ANTÔNIO ARNOT CRESPO

19ª edição
Atualizada

ESTATÍSTICA FÁCIL

Editora Saraiva

A *Série Fácil*, criada com o objetivo de facilitar o aprendizado da Contabilidade, tem como principal característica a linguagem clara e acessível.

Os assuntos são tratados sempre de forma gradual, no momento adequado e seguindo uma seqüência lógica, partindo de situações mais fáceis para as menos fáceis, permitindo ao estudante familiarizar-se com a matéria de maneira natural e intuitiva.

O programa desenvolvido pela Série atende ao conteúdo programático dos cursos de nível técnico e de nível superior de Contabilidade, além de servir como instrumento de consulta e orientação para todos os profissionais, tanto da área de Contabilidade quanto de outras áreas, inclusive para os que pretendem se preparar para concursos públicos.

Antônio Arnot Crespo é bacharel em Ciências Econômicas pela Faculdade de Ciências Econômicas de Andradina; licenciado em Matemática pela Faculdade de Filosofia, Ciências e Letras Rui Barbosa, de Andradina, e licenciado em Pedagogia pela Faculdade de Educação, Ciências e Letras Urubupungá, de Pereira Barreto.

É professor efetivo de Matemática, por concurso público, da rede pública de ensino do Estado de São Paulo.

ESTATÍSTICA FÁCIL

Antônio Arnot Crespo

19ª edição
Atualizada

Editora Saraiva

Editora Saraiva

Rua Henrique Schaumann, 270
Pinheiros – São Paulo – SP – CEP: 05413-010
PABX (11) 3613-3000

SAC | **0800-0117875**
De 2ª a 6ª, das 8h30 às 19h30
www.editorasaraiva.com.br/contato

Diretora editorial	Flávia Alves Bravin
Gerente editorial	Rogério Eduardo Alves
Planejamento editorial	Rita de Cássia S. Puoço
Editoras	Luciana Cruz Patricia Quero
Produtoras editoriais	Daniela Nogueira Secondo Rosana Peroni Fazolari
Comunicação e produção digital	Nathalia Setrini Luiz
Suporte editorial	Najla Cruz Silva
Arte, produção e capa	Casa de ideias
Produção gráfica	Liliane Cristina Gomes
Atualização da 10ª tiragem	ERJ Composição Editorial
Impressão e acabamento	*Bartira Gráfica e Editora LTDA.*

ISBN 978-85-02-08106-2

CIP-BRASIL. CATALOGAÇÃO NA FONTE
SINDICATO NACIONAL DOS EDITORES DE LIVROS, RJ.

C94e

19.ed.
Crespo, Antônio Arnot
Estatística fácil / Antônio Arnot Crespo. – 19.ed. atual. – São Paulo : Saraiva, 2009.

Anexos
Contém questões e respectivas respostas
ISBN 978-85-02-08106-2

1. Estatística. I. Título.

09-0539 CDD-519.5
CDU-519.2

19ª edição
1ª tiragem: 2009
2ª tiragem: 2009
3ª tiragem: 2010
4ª tiragem: 2010
5ª tiragem: 2011
6ª tiragem: 2011
7ª tiragem: 2012
8ª tiragem: 2013
9ª tiragem: 2013
10ª tiragem: 2014

360.803.019.010

APRESENTAÇÃO

Este livro é o resultado de vários anos de estudo dirigidos ao ensino de **Estatística** e destina-se aos estudantes de cursos técnicos (Secretariado, Contabilidade, Administração, etc.) e, também, aos alunos de cursos superiores que necessitam de um estudo **introdutório de Estatística**.

Preocupou-nos apresentar todos os tópicos exigidos pelo programa estabelecido para os cursos profissionalizantes da rede de ensino particular e oficial, de uma forma acessível ao aluno, dentro de um esquema de ensino objetivo e prático.

Por essa razão, as características deste livro são eminentemente didáticas. Foram evitadas demonstrações, sendo apresentados comentários e análises objetivas dos assuntos. O estudo é complementado por exercícios em abundância, nos quais procuramos trabalhar com situações práticas.

Após ampla **reformulação**, que promoveu a atualização do texto e a inclusão e redistribuição de alguns assuntos, a estrutura da obra ficou assim:

- Nos oito primeiros capítulos, desenvolvemos os tópicos de **Estatística Descritiva**, dando um especial destaque à Distribuição de Frequência.
- No Capítulo 9, enfocamos o estudo de **Probabilidades**, de forma elementar, enfatizando o uso do raciocínio. No Capítulo 10, entreabrimos a porta para um

primeiro contato com os dois principais modelos teóricos de **Distribuição de Probabilidade**: Distribuição Binomial e Distribuição Normal.

- No Capítulo 11, apresentamos um estudo elementar de **Correlação e Regressão**, que nos ajudará a compreender e medir a relação entre variáveis. Os **Números-índices**, de interesse permanente no aspecto econômico de nosso dia a dia, passaram por uma revisão, na qual procuramos dar ênfase à realidade prática de sua formação e de seu emprego (Capítulo 12).
- Finalmente, o **Apêndice — Instrumental Matemático**, a ser consultado de acordo com as necessidades de cada aluno, foi complementado.

Os exercícios, sempre colocados em pontos estratégicos de cada capítulo, estão divididos em três seções:

- **Exercícios resolvidos** — exemplos para a fixação da matéria estudada;
- **Resolva** — exercícios de aprendizagem imediata, algumas vezes com o raciocínio já encaminhado;
- **Exercícios** — sequência graduada de exercícios propostos.

No final do livro, apresentamos uma **Coletânea de Questões Objetivas**, que poderão ser usadas nas verificações de aprendizagem.

Todos os exercícios deverão ser resolvidos num caderno à parte. As respostas estão no final do livro.

Consideramos a Matemática, a Música e a Estatística linguagens universais; lembramos que, "embora uma nova linguagem pareça um enigma antes de ser conquistada, é um poder, em seguida". Nosso desejo é que aqueles que fizerem uso deste livro conquistem a linguagem estatística, utilizando-a proveitosamente.

Aproveitamos para agradecer a todos aqueles que confiaram em nosso trabalho, utilizando este livro, e, em especial, àqueles que, fazendo suas críticas, deram-nos a oportunidade de melhorá-lo.

Continuamos a acolher os pareceres e sugestões para o aperfeiçoamento deste trabalho.

O autor

SUMÁRIO

CAPÍTULO 3 – SÉRIES ESTATÍSTICAS

CAPÍTULO 4 – GRÁFICOS ESTATÍSTICOS

CAPÍTULO 5 – DISTRIBUIÇÃO DE FREQUÊNCIA

1 A NATUREZA DA ESTATÍSTICA

1.1 Panorama histórico

Todas as ciências têm suas raízes na história do homem.

A Matemática, que é considerada "a ciência que une à clareza do raciocínio a síntese da linguagem", originou-se do convívio social, das trocas, da contagem, com caráter prático, utilitário, empírico.

A Estatística, ramo da Matemática Aplicada, teve origem semelhante.

Desde a Antiguidade, vários povos já registravam o número de habitantes, de nascimentos, de óbitos, faziam estimativas das riquezas individual e social, distribuíam equitativamente terras ao povo, cobravam impostos e realizavam inquéritos quantitativos por processos que, hoje, chamaríamos de "estatísticas".

Na Idade Média colhiam-se informações, geralmente com finalidades tributárias ou bélicas.

A partir do século XVI começaram a surgir as primeiras análises sistemáticas de fatos sociais, como batizados, casamentos, funerais, originando as primeiras tábuas e tabelas e os primeiros números relativos.

No século XVIII o estudo de tais fatos foi adquirindo, aos poucos, feição verdadeiramente científica. **Godofredo Achenwall** batizou a nova ciência (ou método) com o nome de **Estatística**, determinando o seu objetivo e suas relações com as ciências.

As tabelas tornaram-se mais completas, surgiram as representações gráficas e o cálculo das probabilidades, e a Estatística deixou de ser simples catalogação de dados numéricos coletivos para se tornar o estudo de **como chegar a conclusões sobre o todo** (população[1]), **partindo da observação de partes desse todo** (amostras[1]).

Atualmente, o público leigo (leitor de jornais e revistas) posiciona-se em dois extremos divergentes e igualmente errôneos quanto à validade das conclusões estatísticas: ou crê em sua infalibilidade ou afirma que elas nada provam. Os que assim pensam ignoram os objetivos, o campo e o rigor do método estatístico; ignoram a Estatística, quer teórica quer prática, ou a conhecem muito superficialmente.

Na era da energia nuclear, os estudos estatísticos têm avançado rapidamente e, com seus processos e técnicas, têm contribuído para a organização dos negócios e recursos do mundo moderno.

1.2 Método estatístico

1.2.1 O método científico

Muitos dos conhecimentos que temos foram obtidos na Antiguidade por acaso e, outros, por necessidades práticas, sem aplicação de um método.

Atualmente, quase todo acréscimo de conhecimento resulta da observação e do estudo. Se bem que muito desse conhecimento possa ter sido observado inicialmente por acaso, a verdade é que desenvolvemos processos científicos para seu estudo e para adquirirmos tais conhecimentos.

Podemos dizer, então, que:

> **Método** é um conjunto de meios dispostos convenientemente para se chegar a um fim que se deseja.

Dos métodos científicos, vamos destacar o método **experimental** e o **estatístico**.

1.2.2 O método experimental

> O **método experimental** consiste em manter constantes todas as causas (fatores), menos uma, e variar esta causa de modo que o pesquisador possa descobrir seus efeitos, caso existam.

É o método preferido no estudo da Física, da Química etc.

---

[1] Capítulo 2.

1.2.3 O método estatístico

Muitas vezes temos necessidade de descobrir fatos em um campo em que o método experimental não se aplica (nas ciências sociais), já que os vários fatores que afetam o fenômeno em estudo não podem permanecer constantes enquanto fazemos variar a causa que, naquele momento, nos interessa.

Como exemplo, podemos citar a determinação das causas que definem o preço de uma mercadoria. Para aplicarmos o método experimental, teríamos de fazer variar a quantidade da mercadoria e verificar se tal fato iria influenciar seu preço.

Contudo, seria necessário que não houvesse alteração nos outros fatores. Assim, deveria existir, no momento da pesquisa, uma uniformidade dos salários, o gosto dos consumidores deveria permanecer constante, seria necessária a fixação do nível geral dos preços das outras necessidades etc. Mas isso tudo é impossível.

Nesses casos, lançamos mão de outro método, embora mais difícil e menos preciso, denominado método estatístico.

O **método estatístico**, diante da impossibilidade de manter as causas constantes, admite todas essas causas presentes variando-as, registrando essas variações e procurando determinar, no resultado final, que influências cabem a cada uma delas.

1.3 A Estatística

Exprimindo por meio de números as observações que se fazem de elementos com, pelo menos, uma característica comum (por exemplo: os alunos do sexo masculino de uma comunidade), obtemos os chamados **dados** referentes a esses elementos.

Podemos dizer, então, que:

A **Estatística** é uma parte da Matemática Aplicada que fornece métodos para a coleta, organização, descrição, análise e interpretação de dados e para a utilização dos mesmos na tomada de decisões.

A **coleta**, a **organização** e a **descrição dos dados** estão a cargo da **Estatística Descritiva**, enquanto a **análise** e a **interpretação** desses dados ficam a cargo da **Estatística Indutiva** ou **Inferencial**.

Em geral, as pessoas, quando se referem ao termo estatística, o fazem no sentido da organização e descrição dos dados (estatística do Ministério da Educação, estatística dos acidentes de tráfego etc.), desconhecendo que **o aspecto essencial da Estatística é o de proporcionar métodos inferenciais, que permitam conclusões que transcendam os dados obtidos inicialmente.**

Assim, a análise e a interpretação dos dados estatísticos tornam possível o diagnóstico de uma empresa (por exemplo, de uma escola), o conhecimento de seus problemas (condições de funcionamento, produtividade), a formulação de soluções apropriadas e um planejamento objetivo de ação.

1.4 Fases do método estatístico

Podemos distinguir no método estatístico as seguintes fases:

1.4.1 Coleta de dados

Após cuidadoso planejamento e a devida determinação das características mensuráveis do fenômeno coletivamente típico[2] que se quer pesquisar, damos início à **coleta dos dados** numéricos necessários à sua descrição.

A coleta pode ser direta e indireta.

A coleta é **direta** quando feita sobre elementos informativos de registro obrigatório (nascimentos, casamentos e óbitos, importação e exportação de mercadorias), elementos pertinentes aos prontuários dos alunos de uma escola ou, ainda, quando os dados são coletados pelo próprio pesquisador através de inquéritos e questionários, como é o caso das notas de verificação e de exames, do censo demográfico etc.

A coleta direta de dados pode ser classificada relativamente ao fator tempo em:

a. **contínua** (registro) — quando feita continuamente, tal como a de nascimentos e óbitos e a de frequência dos alunos às aulas;
b. **periódica** — quando feita em intervalos constantes de tempo, como os censos (de 10 em 10 anos) e as avaliações mensais dos alunos;
c. **ocasional** — quando feita extemporaneamente, a fim de atender a uma conjuntura ou a uma emergência, como no caso de epidemias que assolam ou dizimam rebanhos inteiros.

---

[2] **Fenômeno coletivamente típico** é aquele que não apresenta regularidade na observação de casos isolados, mas na massa de observações. (ROCHA, Marcos Vinícius da. *Curso de Estatística*. 3. ed. Rio de Janeiro, Fundação IBGE, 1975.)

