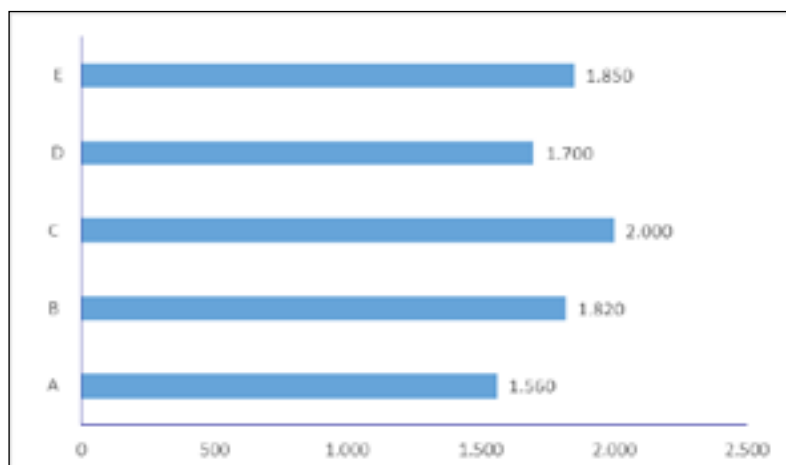
### 3.2

a) Gráfico 3.2.a – Produtividade (em kg) das variedades de fumo na empresa agrícola Verde Vida, safra de 2013/2014



Fonte: Empresa agrícola Verde Vida.

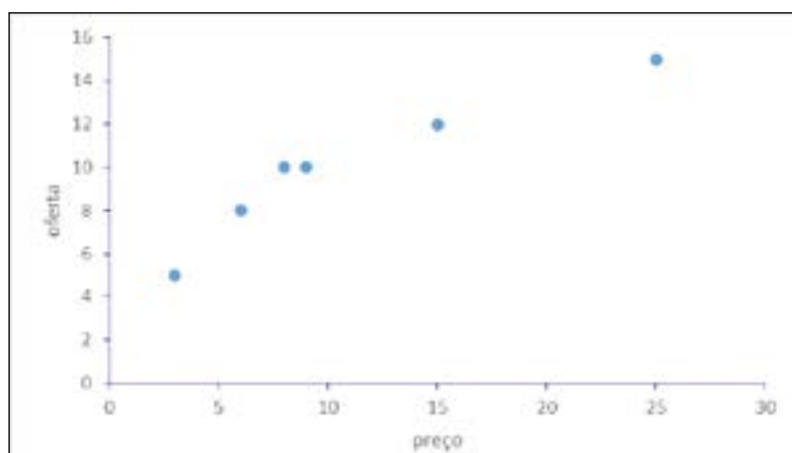
b) Gráfico 3.2.b – Produtividade (em kg) das variedades de fumo na empresa agrícola Verde Vida, safra de 2013/2014



Fonte: Empresa agrícola Verde Vida.

### 3.3

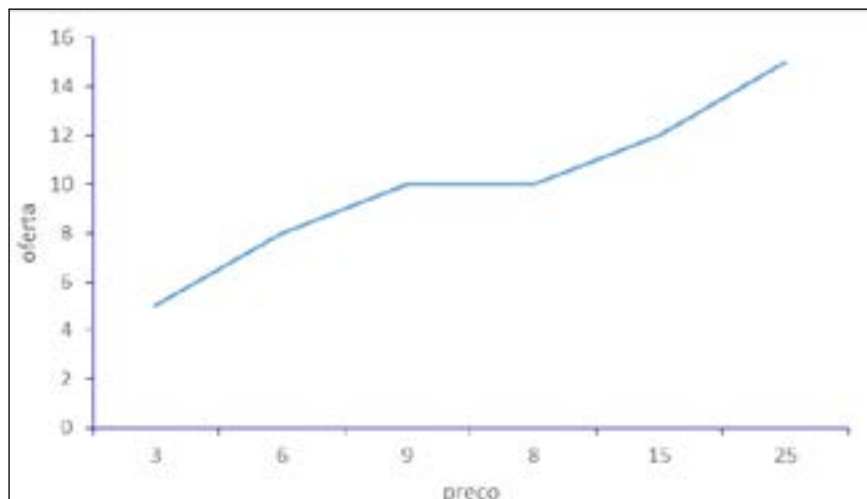
a) Gráfico 3.3.a – Relação do preço (R\$) e oferta (unidades) de um bem em determinado município gaúcho no mês de junho de 2014



Fonte: Determinado município gaúcho.