A coleta se diz **indireta** quando é inferida de elementos conhecidos (coleta direta) e/ou do conhecimento de outros fenômenos relacionados com o fenômeno estudado. Como exemplo, podemos citar a pesquisa sobre a mortalidade infantil, que é feita através de dados colhidos por uma coleta direta.

1.4.2 Crítica dos dados

Obtidos os dados, eles devem ser cuidadosamente criticados, à procura de possíveis falhas e imperfeições, a fim de não incorrermos em erros grosseiros ou de certo vulto, que possam influir sensivelmente nos resultados.

A crítica é **externa** quando visa às causas dos erros por parte do informante, por distração ou má interpretação das perguntas que lhe foram feitas; é **interna** quando visa a observar os elementos originais dos dados da coleta.

1.4.3 Apuração dos dados

Nada mais é do que a soma e o processamento dos dados obtidos e a disposição mediante critérios de classificação. Pode ser **manual**, **eletromecânica** ou **eletrônica**.

1.4.4 Exposição ou apresentação dos dados

Por mais diversa que seja a finalidade que se tenha em vista, os dados devem ser apresentados sob forma adequada (**tabelas** ou **gráficos**[3]), tornando mais fácil o exame daquilo que está sendo objeto de tratamento estatístico e ulterior obtenção de medidas típicas[4].

1.4.5 Análise dos resultados

Como já dissemos, o objetivo último da Estatística é tirar conclusões sobre o todo (população) a partir de informações fornecidas por parte representativa do todo (amostra). Assim, realizadas as fases anteriores (**Estatística Descritiva**), fazemos uma análise dos resultados obtidos, através dos métodos da **Estatística Indutiva** ou **Inferencial**, que tem por base a indução ou inferência, e tiramos desses resultados conclusões e previsões.

1.5 A Estatística nas empresas

No mundo atual, a **empresa** é uma das vigas mestras da Economia dos povos.

---

[3] Capítulo 3 e 4.

[4] Capítulo 6.

A direção de uma empresa, de qualquer tipo, incluindo as estatais e governamentais, exige de seu administrador a importante tarefa de **tomar decisões**, e o conhecimento e o uso da Estatística facilitarão seu tríplice trabalho de organizar, dirigir e controlar a empresa.

Por meio de **sondagem**, de **coleta de dados** e de **recenseamento de opiniões**, podemos conhecer a realidade geográfica e social, os recursos naturais, humanos e financeiros disponíveis, as expectativas da comunidade sobre a empresa, e estabelecer suas **metas**, seus objetivos com maior possibilidade de serem alcançados a curto, médio ou longo prazos.

A Estatística ajudará em tal trabalho, como também na seleção e organização da estratégia a ser adotada no empreendimento e, ainda, na escolha das técnicas de **verificação** e **avaliação** da quantidade e da qualidade do produto e mesmo dos possíveis lucros e/ou perdas.

Tudo isso que se pensou, que se planejou, precisa ficar registrado, documentado para evitar esquecimentos, a fim de garantir o bom uso do tempo, da energia e do material e, ainda, para um controle eficiente do trabalho.

O esquema do planejamento é o **plano**, que pode ser resumido, com auxílio da Estatística, em **tabelas** e **gráficos**, que facilitarão a compreensão visual dos cálculos matemático-estatísticos que lhes deram origem.

O homem de hoje, em suas múltiplas atividades, lança mão de processos e técnicas estatísticos, e só estudando-os evitaremos o erro das generalizações apressadas a respeito de tabelas e gráficos apresentados em jornais, revistas e televisão, frequentemente cometido quando se conhece apenas "por cima" um pouco de Estatística.

Exercícios

1. Complete:
   O método experimental é o mais usado por ciências como: ______________________

2. As ciências humanas e sociais, para obter os dados que buscam, lançam mão de que método?

3. O que é Estatística?

4. Cite as fases do método estatístico.

5. Para você, o que é coletar dados?

6. Para que serve a crítica dos dados?

7. O que é apurar dados?

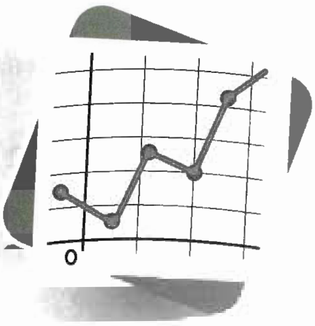
8. Como podem ser apresentados ou expostos os dados?

9. As conclusões, as inferências pertencem a que parte da Estatística?

10. Cite três ou mais atividades do planejamento empresarial em que a Estatística se faz necessária.

11. O método estatístico tem como um de seus fins:

    a. estudar os fenômenos estatísticos.
    b. estudar qualidades concretas dos indivíduos que formam grupos.
    c. determinar qualidades abstratas dos indivíduos que formam grupos.
    d. determinar qualidades abstratas de grupos de indivíduos.
    e. estudar fenômenos numéricos.



2 POPULAÇÃO E AMOSTRA[1]

2.1 Variáveis

A cada fenômeno corresponde um número de resultados possíveis. Assim, por exemplo:

- para o fenômeno "sexo" são dois os resultados possíveis: sexo masculino e sexo feminino;
- para o fenômeno "número de filhos" há um número de resultados possíveis expresso através dos números naturais:

  0, 1, 2, 3, ..., n;
- para o fenômeno "estatura" temos uma situação diferente, pois os resultados podem tomar um número infinito de valores numéricos dentro de um determinado intervalo.

Variável é, convencionalmente, o conjunto de resultados possíveis de um fenômeno.

[1] Consulte o **Apêndice — Instrumental Matemático**, para uma revisão dos assuntos **Arredondamento de Dados** (p. 172) e **Compensação** (p. 173).

Os exemplos acima nos dizem que uma variável pode ser:

a. **qualitativa** — quando seus valores são expressos por atributos: sexo (masculino — feminino), cor da pele (branca, preta, amarela, vermelha, parda) etc.;

b. **quantitativa** — quando seus valores são expressos em números (salários dos operários, idade dos alunos de uma escola etc.). Uma variável quantitativa que pode assumir, teoricamente, qualquer valor entre dois limites recebe o nome de **variável contínua**; uma variável que só pode assumir valores pertencentes a um conjunto enumerável recebe o nome de **variável discreta**.

Assim, o número de alunos de uma escola pode assumir qualquer um dos valores do conjunto $\mathbb{N} = \{1, 2, 3, ..., 58, ...\}$, mas nunca valores como 2,5 ou 3,78 ou 4,325 etc. Logo, é uma **variável discreta**. Já o peso desses alunos é uma **variável contínua**, pois um dos alunos tanto pode pesar 72 kg, como 72,5 kg, como 72,54 kg etc., dependendo esse valor da precisão da medida.

De modo geral, as **medições** dão origem a variáveis contínuas e as **contagens** ou **enumerações**, a variáveis discretas.

Designamos as variáveis por letras latinas, em geral, as últimas:

$$\mathbf{x, y, z.}$$

Por exemplo, sejam 2, 3, 5 e 8 todos os resultados possíveis de um dado fenômeno. Fazendo uso da letra **x** para indicar a variável relativa ao fenômeno considerado, temos:

$$x \in \{2, 3, 5, 8\}$$

Resolva

1. Classifique as variáveis em qualitativas ou quantitativas (contínuas ou descontínuas):
   a. Universo: alunos de uma escola.
   Variável: cor dos cabelos — ....
   b. Universo: casais residentes em uma cidade.
   Variável: número de filhos — ....
   c. Universo: as jogadas de um dado.
   Variável: o ponto obtido em cada jogada — ....
   d. Universo: peças produzidas por certa máquina.
   Variável: número de peças produzidas por hora — ....
   e. Universo: peças produzidas por certa máquina.
   Variável: diâmetro externo — ....

Exercício

1. Diga quais das variáveis abaixo são discretas e quais são contínuas:
   - a. População: alunos de uma cidade.
     Variável: cor dos olhos.
   - b. P.: estação meteorológica de uma cidade.
     V.: precipitação pluviométrica, durante um ano.
   - c. P.: Bolsa de Valores de São Paulo.
     V.: número de ações negociadas.
   - d. P.: funcionários de uma empresa.
     V.: salários.
   - e. P.: pregos produzidos por uma máquina.
     V.: comprimento.
   - f. P.: casais residentes em uma cidade.
     V.: sexo dos filhos.
   - g. P.: propriedades agrícolas do Brasil.
     V.: produção de algodão.
   - h. P.: segmentos de reta.
     V.: comprimento.
   - i. P.: bibliotecas da cidade de São Paulo.
     V.: número de volumes.
   - j. P.: aparelhos produzidos em uma linha de montagem.
     V.: número de defeitos por unidade.
   - l. P.: indústrias de uma cidade.
     V.: índice de liquidez.

2.2 População e amostra

Ao conjunto de entes portadores de, pelo menos, uma característica comum denominamos **população estatística** ou **universo estatístico.**

Assim, os estudantes, por exemplo, constituem uma população, pois apresentam pelo menos uma característica comum: são os que estudam.

Como em qualquer estudo estatístico, temos em mente pesquisar uma ou mais características dos elementos de alguma população, esta característica deve estar perfeitamente definida. E isto se dá quando, considerado um elemento qualquer, podemos afirmar, sem ambiguidade, se esse elemento pertence ou não à população. É necessário, pois, existir um critério de constituição da população, válido para qualquer pessoa, no tempo ou no espaço.

Por isso, quando pretendemos fazer uma pesquisa entre os alunos das escolas de 1º grau, precisamos definir quais são os alunos que formam o universo: os que atualmente ocupam as carteiras das escolas, ou devemos incluir também os que já passaram pela escola? É claro que a solução do problema vai depender de cada caso em particular.

Na maioria das vezes, por impossibilidade ou inviabilidade econômica ou temporal, limitamos as observações referentes a uma determinada pesquisa a apenas uma parte da população. A essa parte proveniente da população em estudo denominamos **amostra**.

Uma **amostra** é um subconjunto finito de uma população.

Como vimos no capítulo anterior, a Estatística Indutiva tem por objetivo tirar conclusões sobre as populações, com base em resultados verificados em amostras retiradas dessa população.

Mas, para as inferências serem corretas, é necessário garantir que a amostra seja **representativa** da população, isto é, a amostra deve possuir as mesmas características básicas da população, no que diz respeito ao fenômeno que desejamos pesquisar. É preciso, pois, que a amostra ou as amostras que vão ser usadas sejam obtidas por processos adequados.

Há casos, como o de pesquisas sociais, econômicas e de opinião, em que os problemas de amostragem são de extrema complexidade. Mas existem também casos em que os problemas de amostragem são bem mais fáceis. Como exemplo, podemos citar a retirada de amostras para controle de qualidade dos produtos ou materiais de determinada indústria.

2.3 Amostragem

Existe uma técnica especial — **amostragem** — para recolher amostras, que garante, tanto quanto possível, o acaso na escolha.

Dessa forma, cada elemento da população passa a ter a mesma chance de ser escolhido, o que garante à amostra o caráter de representatividade, e isto é muito importante, pois, como vimos, nossas conclusões relativas à população vão estar baseadas nos resultados obtidos nas amostras dessa população.

Daremos, a seguir, três das principais técnicas de amostragem.

2.3.1 Amostragem casual ou aleatória simples

Este tipo de amostragem é equivalente a um sorteio lotérico.

Na prática, a **amostragem casual** ou **aleatória simples** pode ser realizada numerando-se a população de **1** a **n** e sorteando-se, a seguir, por meio de um dispositivo aleatório qualquer, **k** números dessa sequência, os quais corresponderão aos elementos pertencentes à amostra.

Exemplo:

Vamos obter uma amostra representativa para a pesquisa da estatura de noventa alunos de uma escola:

a. Numeramos os alunos de 01 a 90.

b. Escrevemos os números, de 01 a 90, em pedaços iguais de um mesmo papel, colocando-os dentro de uma caixa. Agitamos sempre a caixa para misturar bem os pedaços de papel e retiramos, um a um, nove números que formarão a amostra. Neste caso, 10% da população.

Quando o número de elementos da amostra é grande, esse tipo de sorteio torna-se muito trabalhoso. A fim de facilitá-lo, foi elaborada uma tabela — **Tabela de Números Aleatórios** —, construída de modo que os dez algarismos (0 a 9) são distribuídos ao acaso nas linhas e colunas (Anexo I, p. 217).

Para obtermos os elementos da amostra usando a tabela, sorteamos um algarismo qualquer da mesma, a partir do qual iremos considerar números de dois, três ou mais algarismos, conforme nossa necessidade. Os números assim obtidos irão indicar os elementos da amostra.

A leitura da tabela pode ser feita horizontalmente (da direita para a esquerda ou vice-versa), verticalmente (de cima para baixo ou vice-versa), diagonalmente (no sentido ascendente ou descendente) ou formando o desenho de uma letra qualquer. A opção, porém, deve ser feita antes de iniciado o processo.

Assim, para o nosso exemplo, considerando a 18ª linha, tomamos os números de dois algarismos (tantos algarismos quantos formam o maior número da população), obtendo:

61 02 01 81 73 92 60 66 73 58 53 34

Evidentemente, o numeral 92 será desprezado, pois não consta da população, como será também abandonado um numeral que já tenha aparecido. Temos, então:

61 02 01 81 73 60 66 58 53

Medindo as alturas dos alunos correspondentes aos números sorteados, obteremos uma amostra das estaturas dos noventa alunos.

2.3.2 Amostragem proporcional estratificada

Muitas vezes a população se divide em subpopulações — **estratos**.

Como é provável que a variável em estudo apresente, de estrato em estrato, um comportamento heterogêneo e, dentro de cada estrato, um comportamento homogêneo, convém que o sorteio dos elementos da amostra leve em consideração tais estratos.

É exatamente isso que fazemos quando empregamos a **amostragem proporcional estratificada**, que, além de considerar a existência dos estratos, obtém os elementos da amostra proporcional ao número de elementos dos mesmos.

Exemplo:

Supondo, no exemplo anterior, que, dos noventa alunos, 54 sejam meninos e 36 sejam meninas, vamos obter a amostra proporcional estratificada.

São, portanto, dois estratos (sexo masculino e sexo feminino) e queremos uma amostra de 10% da população. Logo, temos:

a.

SEXO	POPULAÇÃO	10%	AMOSTRA
M	54	$\frac{10 \times 54}{100} = 5{,}4$	5
F	36	$\frac{10 \times 36}{100} = 3{,}6$	4
Total	90	$\frac{10 \times 90}{100} = 9{,}0$	9

b. Numeramos os alunos de 01 a 90, sendo que de 01 a 54 correspondem meninos e de 55 a 90, meninas. Tomando na Tabela de Números Aleatórios a primeira e a segunda colunas da esquerda, de cima para baixo, obtemos os seguintes números:

57 28 ~~92~~ 90 80 22 56 79 53 18 ~~53~~ 03 27 05 40

Temos, então:

28 22 53 18 03 — para os meninos;

57 90 80 56 — para as meninas.

Resolva

1. Pesquisa — peso dos colegas de sua classe (incluindo você).

   Amostra — correspondente a 30% da população.

   Sugestão — faça uso da caderneta de seu professor e da Tabela dos Números Aleatórios (5ª e 6ª colunas, de baixo para cima).

2. Pesquisa — estatura dos alunos das 1as séries de sua escola.

   Amostra — 15% da população.

   Sugestão — use a Tabela de Números Aleatórios (25ª linha, da esquerda para a direita).

SÉRIES	POPULAÇÃO	15%	AMOSTRA
A			
B			

3. Em uma escola existem 250 alunos, sendo 35 na 1ª série, 32 na 2ª, 30 na 3ª, 28 na 4ª, 35 na 5ª, 32 na 6ª, 31 na 7ª e 27 na 8ª. Obtenha uma amostra de 40 alunos e preencha o quadro da página seguinte.

   Como, neste caso, foi dado o número de elementos da amostra, devemos, então, calcular o número de elementos de cada estrato proporcionalmente ao número de elementos da amostra. Assim, para a 1ª série, temos:

$$\begin{array}{cc} 250 & 40 \\ 35 & x \end{array} \Rightarrow x = \frac{35 \times 40}{250} = 5{,}6 \Rightarrow x = 6$$

   Logo:

SÉRIES	POPULAÇÃO	CÁLCULO PROPORCIONAL	AMOSTRA
1ª	35	$\frac{35 \times 40}{250} = 5{,}6$	6
2ª	....	....	....
3ª	....	....	....
4ª	28	....	....
5ª	....	....	6
6ª	....	....	....
7ª	....	$\frac{31 \times 40}{250} = ...$	....
8ª	....	....	....
Total	250	—	40

2.3.3 Amostragem sistemática

Quando os elementos da população já se acham ordenados, não há necessidade de construir o sistema de referência. São exemplos os prontuários médicos de um hospital, os prédios de uma rua, as linhas de produção etc. Nestes casos, a seleção dos elementos que constituirão a amostra pode ser feita por um sistema imposto pelo pesquisador. A esse tipo de amostragem denominamos **sistemática**.

Assim, no caso de uma linha de produção, podemos, a cada dez itens produzidos, retirar um para pertencer a uma amostra da produção diária. Neste caso, estaríamos fixando o tamanho da amostra em 10% da população.

Exemplo:

Suponhamos uma rua contendo novecentos prédios, dos quais desejamos obter uma amostra formada de cinquenta prédios. Podemos, neste caso, usar o seguinte procedimento: como $\frac{900}{50} = 18$, escolhemos por sorteio casual um número de 1 a 18 (inclusive), o qual indicaria o primeiro elemento sorteado para a amostra; os demais elementos seriam periodicamente considerados de 18 em 18. Assim, se o número sorteado fosse o 4, tomaríamos, pelo lado direito da rua, o 4º prédio, o 22º, o 40º etc., até voltarmos ao início da rua, pelo lado esquerdo.

Exercícios

1. Uma escola de 1º grau abriga 124 alunos. Obtenha uma amostra representativa correspondendo a 15% da população.
   Sugestão: use a 8ª, 9ª e 10ª colunas, a partir da 1ª linha, da Tabela de Números Aleatórios (de cima para baixo).

2. Em uma escola há oitenta alunos. Obtenha uma amostra de doze alunos.
   Sugestão: decida, juntamente com a classe e seu professor, o uso da Tabela de Números Aleatórios.

3. Uma população é formada por 140 notas resultantes da aplicação de um teste de inteligência:

62	129	95	123	81	93	105	95	96	80	87	110	139	75
123	60	72	86	108	120	57	113	65	108	90	137	74	106
109	84	121	60	128	100	72	119	103	128	80	99	149	85
77	91	51	100	63	107	76	82	110	63	131	65	114	103
104	107	63	117	116	86	115	62	122	92	102	113	74	78
69	116	82	95	72	121	52	80	100	85	117	85	102	106
94	84	123	42	90	91	81	116	73	79	98	82	69	102
100	79	101	98	110	95	67	77	91	95	74	90	134	94
79	92	73	83	74	125	101	82	71	75	101	102	78	108
125	56	86	98	106	72	117	89	99	86	82	57	106	90

   Obtenha uma amostra formada de 26 elementos, tomando, inicialmente, a 1ª linha da esquerda para a direita.

4. O diretor de uma escola, na qual estão matriculados 280 meninos e 320 meninas, desejoso de conhecer as condições de vida extraescolar de seus alunos e não dispondo de tempo para entrevistar todas as famílias, resolveu fazer um levantamento, por amostragem, em 10% dessa clientela. Obtenha, para esse diretor, os elementos componentes da amostra.

ESCOLAS	Nº DE ESTUDANTES	
	MASCULINO	FEMININO
A	80	95
B	102	120
C	110	92
D	134	228
E	150	130
F	300	290
Total	876	955

5. Uma cidade **X** apresenta o seguinte quadro relativo às suas escolas de 1º grau:

   Obtenha uma amostra proporcional estratificada de 120 estudantes.

6. Uma população encontra-se dividida em três estratos, com tamanhos, respectivamente, $n_1 = 40$, $n_2 = 100$ e $n_3 = 60$. Sabendo que, ao ser realizada uma amostragem estratificada proporcional, nove elementos da amostra foram retirados do 3º estrato, determine o número total de elementos da amostra.

7. Mostre como seria possível retirar uma amostra de 32 elementos de uma população ordenada formada por 2.432 elementos.
Na ordenação geral, qual dos elementos abaixo seria escolhido para pertencer à amostra, sabendo-se que o elemento de ordem 1.420 a ela pertence?
1.648º, 290º, 725º, 2.025º, 1.120º.

3 SÉRIES ESTATÍSTICAS[1]

3.1 Tabelas

Um dos objetivos da Estatística é sintetizar os valores que uma ou mais variáveis podem assumir, para que tenhamos uma visão global da variação dessa ou dessas variáveis. E isso ela consegue, inicialmente, apresentando esses valores em **tabelas** e **gráficos**, que irão nos fornecer rápidas e seguras informações a respeito das variáveis em estudo, permitindo-nos determinações administrativas e pedagógicas mais coerentes e científicas.

Tabela é um quadro que resume um conjunto de observações.

Uma tabela compõe-se de:

a. **corpo** — conjunto de linhas e colunas que contém informações sobre a variável em estudo;
b. **cabeçalho** — parte superior da tabela que especifica o conteúdo das colunas;
c. **coluna indicadora** — parte da tabela que especifica o conteúdo das linhas;

[1] Consulte o **Apêndice — Instrumental Matemático**, para uma revisão dos assuntos **Frações** (p. 174), **Razões** (p. 182) e **Percentagem** (p. 183).

d. **linhas** — retas imaginárias que facilitam a leitura, no sentido horizontal, de dados que se inscrevem nos seus cruzamentos com as colunas;

e. **casa** ou **célula** — espaço destinado a um só número;

f. **título** — conjunto de informações, as mais completas possíveis, respondendo às perguntas: **O quê?, Quando?, Onde?,** localizado no topo da tabela.

Há ainda a considerar os elementos complementares da tabela, que são a **fonte**, as **notas** e as **chamadas**, colocados, de preferência, no seu rodapé.

Exemplo:

MÉDIA DE ANOS DE ESTUDO DAS ← TÍTULO
PESSOAS DE 10 ANOS OU MAIS DE IDADE
BRASIL — 2003-2007

CABEÇALHO — COLUNA INDICADORA — COLUNA NUMÉRICA — CORPO — CASA OU CÉLULA — LINHAS — RODAPÉ

ANOS	MÉDIA DE ANOS DE ESTUDO
2003	$\boxed{7{,}2}$
2004	7,3
2005	7,4
2006	7,7
2007	7,8

FONTE: IBGE.

De acordo com as normas da Fundação IBGE, nas **casas** ou **células** devemos colocar:

- um traço horizontal (—) quando o valor é zero, não só quanto à natureza das coisas, como quanto ao resultado do inquérito;
- três pontos (...) quando não temos os dados;
- um ponto de interrogação (?) quando temos dúvida quanto à exatidão de determinado valor;
- zero (0) quando o valor é muito pequeno para ser expresso pela unidade utilizada. Se os valores são expressos em numerais decimais, precisamos acrescentar à parte decimal um número correspondente de zeros (0,0; 0,00; 0,000; ...).

3.2 Séries estatísticas

Denominamos **série estatística** toda tabela que apresenta a distribuição de um conjunto de dados estatísticos em função da época, do local ou da espécie.

Daí, podemos inferir que numa série estatística observamos a existência de três elementos ou fatores: o **tempo**, o **espaço** e a **espécie**.

Conforme varie um dos elementos da série, podemos classificá-la em **histórica, geográfica** e **específica**.

3.2.1 Séries históricas, cronológicas, temporais ou marchas

Descrevem os valores da variável, em determinado local, discriminados segundo intervalos de tempo variáveis.

Exemplo:

FRANGO — PREÇOS MÉDIOS
EM SÃO PAULO — 2003-2008

ANOS	PREÇO MÉDIO (R$)
2003	2,56
2004	2,64
2005	2,67
2006	2,53
2007	3,20
2008	3,64

FONTE: Associação Paulista de Avicultura.

3.2.2 Séries geográficas, espaciais, territoriais ou de localização

Descrevem os valores da variável, em determinado instante, discriminados segundo regiões.

Exemplo:

DURAÇÃO MÉDIA DOS
ESTUDOS SUPERIORES
1994

PAÍSES	NÚMERO DE ANOS
Itália	7,5
Alemanha	7,0
França	7,0
Holanda	5,9
Inglaterra	Menos de 4

FONTE: Revista *Veja*.

3.2.3 Séries específicas ou categóricas

Descrevem os valores da variável, em determinado tempo e local, discriminados segundo especificações ou categorias.

Exemplo:

REBANHOS BRASILEIROS — EFETIVO
NOS ESTABELECIMENTOS
AGROPECUÁRIOS
2006

ESPÉCIES	QUANTIDADE
Bovinos	205.886.244
Bubalinos	1.156.870
Aves	821.541.630
Suínos	35.173.824
Ovinos	16.019.170
Caprinos	10.401.449

FONTE: IBGE.

3.3 Séries conjugadas
Tabela de dupla entrada

Muitas vezes temos necessidade de apresentar, em uma única tabela, a variação de valores de mais de uma variável, isto é, fazer uma **conjugação** de duas ou mais séries.

Conjugando duas séries em uma única tabela, obtemos uma **tabela de dupla entrada**. Em uma tabela desse tipo ficam criadas duas ordens de classificação: uma **horizontal** (linha) e uma **vertical** (coluna).

Exemplo:

TERMINAIS TELEFÔNICOS EM SERVIÇO
1991-93

REGIÕES	1991	1992	1993
Norte	342.938	375.658	403.494
Nordeste	1.287.813	1.379.101	1.486.649
Sudeste	6.234.501	6.729.467	7.231.634
Sul	1.497.315	1.608.989	1.746.232
Centro-Oeste	713.357	778.925	884.822

FONTE: Ministério das Comunicações.

A conjugação, no exemplo dado, foi **série geográfica-série histórica**, que dá origem à série geográfico-histórica ou geográfico-temporal.

Podem existir, se bem que mais raramente, pela dificuldade de representação, séries compostas de três ou mais entradas.

3.4 Distribuição de frequência

Por se tratar de um conceito estatístico de suma importância, merecerá no Capítulo 5 um tratamento especial.

Exemplo:

ESTATURAS DE 100 ALUNOS
DA ESCOLA X — 2008

ESTATURAS (cm)	Nº DE ALUNOS
140 ⊢ 145	2
145 ⊢ 150	5
150 ⊢ 155	11
155 ⊢ 160	39
160 ⊢ 165	32
165 ⊢ 170	10
170 ⊢ 175	1
Total	100

Dados fictícios.

Exercícios

1. Classifique as séries:

a.

PRODUÇÃO DE BORRACHA NATURAL
1991-93

ANOS	TONELADAS
1991	29.543
1992	30.712
1993	40.663

FONTE: IBGE.

b.

AVICULTURA BRASILEIRA
1992

ESPÉCIES	NÚMERO (1.000 cabeças)
Galinhas	204.160
Galos, frangos, frangas e pintos	435.465
Codornas	2.488

FONTE: IBGE.

c.