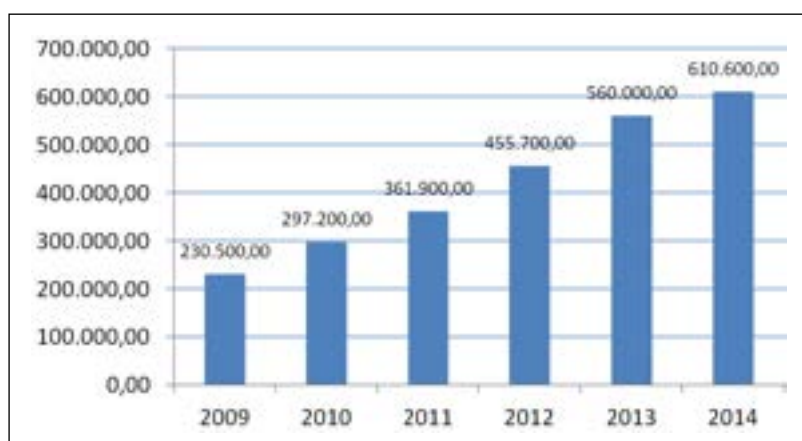
b) Gráfico 3.3.b – Relação do preço (R\$) e oferta (unidades) de um bem em determinado município gaúcho no mês de junho de 2014



Fonte: Determinado município gaúcho.

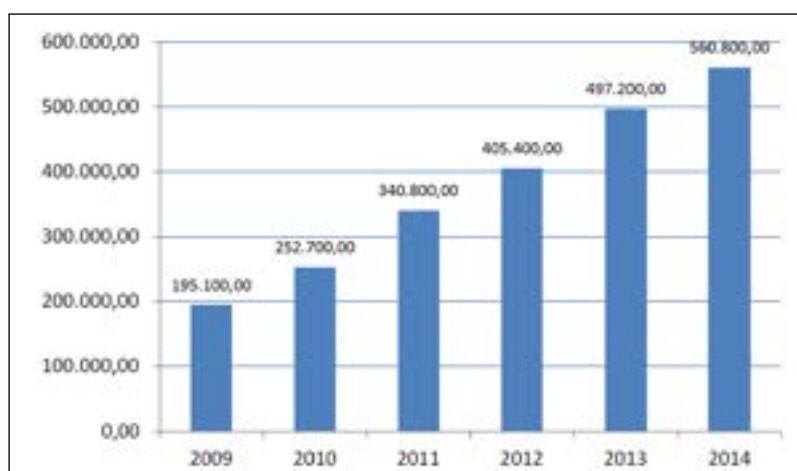
### 3.4

a) Gráfico 3.4.a1 – Faturamento (R\$) de uma empresa de Lajeado – 2009/2014



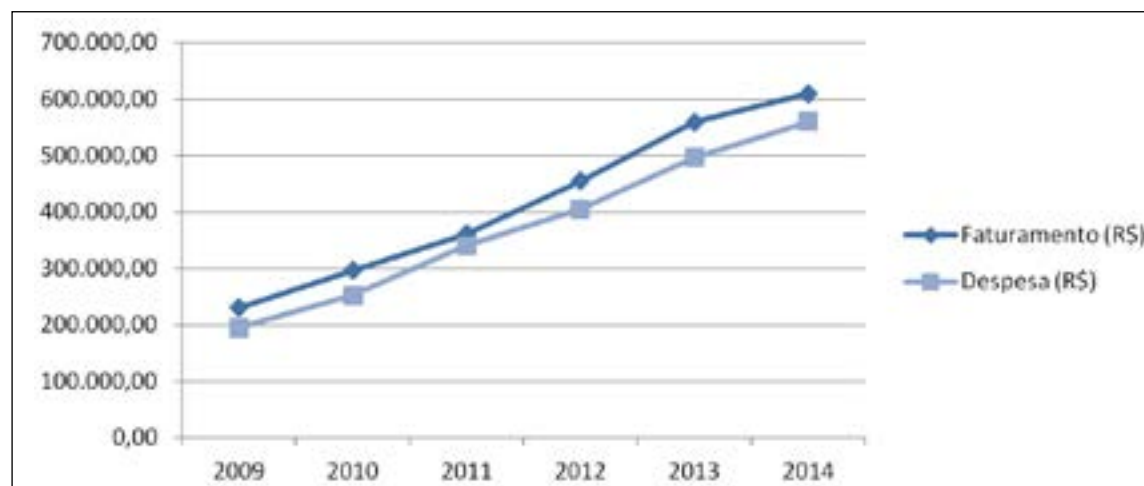
Fonte: Empresa de Lajeado.