VACINAÇÃO CONTRA A
POLIOMIELITE — 1993

REGIÕES	QUANTIDADE
Norte	211.209
Nordeste	631.040
Sudeste	1.119.708
Sul	418.785
Centro-Oeste	185.823

FONTE: Ministério da Saúde.

d.

AQUECIMENTO DE UM MOTOR
DE AVIÃO DE MARCA X

MINUTOS	TEMPERATURA (°C)
0	20
1	27
2	34
3	41
4	49
5	56
6	63

Dados fictícios.

e.

PRODUÇÃO BRASILEIRA DE AÇO BRUTO
1991-93

PROCESSOS	QUANTIDADE (1.000 t)		
	1991	1992	1993
Oxigênio básico	17.934	18.849	19.698
Forno elétrico	4.274	4.637	5.065
EOF	409	448	444

FONTE: Instituto Brasileiro de Siderurgia.

f.

EXPORTAÇÃO BRASILEIRA
1985-1990-1995

IMPORTADORES	1985 %	1990 %	1995 %
América Latina	13,0	13,4	25,6
EUA e Canadá	28,2	26,3	22,2
Europa	33,9	35,2	20,7
Ásia e Oceania	10,9	17,7	15,4
África e Oriente Médio	14,0	8,8	5,5

FONTES: MIC e SECEX.

2. Procure exemplos de séries estatísticas em jornais e revistas e copie-os, classificando essas séries.

3. Pesquise, junto à secretaria de sua escola, os dados necessários ao preenchimento da tabela abaixo:

MATRÍCULAS NA ESCOLA...
EM 19...

SÉRIES	SEXO	
	MASCULINO	FEMININO

4. Verificou-se, em 1993, o seguinte movimento de importação de mercadorias: 14.839.804 t, oriundas da Arábia Saudita, no valor de US$ 1.469.104.000; 10.547.889 t, dos Estados Unidos, no valor de US$ 6.034.946.000; e 561.024 t, do Japão, no valor de US$ 1.518.843.000. Confeccione a série correspondente e classifique-a, sabendo que os dados acima foram fornecidos pelo Ministério da Fazenda.

3.5 Dados absolutos e dados relativos

Os dados estatísticos resultantes da coleta direta da fonte, sem outra manipulação senão a contagem ou medida, são chamados **dados absolutos**.

A leitura dos dados absolutos é sempre enfadonha e inexpressiva; embora esses dados traduzam um resultado exato e fiel, não têm a virtude de ressaltar de imediato as suas conclusões numéricas. Daí o uso imprescindível que faz a Estatística dos dados relativos.

Dados relativos são o resultado de comparações por quociente (razões) que se estabelecem entre dados absolutos e têm por finalidade realçar ou facilitar as comparações entre quantidades.

Traduzem-se os dados relativos, em geral, por meio de **percentagens, índices, coeficientes** e **taxas**.

3.5.1 As percentagens

Consideremos a série:

MATRÍCULAS NAS ESCOLAS
DA CIDADE A — 2008

CATEGORIAS	NÚMERO DE ALUNOS
Ensino Fundamental	19.286
Ensino Médio	1.681
Ensino Superior	234
Total	21.201

Dados fictícios.

Calculemos as percentagens dos alunos de cada nível de ensino:

$$\text{Ensino Fundamental} \rightarrow \frac{19.286 \times 100}{21.201} = 90{,}96 = 91{,}0$$

$$\text{Ensino Médio} \rightarrow \frac{1.681 \times 100}{21.201} = 7{,}92 = 7{,}9$$

$$\text{Ensino Superior} \rightarrow \frac{234 \times 100}{21.201} = 1{,}10 = 1{,}1$$

Com esses dados, podemos formar uma nova coluna na série em estudo:

MATRÍCULAS NAS ESCOLAS
DA CIDADE A — 2008

CATEGORIAS	Nº DE ALUNOS	%
Ensino Fundamental	19.286	91,0
Ensino Médio	1.681	7,9
Ensino Superior	234	1,1
Total	21.201	100,0

Dados fictícios.

Os valores dessa nova coluna nos dizem que, de cada 100 alunos da cidade **A**, 91 estão matriculados no Ensino Fundamental, 8, aproximadamente, no Ensino Médio e 1 no Ensino Superior.

O emprego da percentagem é de grande valia quando é nosso intuito destacar a participação da parte no todo.

Consideremos, agora, a série:

MATRÍCULAS NAS ESCOLAS
DAS CIDADES A E B — 2008

CATEGORIAS	Nº DE ALUNOS	
	CIDADE A	CIDADE B
Ensino Fundamental	19.286	38.660
Ensino Médio	1.681	3.399
Ensino Superior	234	424
Total	21.201	42.483

Dados fictícios.

Qual das cidades tem, comparativamente, maior número de alunos em cada nível de ensino?

Como o número total de alunos é diferente nas duas cidades, não é fácil concluir a respeito usando os dados absolutos. No entanto, usando as percentagens, tal tarefa fica bastante facilitada. Assim, acrescentando na tabela anterior as colunas correspondentes às percentagens, obtemos:

MATRÍCULAS NAS ESCOLAS
DAS CIDADES A E B — 2008

CATEGORIAS	CIDADE A		CIDADE B	
	Nº DE ALUNOS	%	Nº DE ALUNOS	%
Ensino Fundamental	19.286	91,0	38.660	91,0
Ensino Médio	1.681	7,9	3.399	8,0
Ensino Superior	234	1,1	424	1,0
Total	21.201	100,0	42.483	100,0

o que nos permite dizer que, comparativamente, contam, praticamente, com o mesmo número de alunos em cada nível de ensino.

NOTAS:

- Do mesmo modo que tomamos **100** para base de comparação, também podemos tomar outro número qualquer, entre os quais destacamos o número **1**. É claro que, supondo o total igual a **1**, os dados relativos das parcelas serão todos menores que **1**.
- Em geral, quando usamos **100** para base, os dados são arredondados até a primeira casa decimal; e quando tomamos **1** por base, são arredondados até a terceira casa decimal.

Resolva

1. Complete a tabela abaixo:

ESCOLAS	Nº DE ALUNOS	DADOS RELATIVOS	
		POR 1	POR 100
A	175	0,098	9,8
B	222	....	....
C	202	....	....
D	362	....	....
E	280	....	....
F	540	....	....
Total	1.781	1,000	100,0

Cálculos:

$$A \rightarrow \frac{175 \times 1}{1.781} = 0{,}098$$

3.5.2 Os índices. Índices econômicos

Os **índices** são razões entre duas grandezas tais que uma não inclui a outra.

São exemplos de índices:

$$\textbf{Índice cefálico} = \frac{\text{diâmetro transverso do crânio}}{\text{diâmetro longitudinal do crânio}} \times 100$$

$$\textbf{Quociente intelectual} = \frac{\text{idade mental}}{\text{idade cronológica}} \times 100$$

$$\textbf{Densidade demográfica} = \frac{\text{população}}{\text{superfície}}$$

Índices econômicos:

$$\text{Produção } \textit{per capita} = \frac{\text{valor total da produção}}{\text{população}}$$

$$\text{Consumo } \textit{per capita} = \frac{\text{consumo do bem}}{\text{população}}$$

$$\text{Renda } \textit{per capita} = \frac{\text{renda}}{\text{população}}$$

$$\text{Receita } \textit{per capita} = \frac{\text{receita}}{\text{população}}$$

3.5.3 Os coeficientes

Os **coeficientes** são razões entre o número de ocorrências e o número total (número de ocorrências e número de não ocorrências).

São exemplos de coeficientes:

$$\textbf{Coeficiente de natalidade} = \frac{\text{número de nascimentos}}{\text{população total}}$$

$$\textbf{Coeficiente de mortalidade} = \frac{\text{número de óbitos}}{\text{população total}}$$

Coeficientes educacionais:

$$\text{Coeficiente de evasão escolar} = \frac{\text{número de alunos evadidos}}{\text{número inicial de matrículas}}$$

$$\text{Coeficiente de aproveitamento escolar} = \frac{\text{número de alunos aprovados}}{\text{número final de matrículas}}$$

$$\text{Coeficiente de recuperação escolar} = \frac{\text{número de alunos recuperados}}{\text{número de alunos em recuperação}}$$

3.5.4 As taxas

As **taxas** são os coeficientes multiplicados por uma potência de 10 (10, 100, 1.000 etc.) para tornar o resultado mais inteligível.

São exemplos de taxas:

Taxa de mortalidade = coeficiente de mortalidade × 1.000

Taxa de natalidade = coeficiente de natalidade × 1.000

Taxa de evasão escolar = coeficiente de evasão escolar × 100

Exercício resolvido

1. O Estado **A** apresentou 733.986 matrículas na 1ª série, no início do ano de 1994, e 683.816 no fim do ano. O Estado **B** apresentou, respectivamente, 436.127 e 412.457 matrículas. Qual o Estado que apresentou maior evasão escolar?

$$A \longrightarrow TEE = \frac{733.986 - 683.816}{733.986} \times 100 =$$
$$0{,}0683 \times 100 = 6{,}83 = 6{,}8\%$$

$$B \longrightarrow TEE = \frac{436.127 - 412.457}{436.127} \times 100 =$$
$$0{,}0542 \times 100 = 5{,}42 = 5{,}4\%$$

O Estado que apresentou maior evasão escolar foi **A**.

Resolva

1. Uma escola registrou em março, na 1ª série, a matrícula de 40 alunos e a matrícula efetiva, em dezembro, de 35 alunos. A taxa de evasão foi de:

$$TEE = \frac{\text{nº de evadidos}}{\text{nº matrícula inicial}} \times 100 =$$
$$\frac{40 - 35}{40} \times 100 = \frac{....}{....} \times 100 = 12{,}5\%$$

2. Calcule a taxa de aprovação de um professor de uma classe de 45 alunos, sabendo que obtiveram aprovação 36 alunos.

$$TAE = \frac{\text{nº de aprovação}}{\text{nº matrícula final}} \times 100 =$$
$$\frac{....}{....} \times \frac{....}{....} = 80\%$$

Exercícios

1. Considere a série estatística:

SÉRIES	ALUNOS MATRICULADOS	%
1ª	546	
2ª	328	
3ª	280	
4ª	120	
Total	1.274	

Complete-a, determinando as percentagens com uma casa decimal e fazendo a compensação, se necessário.

2. Uma escola apresentava, no final do ano, o seguinte quadro:

SÉRIES	MATRÍCULAS	
	MARÇO	NOVEMBRO
1ª	480	475
2ª	458	456
3ª	436	430
4ª	420	420
Total	1.794	1.781

a. Calcule a taxa de evasão por série.
b. Calcule a taxa de evasão da escola.

3. Considere a tabela abaixo:

EVOLUÇÃO DAS RECEITAS DO CAFÉ INDUSTRIALIZADO
JAN./ABR. — 2008

MESES	VALOR (US$ milhões)
Janeiro	33,3
Fevereiro	54,1
Março	44,5
Abril	52,9
Total	184,8

Dados fictícios.

a. Complete-a com uma coluna de taxas percentuais.
b. Como se distribuem as receitas em relação ao total?
c. Qual o desenvolvimento das receitas de um mês para o outro?
d. Qual o desenvolvimento das receitas em relação ao mês de janeiro?

4. São Paulo tinha, em 1992, uma população de 32.182,7 mil habitantes. Sabendo que sua área terrestre é de 248.256 km^2, calcule a sua densidade demográfica.

5. Considerando que Minas Gerais, em 1992, apresentou (dados fornecidos pelo IBGE):
- população: 15.957,6 mil habitantes;
- superfície: 586.624 km^2;
- nascimentos: 292.036;
- óbitos: 99.281.

Calcule:
a. o índice da densidade demográfica;
b. a taxa de natalidade;
c. a taxa de mortalidade.

6. Uma frota de 40 caminhões, transportando, cada um, oito toneladas, dirige-se a duas cidades **A** e **B**. Na cidade **A** são descarregados 65% desses caminhões, por sete homens, trabalhando sete horas. Os caminhões restantes seguem para a cidade **B**, onde quatro homens gastam cinco horas para o seu descarregamento. Em que cidade se obteve melhor produtividade?

7. Um professor preencheu um quadro, enviado pela D.E., com os seguintes dados:

SÉRIE E TURMA	Nº DE ALUNOS 30.03	Nº DE ALUNOS 30.11	PROMOVIDOS SEM RECUPERAÇÃO	RETIDOS SEM RECUPERAÇÃO	EM RECUPERAÇÃO	RECUPERADOS	NÃO RECUPERADOS	TOTAL GERAL	
								PROMOVIDOS	RETIDOS
1º B	49	44	35	03	06	05	01	40	04
1º C	49	42	42	00	00	00	00	42	00
1º E	47	35	27	00	08	03	05	30	05
1º F	47	40	33	06	01	00	01	33	07
Total	192	161	137	09	15	08	07	145	16

Calcule:

a. a taxa de evasão, por classe;

b. a taxa de evasão total;

c. a taxa de aprovação, por classe;

d. a taxa de aprovação geral;

e. a taxa de recuperação, por classe;

f. a taxa de recuperação geral;

g. a taxa de reprovação na recuperação geral;

h. a taxa de aprovação, sem a recuperação;

i. a taxa de retidos, sem a recuperação.

4 GRÁFICOS ESTATÍSTICOS

4.1 Gráfico estatístico

O **gráfico estatístico** é uma forma de apresentação dos dados estatísticos, cujo objetivo é o de produzir, no investigador ou no público em geral, uma impressão mais rápida e viva do fenômeno em estudo, já que os gráficos falam mais rápido à compreensão que as séries.

Para tornarmos possível uma representação gráfica, estabelecemos uma correspondência entre os termos da série e determinada figura geométrica, de tal modo que cada elemento da série seja representado por uma figura proporcional.

A representação gráfica de um fenômeno deve obedecer a certos requisitos fundamentais para ser realmente útil:

a. **Simplicidade** — o gráfico deve ser destituído de detalhes de importância secundária, assim como de traços desnecessários que possam levar o observador a uma análise morosa ou com erros.
b. **Clareza** — o gráfico deve possibilitar uma correta interpretação dos valores representativos do fenômeno em estudo.
c. **Veracidade** — o gráfico deve expressar a verdade sobre o fenômeno em estudo.

Os principais tipos de gráficos são os **diagramas**, os **cartogramas** e os **pictogramas**.

4.2 Diagramas

Os **diagramas** são gráficos geométricos de, no máximo, duas dimensões; para sua construção, em geral, fazemos uso do sistema cartesiano.

Dentre os principais diagramas, destacamos:

4.2.1 Gráfico em linha ou em curva

Este tipo de gráfico se utiliza da linha poligonal para representar a série estatística.

O **gráfico em linha** constitui uma aplicação do processo de representação das funções num sistema de coordenadas cartesianas.

Como sabemos, nesse sistema fazemos uso de duas retas perpendiculares; as retas são os **eixos coordenados** e o ponto de intersecção, a **origem**. O eixo horizontal é denominado **eixo das abscissas** (ou eixo dos **x**) e o vertical, **eixo das ordenadas** (ou eixo dos **y**).

Para tornar bem clara a explanação, consideremos a seguinte série:

PRODUÇÃO BRASILEIRA
DE ÓLEO DE DENDÊ
1987-92

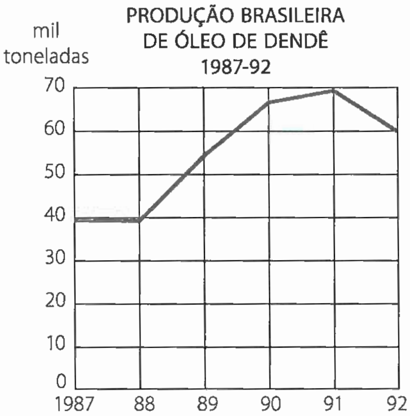
ANOS	QUANTIDADE (1.000 t)
1987	39,3
1988	39,1
1989	53,9
1990	65,1
1991	69,1
1992	59,5

FONTE: Agropalma.

Vamos tomar os anos como abscissas e as quantidades como ordenadas.

Assim, um ano dado (**x**) e a respectiva quantidade (**y**) formam um par ordenado (**x**, **y**), que pode ser representado num sistema cartesiano.

Determinados, graficamente, todos os pontos da série, usando as coordenadas, ligamos todos esses pontos, dois a dois, por segmentos de reta, o que irá nos dar uma **poligonal**, que é o **gráfico em linha** ou **em curva** correspondente à série em estudo (Figura 4.1).



FONTE: Agropalma.

FIGURA 4.1

NOTAS:

- No exemplo dado, o zero foi indicado no eixo vertical, mas, por razões óbvias, não foi indicado no eixo horizontal. Observe que o zero, de modo geral, deverá ser indicado sempre que possível, especialmente no eixo vertical. Se, por alguma razão, for impossível tal indicação e se essa omissão puder levar o observador a conclusões errôneas, é prudente chamar a atenção para a omissão por um dos meios indicados nas Figuras 4.2, 4.3 e 4.4:

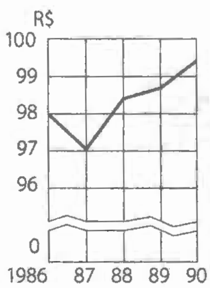


FIGURA 4.2

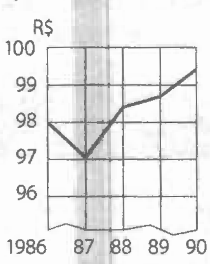


FIGURA 4.3

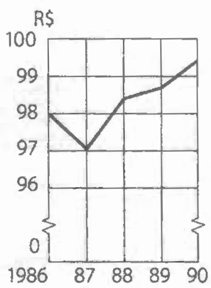


FIGURA 4.4

- Com o intuito de melhorar o aspecto visual, podemos sombrear ou hachurar o gráfico. Assim, o gráfico da Figura 4.3 toma o seguinte aspecto:

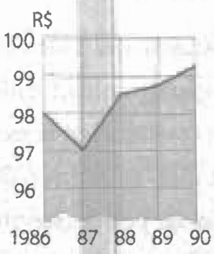


FIGURA 4.5

- Quando representamos, em um mesmo sistema de coordenadas, a variação de dois fenômenos, a parte interna da figura formada pelos gráficos desses fenômenos é denominada **área de excesso**:

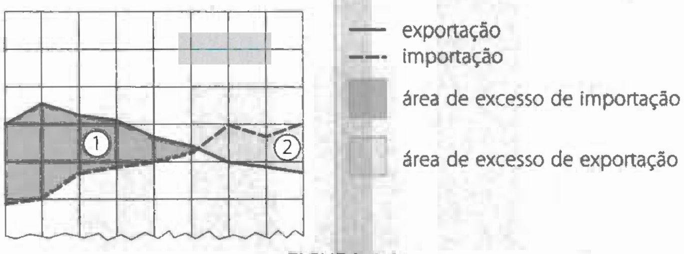


FIGURA 4.6

4.2.2 Gráfico em colunas ou em barras

É a representação de uma série por meio de **retângulos**, dispostos **verticalmente** (em colunas) ou **horizontalmente** (em barras).

Quando em colunas, os retângulos têm a mesma base e as alturas são proporcionais aos respectivos dados.

Quando em barras, os retângulos têm a mesma altura e os comprimentos são proporcionais aos respectivos dados.

Assim estamos assegurando a proporcionalidade entre as áreas dos retângulos e os dados estatísticos.

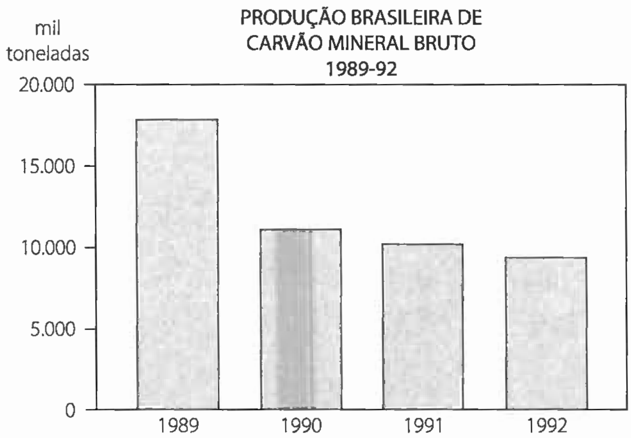
Exemplos:

a. **Gráfico em colunas**

PRODUÇÃO BRASILEIRA
DE CARVÃO MINERAL BRUTO
1989-92

ANOS	QUANTIDADE PRODUZIDA (1.000 t)
1989	18.196
1990	11.168
1991	10.468
1992	9.241

FONTE: Ministério da Agricultura.



FONTE: Ministério da Agricultura.

FIGURA 4.7

b. Gráfico em barras

EXPORTAÇÕES BRASILEIRAS
MARÇO — 1995

ESTADOS	VALOR (US$ milhões)
São Paulo	1.344
Minas Gerais	542
Rio Grande do Sul	332
Espírito Santo	285
Paraná	250
Santa Catarina	202

FONTE: SECEX.

FONTE: SECEX.

FIGURA 4.8

NOTAS:

- Sempre que os dizeres a serem inscritos são extensos, devemos dar preferência ao gráfico em barras (séries geográficas e específicas). Se, porém, ainda assim preferirmos o gráfico em colunas, os dizeres deverão ser dispostos de baixo para cima, nunca ao contrário.
- A ordem a ser observada é a **cronológica**, se a série for histórica, e a **decrescente**, se for geográfica ou categórica.
- A distância entre as colunas (ou barras), por questões estéticas, não deverá ser menor que a metade nem maior que os dois terços da largura (ou da altura) dos retângulos.

4.2.3 Gráfico em colunas ou em barras múltiplas

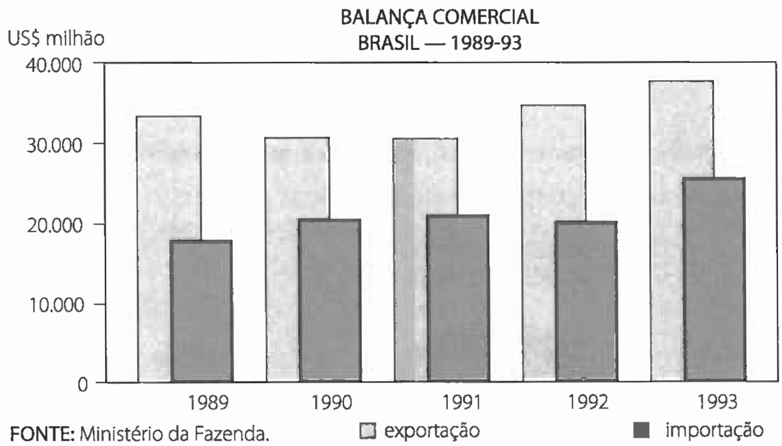
Este tipo de gráfico é geralmente empregado quando queremos representar, simultaneamente, dois ou mais fenômenos estudados com o propósito de comparação.

Exemplo:

BALANÇA COMERCIAL DO BRASIL
1989-93

ESPECIFICAÇÕES	VALOR (US$ 1.000.000)				
	1989	1990	1991	1992	1993
Exportação (FOB)	34.383	31.414	31.620	35.793	38.783
Importação	18.263	20.661	21.041	20.554	25.711

FONTE: Ministério da Fazenda.



FONTE: Ministério da Fazenda.

FIGURA 4.9

4.2.4 Gráfico em setores

Este gráfico é construído com base em um círculo, e é empregado sempre que desejamos ressaltar a participação do dado no total.

O total é representado pelo círculo, que fica dividido em tantos setores quantas são as partes.

Os setores são tais que suas áreas são respectivamente proporcionais aos dados da série.

Obtemos cada setor por meio de uma regra de três simples e direta, lembrando que o total da série corresponde a 360°.

Exemplo:

Dada a série:

REBANHO SUÍNO DO
SUDESTE DO BRASIL
1992

ESTADOS	QUANTIDADE (mil cabeças)
Minas Gerais	3.363,7
Espírito Santo	430,4
Rio de Janeiro	308,5
São Paulo	2.035,9
Total	6.138,5

FONTE: IBGE.

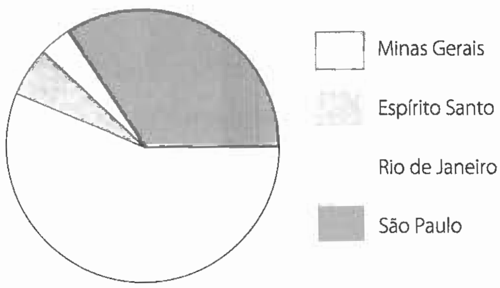
temos:

$$\begin{array}{l} 6.138,5 — 360° \\ 3.363,7 — x_1 \end{array} \Rightarrow x_1 = 197,2 \Rightarrow x_1 = 197°$$

$$x_2 = 25,2 \Rightarrow x_2 = 25°$$
$$x_3 = 18,0 \Rightarrow x_3 = 18°$$
$$x_4 = 119,3 \Rightarrow x_4 = 120°$$

Com esses dados (valores em graus), marcamos num círculo de raio arbitrário, com um transferidor, os arcos correspondentes, obtendo o gráfico:

REBANHO SUÍNO DO
SUDESTE DO BRASIL
1992



FONTE: IBGE.

FIGURA 4.10

NOTAS:

- O gráfico em setores só deve ser empregado quando há, no máximo, sete dados.
- Se a série já apresenta os dados percentuais, obtemos os respectivos valores em graus multiplicando o valor percentual por 3,6.

4.3 Gráfico polar

É o gráfico ideal para representar **séries temporais cíclicas**, isto é, séries temporais que apresentam em seu desenvolvimento determinada periodicidade, como, por exemplo, a variação da precipitação pluviométrica ao longo do ano ou da temperatura ao longo do dia, a arrecadação da Zona Azul durante a semana, o consumo de energia elétrica durante o mês ou o ano, o número de passageiros de uma linha de ônibus ao longo da semana etc.

O gráfico polar faz uso do sistema de coordenadas polares.

Exemplo:

Dada a série:

PRECIPITAÇÃO PLUVIOMÉTRICA
RECIFE — 1993

MESES	MILÍMETROS
Janeiro	49,6
Fevereiro	93,1
Março	63,6
Abril	135,3
Maio	214,7
Junho	277,9
Julho	183,6
Agosto	161,3
Setembro	49,2
Outubro	40,8
Novembro	28,6
Dezembro	33,3

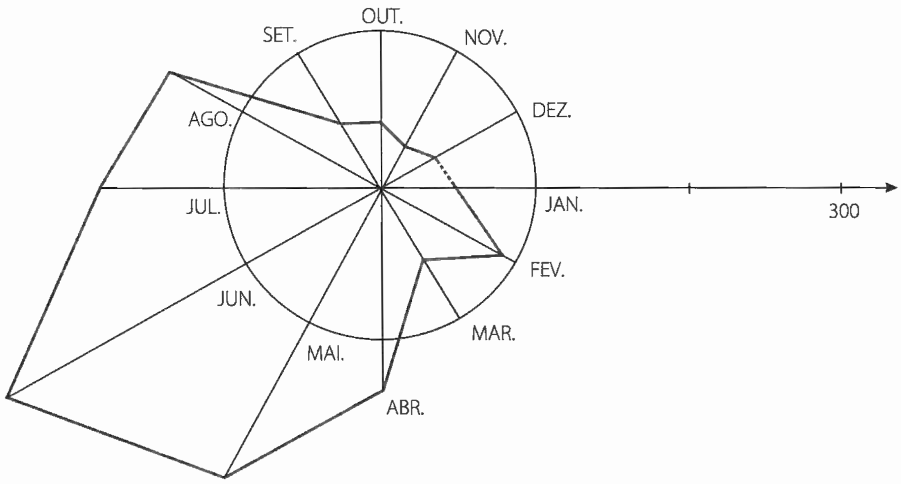
FONTE: Ministério da Agricultura.