Gráfico 3.4.a2 – Despesa (R\$) de uma empresa de Lajeado – 2009/2014



Fonte: Empresa de Lajeado.

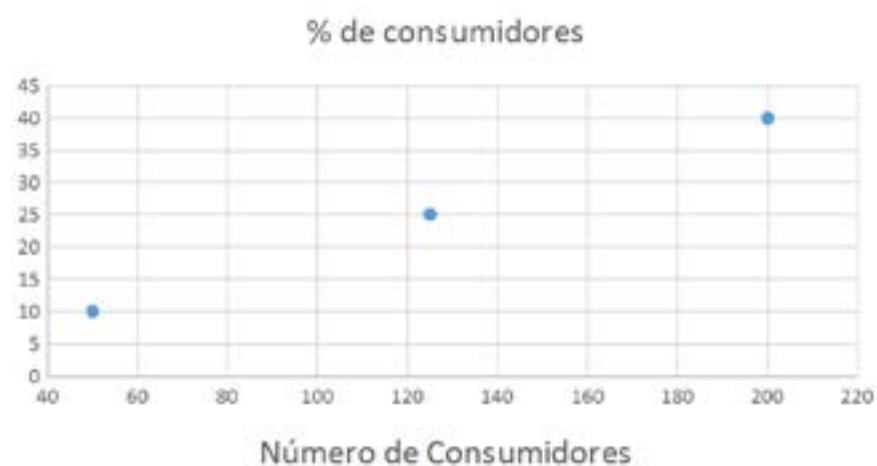
## b) Gráfico 3.4.b – Faturamento (R\$) e despesa (R\$) de uma empresa de Lajeado – 2009/2014



Fonte: Empresa de Lajeado.

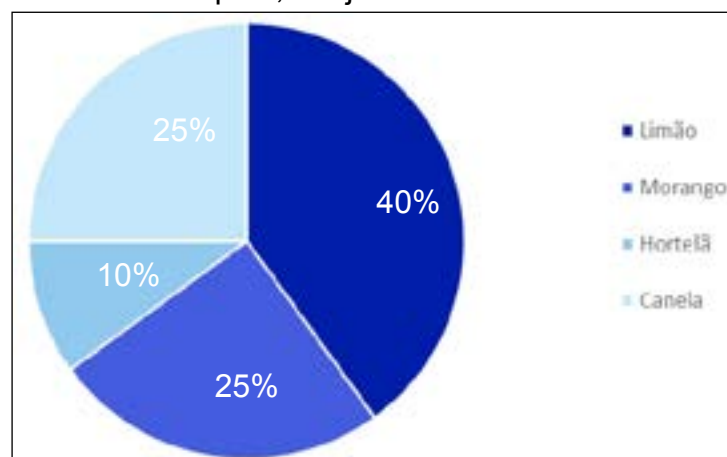
## 3.5

## a) Gráfico 3.5.a – Preferências por sabor de goma de mascar dos consumidores num município do Vale do Taquari, em janeiro de 2014



Fonte: Pesquisa com os consumidores.

## b) Gráfico 3.5.b – Preferências por sabor de goma de mascar dos consumidores num município do Vale do Taquari, em janeiro de 2014



Fonte: Pesquisa com os consumidores.

**APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 4**

4.1

Grupo I) 14; 3; 1; 7; 4; 3; 29; 41; 16; 8.