- traçamos uma circunferência de raio arbitrário (em particular, damos preferência ao raio de comprimento proporcional à média dos valores da série);
- construímos uma semirreta (de preferência na horizontal) partindo de **O** (polo) e com uma escala (eixo polar);

- dividimos a circunferência em tantos arcos quantas forem as unidades temporais;
- traçamos, a partir do centro **O** (polo), semirretas passando pelos pontos de divisão;
- marcamos os valores correspondentes da variável, iniciando pela semirreta horizontal (eixo polar);
- ligamos os pontos encontrados com segmentos de reta;
- se pretendemos fechar a poligonal obtida, empregamos uma linha interrompida.

Assim, para o nosso exemplo, temos:

PRECIPITAÇÃO PLUVIOMÉTRICA
RECIFE — 1993



FONTE: Ministério da Agricultura.

FIGURA 4.11

4.4 Cartograma

O **cartograma** é a representação sobre uma carta geográfica.

Este gráfico é empregado quando o objetivo é o de figurar os dados estatísticos diretamente relacionados com áreas geográficas ou políticas.

Distinguimos duas aplicações:

a. Representar dados absolutos (população) — neste caso, lançamos mão, em geral, dos **pontos**, em número proporcional aos dados (Figura 4.12).

b. Representar dados relativos (densidade) — neste caso, lançamos mão, em geral, de **hachuras** (Figura 4.13).

Exemplo:

Dada a série:

POPULAÇÃO PROJETADA DA REGIÃO SUL DO BRASIL — 1994

ESTADOS	POPULAÇÃO (hab.)	ÁREA (km^2)	DENSIDADE
Paraná	8.651.100	199.324	43,4
Santa Catarina	4.767.800	95.318	50,0
Rio Grande do Sul	9.475.900	280.674	33,8

FONTE: IBGE.

obtemos os seguintes cartogramas:

POPULAÇÃO PROJETADA DA REGIÃO SUL BRASIL — 1994

FIGURA 4.12

DENSIDADE POPULACIONAL PROJETADA DA REGIÃO SUL DO BRASIL — 1994

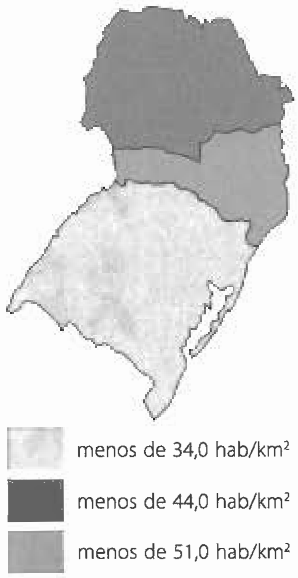


FIGURA 4.13

NOTA:

- Quando os números absolutos a ser representados forem muito grandes, no lugar de pontos podemos empregar hachuras.

4.5 Pictograma

O **pictograma** constitui um dos processos gráficos que melhor fala ao público, pela sua forma ao mesmo tempo atraente e sugestiva. A representação gráfica consta de **figuras**.

Exemplo:

Para a série:

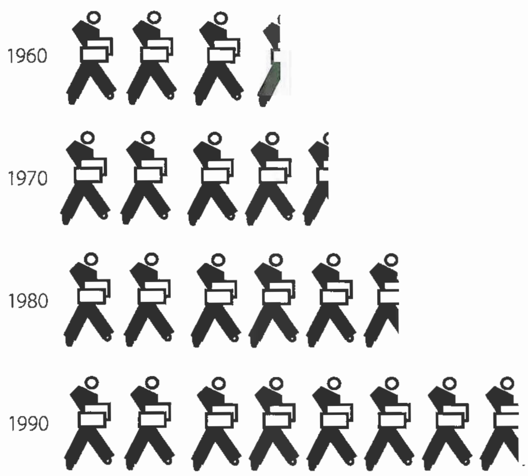
POPULAÇÃO DO BRASIL
1960-90

ANOS	HABITANTES (milhares)
1960	70.070,4
1970	93.139,0
1980	118.562,5
1990	155.822,4

FONTE: IBGE.

temos a seguinte representação pictórica:

POPULAÇÃO DO BRASIL
1960-90



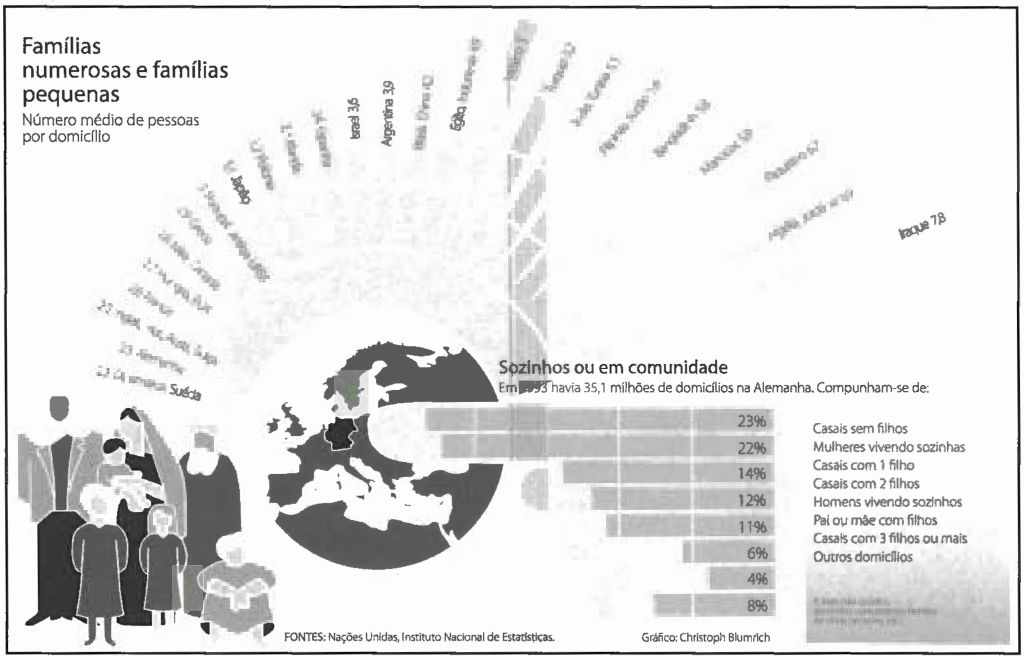
Cada símbolo representa 20.000.000 de habitantes.

FONTE: IBGE.

FIGURA 4.14

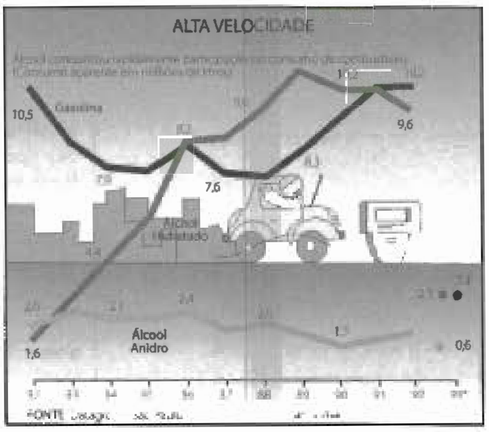
Na verdade, o gráfico referente à Figura 4.14 é essencialmente um gráfico em barras; porém, as figuras o tornam mais atrativo, o que, provavelmente, despertará a atenção do leitor para o seu exame.

Na confecção de gráficos pictóricos temos de utilizar muita criatividade, procurando obter uma otimização na união da arte com a técnica. Eis alguns exemplos:



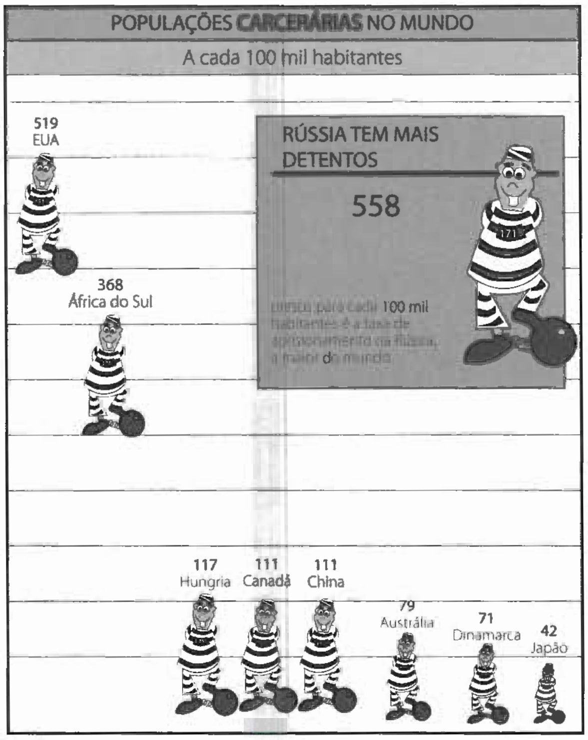
Doutschland, ago. 1993.

FIGURA 4.15



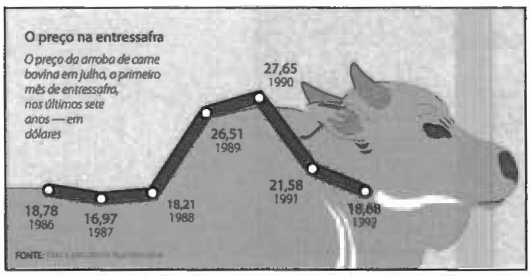
Globo Rural, jul. 1993.

FIGURA 4.16



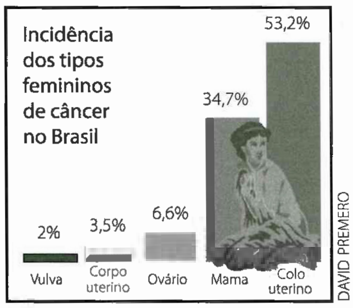
Folha de S. Paulo, set. 1994.

FIGURA 4.17



Veja, 10 out. 1992.

FIGURA 4.18



Veja, 12 abr. 1995.

FIGURA 4.19

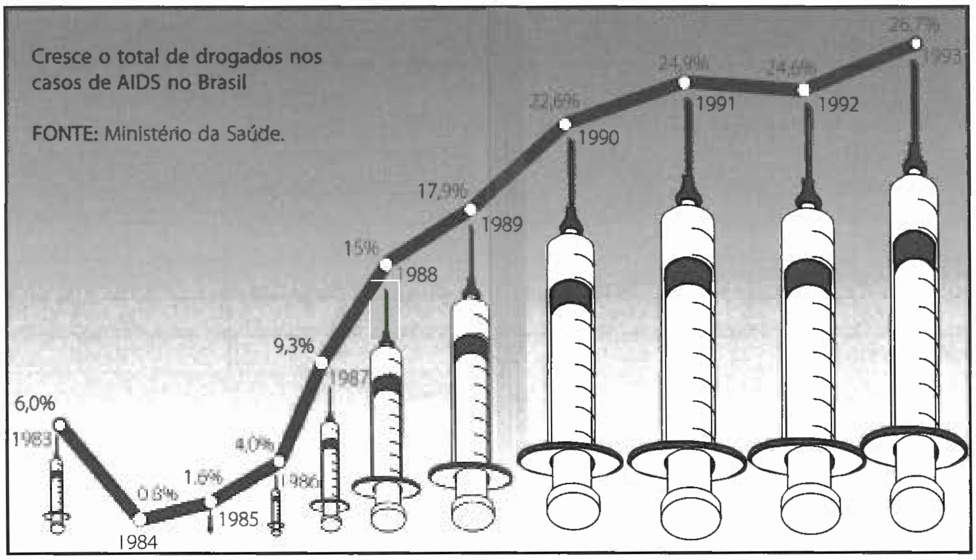


FIGURA 4.20

Exercícios

1. Represente a série abaixo usando o gráfico em linha:

COMÉRCIO EXTERIOR
BRASIL — 1984-93

ANOS	QUANTIDADE (1.000 t)	
	EXPORTAÇÃO	IMPORTAÇÃO
1984	141.737	53.988
1985	146.351	48.870
1986	133.832	60.597
1987	142.378	61.975
1988	169.666	58.085
1989	177.033	57.293
1990	168.095	57.184
1991	165.974	63.278
1992	167.295	68.059
1993	182.561	77.813

FONTE: Min. Indústria, Comércio e Turismo.

2. Represente as tabelas usando o gráfico em colunas:

a.

PRODUÇÃO BRASILEIRA DE PETRÓLEO BRUTO
1991-93

ANOS	QUANTIDADE (1.000 m³)
1991	36.180,4
1992	36.410,5
1993	37.164,3

FONTE: Petrobras.

b.

ENTREGA DE GASOLINA PARA
CONSUMO BRASIL — 1988-91

ANOS	VOLUME (1.000 m³)
1988	9.267,7
1989	9.723,1
1990	10.121,3
1991	12.345,4

FONTE: IBGE.

3. Usando o gráfico em barras, represente as tabelas:

a.

PRODUÇÃO DE OVOS DE GALINHA
BRASIL — 1992

REGIÕES	QUANTIDADE (1.000 dúzias)
Norte	57.297
Nordeste	414.804
Sudeste	984.659
Sul	615.978
Centro-Oeste	126.345

FONTE: IBGE.

b.

PRODUÇÃO DE VEÍCULOS DE AUTOPROPULSÃO
BRASIL — 1993

TIPOS	QUANTIDADE
Automóveis	1.100.278
Comerciais leves	224.387
Comerciais pesados	66.771

FONTE: ANFAVEA.

4. Represente as tabelas por meio de gráficos em setores:

a.

ÁREA TERRESTRE
BRASIL

REGIÕES	RELATIVA (%)
Norte	45,25
Nordeste	18,28
Sudeste	10,85
Sul	6,76
Centro-Oeste	18,86
Total	100,00

FONTE: IBGE.

b.

PRODUÇÃO DE FERRO-GUSA
BRASIL — 1993

UNIDADES DA FEDERAÇÃO	PRODUÇÃO (1.000 t)
Minas Gerais	12.888
Espírito Santo	3.174
Rio de Janeiro	5.008
São Paulo	2.912

FONTE: Instituto Brasileiro de Siderurgia.

5. Represente a tabela por meio de um gráfico de colunas múltiplas:

PROPORÇÃO DOS DOMICÍLIOS
POR CONDIÇÃO DE OCUPAÇÃO
BRASIL — 1990-91

ANOS	NATUREZA		
	PRÓPRIOS (%)	ALUGADOS (%)	CEDIDOS (%)
1990	62,7	22,9	14,4
1991	70,3	16,5	13,2

FONTE: IBGE.

6. Represente as tabelas por meio de gráficos polares:

a.

VENDA DE VACINA CONTRA AFTOSA
BRASIL — 1992

MESES	US$ milhões
Janeiro	37,30
Fevereiro	41,20
Março	38,55
Abril	47,70
Maio	40,65
Junho	44,70
Julho	41,20
Agosto	46,00
Setembro	41,00
Outubro	55,00
Novembro	52,80
Dezembro	35,40

FONTE: Sindan.

b.

PRECIPITAÇÃO PLUVIOMÉTRICA
FLORIANÓPOLIS — 1993

MESES	MILÍMETROS
Janeiro	165,7
Fevereiro	106,6
Março	71,6
Abril	34,7
Maio	184,9
Junho	102,7
Julho	198,3
Agosto	36,8
Setembro	72,2
Outubro	147,8
Novembro	175,1
Dezembro	198,3

FONTE: Ministério da Agricultura.

7. Procure, em jornais e revistas especializados, dois exemplos de cada um dos gráficos estudados.

5 DISTRIBUIÇÃO DE FREQUÊNCIA

5.1 Tabela primitiva ROL

Vamos considerar, neste capítulo, em particular, a forma pela qual podemos descrever os dados estatísticos resultantes de variáveis quantitativas, como é o caso de notas obtidas pelos alunos de uma classe, estaturas de um conjunto de pessoas, salários recebidos pelos operários de uma fábrica etc.

Suponhamos termos feito uma coleta de dados relativos às estaturas de quarenta alunos, que compõem uma amostra dos alunos de um colégio **A**, resultando a seguinte tabela de valores:

ESTATURAS DE 40 ALUNOS DO COLÉGIO A

166	160	161	150	162	160	165	167	164	160
162	161	168	163	156	173	160	155	164	168
155	152	163	160	155	155	169	151	170	164
154	161	156	172	153	157	156	158	158	161

TABELA 5.1

A esse tipo de tabela, cujos elementos não foram numericamente organizados, denominamos **tabela primitiva**.

Partindo dos dados acima — tabela primitiva —, é difícil averiguar em torno de que valor tendem a se concentrar as estaturas, qual a menor ou qual a maior estatura ou, ainda, quantos alunos se acham abaixo ou acima de uma dada estatura.

Assim, conhecidos os valores de uma variável, é difícil formarmos uma ideia exata do comportamento do grupo como um todo, a partir dos dados não ordenados.

A maneira mais simples de organizar os dados é através de uma certa ordenação (crescente ou descrescente). A tabela obtida após a ordenação dos dados recebe o nome de **rol**.

ESTATURAS DE 40 ALUNOS DO COLÉGIO A

150	154	155	157	160	161	162	164	166	169
151	155	156	158	160	161	162	164	167	170
152	155	156	158	160	161	163	164	168	172
153	155	156	160	160	161	163	165	168	173

TABELA 5.2

Agora, podemos saber, com relativa facilidade, qual a menor estatura (150 cm) e qual a maior (173 cm); que a amplitude de variação foi de 173 – 150 = 23 cm; e, ainda, a ordem que um valor particular da variável ocupa no conjunto. Com um exame mais acurado, vemos que há uma concentração das estaturas em algum valor entre 160 cm e 165 cm e, mais ainda, que há poucos valores abaixo de 155 cm e acima de 170 cm.

5.2 Distribuição de frequência

No exemplo que trabalhamos, a variável em questão, **estatura**, será observada e estudada muito mais facilmente quando dispusermos valores ordenados em uma coluna e colocarmos, ao lado de cada valor, o número de vezes que aparece repetido.

Denominamos **frequência** o número de alunos que fica relacionado a um determinado valor da variável. Obtemos, assim, uma tabela que recebe o nome de **distribuição de frequência**:

ESTAT. (cm)	FREQ.	ESTAT. (cm)	FREQ.	ESTAT. (cm)	FREQ.
150	1	158	2	167	1
151	1	160	5	168	2
152	1	161	4	169	1
153	1	162	2	170	1
154	1	163	2	172	1
155	4	164	3	173	1
156	3	165	1	**Total**	40
157	1	166	1		

TABELA 5.3[1]

[1] A tabela foi tripartida para não ocupar muito espaço.

Mas o processo dado é ainda incoveniente, já que exige muito espaço, mesmo quando o número de valores da variável (**n**) é de tamanho razoável. Sendo possível, a solução mais aceitável, pela própria natureza da **variável contínua**, é o agrupamento dos valores em vários intervalos.

Assim, se um dos intervalos for, por exemplo, 154 ⊢ 158[2], em vez de dizermos que a estatura de um aluno é de 154 cm; de quatro alunos, 155 cm; de três alunos, 156 cm; e de um aluno, 157 cm, diremos que nove alunos têm estaturas entre 154, inclusive, e 158 cm.

Deste modo, estaremos agrupando os valores da variável em intervalos, sendo que, em Estatística, preferimos chamar os intervalos de **classes**.

Chamando de **frequência de uma classe** o número de valores da variável pertencentes à classe, os dados da Tabela 5.3 podem ser dispostos como na Tabela 5.4, denominada **distribuição de frequência com intervalos de classe**:

ESTATURAS DE 40 ALUNOS
DO COLÉGIO A

ESTATURAS (cm)	FREQUÊNCIA
150 ⊢ 154	4
154 ⊢ 158	9
158 ⊢ 162	11
162 ⊢ 166	8
166 ⊢ 170	5
170 ⊢ 174	3
Total	40

TABELA 5.4

Ao agruparmos os valores da variável em classes, ganhamos em simplicidade, mas perdemos em pormenores. Assim, na Tabela 5.3 podemos verificar, facilmente, que quatro alunos têm 161 cm de altura e que não existe nenhum aluno com 171 cm de altura. Já na Tabela 5.4 não podemos ver se algum aluno tem a estatura de 159 cm. No entanto, sabemos, com segurança, que onze alunos têm estatura compreendida entre 158 e 162 cm.

O que pretendemos com a construção dessa nova tabela é realçar o que há de essencial nos dados e, também, tornar possível o uso de técnicas analíticas para sua total descrição, até porque a Estatística tem por finalidade específica analisar o conjunto de valores, desinteressando-se por casos isolados.

NOTAS:

- Se nosso intuito é, desde o início, a obtenção de uma distribuição de frequência com intervalos de classe, basta, a partir da Tabela 5.1, fazermos uma **tabulação**, como segue, onde cada traço corresponde a um valor:

[2] 154 ⊢ 158 é um intervalo fechado à esquerda e aberto à direita, tal que: $154 \leq x < 158$.

ESTATURAS (cm)	TABULAÇÃO	FREQUÊNCIA
150 ⊢ 154	□	4
154 ⊢ 158	⊠ □	9
158 ⊢ 162	⊠ ⊠ |	11
162 ⊢ 166	⊠ ⊏	8
166 ⊢ 170	⊠	5
170 ⊢ 174	⊏	3
Total		40

TABELA 5.5

- Quando os dados estão organizados em uma distribuição de frequência, são comumente denominados **dados agrupados**.

5.3 Elementos de uma distribuição de frequência

5.3.1 Classe

Classes de frequência ou, simplesmente, **classes** são intervalos de variação da variável.

As classes são representadas simbolicamente por **i**, sendo $i = 1, 2, 3, ..., k$ (onde **k** é o número total de classes da distribuição).

Assim, em nosso exemplo, o intervalo 154 ⊢ 158 define a segunda classe ($i = 2$). Como a distribuição é formada de seis classes, podemos afirmar que $k = 6$.

5.3.2 Limites de classe

Denominamos **limites de classe** os extremos de cada classe.

O menor número é o **limite inferior da classe** (ℓ_i) e o maior número, o **limite superior da classe** (L_i).

Na segunda classe, por exemplo, temos:

$$\ell_2 = 154 \qquad \text{e} \qquad L_2 = 158$$

NOTA:

- Os intervalos de classe devem ser escritos, de acordo com a Resolução 886/66 do IBGE, em termos de **desta quantidade até menos aquela**, empregando, para isso, o símbolo ⊢ (inclusão de ℓ_i e exclusão de L_i). Assim, o indivíduo com uma estatura de 158 cm está incluído na terceira classe ($i = 3$) e não na segunda.

5.3.3 Amplitude de um intervalo de classe

Amplitude de um intervalo de classe ou, simplesmente, **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 e indicada por $\mathbf{h_i}$. Assim:

$$h_i = L_i - \ell_i$$

Na distribuição da Tabela 5.4, temos:

$$h_2 = L_2 - \ell_2 \Rightarrow h_2 = 158 - 154 = 4 \Rightarrow h_2 = 4 \text{ cm}$$

5.3.4 Amplitude total da distribuição

Amplitude total da distribuição (AT) é a diferença entre o limite superior da última classe **(limite superior máximo)** e o limite inferior da primeira classe **(limite inferior mínimo)**:

$$AT = L(\text{máx.}) - \ell\,(\text{mín.})$$

Em nosso exemplo, temos:

$$AT = 174 - 150 = 24 \Rightarrow AT = 24 \text{ cm}$$

NOTA:

- É evidente que, se as classes possuem o mesmo intervalo, verificamos a relação:

$$\frac{AT}{h_i} = k$$

Em nosso exemplo:

$$\frac{24}{4} = 6$$

5.3.5 Amplitude amostral

Amplitude amostral (AA) é a diferença entre o valor máximo e o valor mínimo da amostra:

$$AA = x(máx.) - x(mín.)$$

Em nosso exemplo, temos:

$$AA = 173 - 150 = 23 \Rightarrow AA = 23 \text{ cm}$$

Observe que a amplitude total da distribuição jamais coincide com a amplitude amostral.

5.3.6 Ponto médio de uma classe

Ponto médio de uma classe (x_i) é, como o próprio nome indica, o ponto que divide o intervalo de classe em duas partes iguais.

Para obtermos o ponto médio de uma classe, calculamos a semissoma dos limites da classe (média aritmética):

$$x_i = \frac{\ell_i + L_i}{2}$$

Assim, o ponto médio da segunda classe, em nosso exemplo, é:

$$x_2 = \frac{\ell_2 + L_2}{2} \Rightarrow x_2 = \frac{154 + 158}{2} = 156 \Rightarrow x_2 = 156 \text{ cm}$$

NOTA:

- O ponto médio de uma classe é o valor que a representa.

5.3.7 Frequência simples ou absoluta

Frequência simples ou **frequência absoluta** ou, simplesmente, **frequência** de uma classe ou de um valor individual é o número de observações correspondentes a essa classe ou a esse valor.

A frequência simples é simbolizada por f_i (lemos: **f índice i ou frequência da classe i**).

Assim, em nosso exemplo, temos:

$$f_1 = 4, f_2 = 9, f_3 = 11, f_4 = 8, f_5 = 5 \text{ e } f_6 = 3$$

A soma de todas as frequências é representada pelo símbolo de somatório[3]:

$$\sum_{i=1}^{k} f_i$$

É evidente que:

$$\sum_{i=1}^{k} f_i = n$$

Para a distribuição em estudo, temos:

$$\sum_{i=1}^{6} f_i = 40$$

Não havendo possibilidade de engano, usamos:

$$\sum f_i = 40$$

Podemos, agora, dar à distribuição de frequência das estaturas dos quarenta alunos do Colégio **A** a seguinte representação tabular técnica:

ESTATURAS DE 40 ALUNOS
DO COLÉGIO A

i	ESTATURAS (cm)	f_i
1	150 ⊢ 154	4
2	154 ⊢ 158	9
3	158 ⊢ 162	11
4	162 ⊢ 166	8
5	166 ⊢ 170	5
6	170 ⊢ 174	3
		$\Sigma f_i = 40$

TABELA 5.6

[3] Consulte o **Apêndice — Instrumental Matemático**, para uma revisão dos assuntos **Sequência** e **Somatório** (p. 185).