Grupo II) 15; 2; 7; 4; 19; 4; 43; 31; 10.

a) Grupo I = 12,6 e Grupo II = 15,0

b) Grupo I = 7,5 e Grupo II = 13

c) Grupo I = 3 e Grupo II = 4

4.2

5,9

4.3

4,4

4.4

12,87%

4.5

a) 715,5

b) 708,8

c) 669,2

d) 602,6

4.6

Média: 40,3; Mediana: 40; Moda: 40

4.7

a) 86,7%

b) 31,7%

c) 9,8

d) 3,6

e) 10,1

**APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 5**

5.1

a) 9

b) 2

c) 2,828

5.2

1,5

5.3

0,18

5.4

a) média aritmética = 7,3; média harmônica: 6,8

b) 5

c) 1,7

d) 4,3

e) 0,28

5.5

o segundo grupo

5.6

a) 3,7

b) desvio-padrão: 2,072; variância: 4,293

c) 1,664

d) 0,557

5.7

a) 29,2

b) 4,111

c) 5,142

d) 6,9

**APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 6**

6.1

c

6.2

c

6.3

e

6.4

e

6.5

a) Curva assimétrica positiva

b) excluir essa questão

6.6

a) Curva simétrica ( $x=mod = Med=5,2$ )b) Curva Platicúrtica ( $k=0,290$ )**APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 7**

7.1

a)  $S = \{BB; BP; BM; PP, PB, PM, MM, MB, MP\}$ 

b) 11,11%

c) 55,56%

7.2

a

7.3

c

7.4

a) 70,56%

b) 70%

7.5

a) 4,54%

b) 54,55%

c) 40,91%

7.6

a) 0,0001%

b) 82,95%

c) 4,37%

7.7

a) 0,01%

b) 1,99%

c) 98,01%

d) 1,98%

e) 99,99%

7.8

a) 12,63%

b) 23,68%

c) 3,16%

d) 76,32%

**APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 8**

8.1

- a) 40,96%
- b) 59,87%
- c) 91,39%

8.2

- a) 0,16%
- b) 40,96%
- c) 81,92%

8.3

- a) 2,32%
- b) 97,69%
- c) 78,47%
- d) 99,85%

8.4

- a) 99,41%
- b) 59,51%
- c) 10,28%

8.5

- a) 39,15%
- b) 55,63%
- c) 0,75 alunos
- d) 0,87 alunos

8.6

- a) 12,46%
- b) 0,67%

8.7

22,31%

8.8

66,26%

8.9

- a) 0,000%
- b) 1,04%

**APÊNDICE IV – Respostas dos exercícios propostos – Capítulo 9**

9.1

- a) 34,13%
- b) 41,15%
- c) 96,51%
- d) 24,79%
- e) 8,38%
- f) 91,62%

9.2

- a) 46,02%
- b) 44,04%
- c) 53,28%

9.3

- a) 4,75%
- b) 20,33%
- c) 74,92%
- d) aproximadamente 31 luzes

9.4

- a) 1,43%
- b) 0,62%

9.5

- a) 97,72%
- b) 97,59%
- c) 135.600 km

9.6

- a) 2,28%
- b) 9,18%
- c) 6,92
- d) Porque, conforme a Figura 5.4, mesmo utilizando-se três desvios padrão que representam 99,7% dos dados, a nota 9,8 estaria fora do intervalo de 0,5 e 9,5 (média 5 + 4,5 de desvio-padrão).

9.7

- a) 98,78%
- b) 0,62%

9.8

- a) 53,28%
- b) 20,29%
- c) 22,66%
- d) 15,87%
- e) 17,4 gramas
- f) 21,6 gramas

9.9

- a) 30 lâmpadas
- b) 3.446 lâmpadas

9.10

4,01%

9.11

- a) 65,54%
- b) 2,28%
- c) 34,46%
- d) aproximadamente 34 pessoas

9.12

82,52

9.13

- a) 65,46%
- b) 37,97%
- c) 40,88%
- d) aproximadamente 454 funcionários
- e) R\$ 870,50

9.14

- a) 4,73%
- b) 89,80%
- c) 35,07%
- d) 4,73%
- e) 79,60%

9.15

- a) 0,62%
- b) 78,88%
- c) Média de R\$ 1.000,00 e desvio-padrão de R\$ 80,00

9.16

- a) 1,66%
- b) 91,98%
- c) 64,06%
- d) 49,22%

9.17

- a) 90,53%
- b) 90,01%

9.18

- a) 8,52%
- b) 23,40%
- c) 85,12%
- d) 31,92%
- e) 58,52%



**UNIVATES**

R. Avelino Tallini, 171 | Bairro Universitário | Lajeado | RS | Brasil  
CEP 95900.000 | Cx. Postal 155 | Fone: (51) 3714.7000  
[www.univates.br](http://www.univates.br) | 0800 7 07 08 09