5.4 Número de classes Intervalos de classe

A primeira preocupação que temos, na construção de uma distribuição de frequência, é a determinação do número de classes e, consequentemente, da amplitude e dos limites dos intervalos de classe.

Para a determinação do número de classes de uma distribuição podemos lançar mão da **regra de Sturges**, que nos dá o número de classes em função do número de valores da variável:

$$k \cong 1 + 3{,}3 \cdot \log n$$

onde:

k é o número de classe;

n é o número total de dados.

Essa regra nos permite obter a seguinte tabela:

n	i
3 H 5	3
6 H 11	4
12 H 22	5
23 H 46	6
47 H 90	7
91 H 181	8
182 H 362	9
...	...

TABELA 5.7

Além da regra de Sturges, existem outras fórmulas empíricas que pretendem resolver o problema da determinação do número de classes que deve ter a distribuição[4]. Entretanto, a verdade é que essas fórmulas não nos levam a uma decisão final; esta vai depender, na realidade, de um julgamento pessoal, que deve estar ligado à natureza dos dados, da unidade usada para expressá-los e, ainda, do objetivo que se tem em vista, procurando, sempre que possível, evitar classe com frequência nula ou com frequência relativa[5] muito exagerada etc.

[4] Há quem prefira: $k \cong \sqrt{h}$.

[5] Item 5.5, p. 55.

Decidido o número de classes que deve ter a distribuição, resta-nos resolver o problema da determinação da amplitude do intervalo de classe, o que conseguimos divivindo a amplitude total pelo número de classes:

$$h \cong \frac{AT}{i}$$

Quando o resultado não é exato, devemos arredondá-lo para mais.

Outro problema que surge é a escolha dos limites dos intervalos, os quais deverão ser tais que forneçam, na medida do possível, para pontos médios, números que facilitem os cálculos — **números naturais**.

Em nosso exemplo, temos:

para $n = 40$, pela Tabela 5.7, $i = 6$.

Logo:

$$h = \frac{173 - 150}{6} = \frac{23}{6} = 3{,}8 = 4,$$

isto é, seis classes de intervalos iguais a 4.

Resolva

1. As notas obtidas por 50 alunos de uma classe foram:

1	2	3	4	5	6	6	7	7	8
2	3	3	4	5	6	6	7	8	8
2	3	4	4	5	6	6	7	8	9
2	3	4	5	5	6	6	7	8	9
2	3	4	5	5	6	7	7	8	9

a. Complete a distribuição de frequência abaixo:

i	NOTAS	x_i	f_i
1	$0 \vdash 2$	1	1
2	$2 \vdash 4$	....	....
3	$4 \vdash 6$	....	....
4	$6 \vdash 8$	....	....
5	$8 \vdash 10$	....	....
			$\Sigma f_i = 50$

b. Agora, responda:

1. Qual a amplitude amostral?

2. Qual a amplitude da distribuição?

3. Qual o número de classes da distribuição?

4. Qual o limite inferior da quarta classe?

5. Qual o limite superior da classe de ordem 2?

6. Qual a amplitude do segundo intervalo de classe?

c. Complete:

1. $h_3 = ...$	**3.** $\ell_1 = ...$	**5.** $x_2 = ...$
2. $n = ...$	**4.** $L_3 = ...$	**6.** $f_5 = ...$

5.5 Tipos de frequências

Frequências simples ou absolutas (f_i) são os valores que realmente representam o número de dados de cada classe.

Como vimos, a soma das frequências simples é igual ao número total dos dados:

$$\sum f_i = n$$

Frequências relativas (fr_i) são os valores das razões entre as frequências simples e a frequência total:

$$fr_i = \frac{f_i}{\sum f_i}$$

Logo, a frequência relativa da terceira classe, em nosso exemplo (Tabela 5.6), é:

$$fr_3 = \frac{f_3}{\sum f_i} \Rightarrow fr_3 = \frac{11}{40} = 0{,}275 \Rightarrow fr_3 = 0{,}275$$

Evidentemente:

$$\sum fr_i = 1 \text{ ou } 100\%$$

NOTA:

- O propósito das frequências relativas é o de permitir a análise ou facilitar as comparações.

Frequência acumulada (F_i) é o total das frequências de todos os valores inferiores ao limite superior do intervalo de uma dada classe:

$$F_k = f_1 + f_2 + ... + f_k$$

ou

$$F_k = \Sigma f_i \ (i = 1, 2, ..., k)$$

Assim, no exemplo apresentado no início deste capítulo, a frequência acumulada correspondente à terceira classe é:

$$F_3 = \sum_{i=1}^{3} f_i = f_1 + f_2 + f_3 \Rightarrow F_3 = 4 + 9 + 11 \Rightarrow F_3 = 24,$$

o que significa existirem 24 alunos com estatura inferior a 162 cm (limite superior do intervalo da terceira classe).

Frequência acumulada relativa (Fr_i) de uma classe é a frequência acumulada da classe, dividida pela frequência total da distribuição:

$$Fr_i = \frac{F_i}{\sum f_i}$$

Assim, para a terceira classe, temos:

$$Fr_3 = \frac{F_3}{\sum f_i} \Rightarrow Fr_3 = \frac{24}{40} = 0,600 \Rightarrow Fr_3 = 0,600$$

Considerando a Tabela 5.4, podemos montar a seguinte tabela com as frequências estudadas:

i	ESTATURAS (cm)	f_i	x_i	fr_i	F_i	Fr_i
1	150 ⊢ 154	4	152	0,100	4	0,100
2	154 ⊢ 158	9	156	0,225	13	0,325
3	158 ⊢ 162	11	160	0,275	24	0,600
4	162 ⊢ 166	8	164	0,200	32	0,800
5	166 ⊢ 170	5	168	0,125	37	0,925
6	170 ⊢ 174	3	172	0,075	40	1,000
		$\Sigma = 40$		$\Sigma = 1,000$		

TABELA 5.8

O conhecimento dos vários tipos de frequência ajuda-nos a responder a muitas questões com relativa facilidade, como as seguintes:

a. Quantos alunos têm estatura entre 154 cm, inclusive, e 158 cm?

Esses são os valores da variável que formam a segunda classe. Como $f_2 = 9$, a resposta é: nove alunos.

b. Qual a percentagem de alunos cujas estaturas são inferiores a 154 cm?

Esses valores são os que formam a primeira classe. Como $fr_1 = 0{,}100$, obtemos a resposta multiplicando a frequência relativa por 100:

$$0{,}100 \times 100 = 10$$

Logo, a percentagem de alunos é 10%.

c. Quantos alunos têm estatura abaixo de 162 cm?

É evidente que as estaturas consideradas são aquelas que formam as classes de ordem 1, 2 e 3. Assim, o número de alunos é dado por:

$$f_1 + f_2 + f_3 = \sum_{i=1}^{3} f_i = F_3 = 24$$

Portanto, 24 alunos têm estatura abaixo de 162 cm.

d. Quantos alunos têm estatura não inferior a 158 cm? O número de alunos é dado por:

$$\sum_{i=3}^{6} f_i = f_3 + f_4 + f_5 + f_6 = 11 + 8 + 5 + 3 = 27$$

Ou então:

$$\sum_{i=3}^{6} f_i - F_2 = n - F_2 = 40 - 13 = 27$$

5.6 Distribuição de frequência sem intervalos de classe

Quando se trata de **variável discreta** de variação relativamente pequena, cada valor pode ser tomado como um intervalo de classe (intervalo degenerado) e, nesse caso, a distribuição é chamada **distribuição sem intervalos de classe**, tomando a seguinte forma:

x_i	f_i
x_1	f_1
x_2	f_2
⋮	⋮
x_n	f_n
	$\Sigma f_i = n$

TABELA 5.9

Exemplo:

Seja **x** a variável "número de cômodos das casas ocupadas por vinte famílias entrevistadas":

i	x_i	f_i
1	2	4
2	3	7
3	4	5
4	5	2
5	6	1
6	7	1
		$\Sigma = 20$

TABELA 5.10

Completada com os vários tipos de frequência, temos:

i	x_i	f_i	fr_i	F_i	Fr_i
1	2	4	0,20	4	0,20
2	3	7	0,35	11	0,55
3	4	5	0,25	16	0,80
4	5	2	0,10	18	0,90
5	6	1	0,05	19	0,95
6	7	1	0,05	20	1,00
		$\Sigma = 20$	$\Sigma = 1,00$		

TABELA 5.11

NOTA:

- **Se a variável toma numerosos valores distintos, é comum tratá-la como uma variável contínua, formando intervalos de classe de amplitude diferente de um.**
  Esse tratamento (arbitrário) abrevia o trabalho, mas acarreta alguma perda de precisão.

Resolva

1. Complete a distribuição abaixo, determinando as frequências simples:

i	x_i	f_i	F_i
1	2	....	2
2	3	....	9
3	4	....	21
4	5	....	29
5	6	....	34
		$\Sigma = 34$	

Exercícios

1. Conhecidas as notas de 50 alunos:

84	68	33	52	47	73	68	61	73	77
74	71	81	91	65	55	57	35	85	88
59	80	41	50	53	65	76	85	73	60
67	41	78	56	94	35	45	55	64	74
65	94	66	48	39	69	89	98	42	54

obtenha a distribuição de frequência, tendo 30 para limite inferior da primeira classe e 10 para intervalo de classe.

2. Os resultados do lançamento de um dado 50 vezes foram os seguintes:

6	5	2	6	4	3	6	2	6	5
1	6	3	3	5	1	3	6	3	4
5	4	3	1	3	5	4	4	2	6
2	2	5	2	5	1	3	6	5	1
5	6	2	4	6	1	5	2	4	3

Forme uma distribuição de frequência sem intervalos de classe.

3. Considerando as notas de um teste de inteligência aplicado a 100 alunos:

64	78	66	82	74	103	78	86	103	87
73	95	82	89	73	92	85	80	81	90
78	86	78	101	85	98	75	73	90	86
86	84	86	76	76	83	103	86	84	85
76	80	92	102	73	87	70	85	79	93
82	90	83	81	85	72	81	96	81	85
68	96	86	70	72	74	84	99	81	89
71	73	63	105	74	98	78	78	83	96
95	94	88	62	91	83	98	93	83	76
94	75	67	95	108	98	71	92	72	73

Forme uma distribuição de frequência.

4. A tabela abaixo apresenta as vendas diárias de um determinado aparelho elétrico, durante um mês, por uma firma comercial:

14	12	11	13	14	13
12	14	13	14	11	12
12	14	10	13	15	11
15	13	16	17	14	14

Forme uma distribuição de frequência sem intervalos de classe.

5. Complete a tabela abaixo:

i	CLASSES	f_i	fr_i	F_i	Fr_i
1	$0 \vdash 8$	4	....	....	....
2	$8 \vdash 16$	10	....	....	....
3	$16 \vdash 24$	14	....	....	....
4	$24 \vdash 32$	9	....	....	....
5	$32 \vdash 40$	3	....	....	....
		$\Sigma = 40$	$\Sigma = 1{,}00$		

6. Dada a distribuição de frequência:

	3	4	5	6	7	8
	2	5	12	10	8	3

determine:

a. Σf_i;

b. as frequências relativas;

c. as frequências acumuladas;

d. as frequências relativas acumuladas.

7. A tabela abaixo apresenta uma distribuição de frequência das áreas de 400 lotes:

ÁREAS (m^2)	$300 \vdash 400$	$400 \vdash 500$	$500 \vdash 600$	$600 \vdash 700$	$700 \vdash 800$	$800 \vdash 900$	$900 \vdash 1.000$	$1.000 \vdash 1.100$	$1.100 \vdash 1.200$
Nº DE LOTES	14	46	58	76	68	62	48	22	6

Com referência a essa tabela, determine:

a. a amplitude total;

b. o limite superior da quinta classe;

c. o limite inferior da oitava classe;

d. o ponto médio da sétima classe;

e. a amplitude do intervalo da segunda classe;

f. a frequência da quarta classe;

g. a frequência relativa da sexta classe;

h. a frequência acumulada da quinta classe;

i. o número de lotes cuja área não atinge 700 m^2;

j. o número de lotes cuja área atinge e ultrapassa 800 m^2;

l. a percentagem dos lotes cuja área não atinge 600 m^2;

m. a percentagem dos lotes cuja área seja maior ou igual a 900 m^2;

n. a percentagem dos lotes cuja área é de 500 m^2, no mínimo, mas inferior a 1.000 m^2;

o. a classe do 72º lote;

p. até que classe estão incluídos 60% dos lotes.

8. A distribuição abaixo indica o número de acidentes ocorridos com 70 motoristas de uma empresa de ônibus:

	0	1	2	3	4	5	6	7
	20	10	16	9	6	5	3	1

Determine:

a. o número de motoristas que não sofreram nenhum acidente;

b. o número de motoristas que sofreram pelo menos quatro acidentes;

c. o número de motoristas que sofreram menos de três acidentes;

d. o número de motoristas que sofreram no mínimo três e no máximo cinco acidentes;

e. a percentagem dos motoristas que sofreram no máximo dois acidentes.

9. Complete os dados que faltam na distribuição de frequência:

a.

i	x_i	f_i	fr_i	F_i
1	0	1	0,05	....
2	1	....	0,15	4
3	2	4	....	....
4	3	....	0,25	13
5	4	3	0,15	....
6	5	2	....	18
7	6	....	....	19
8	7	....	....	....
		$\Sigma = 20$	$\Sigma = 1{,}00$	

b.

i	CLASSES	x_i	f_i	F_i	fr_i
1	0 ⊢ 2	1	4	....	0,04
2	2 ⊢ 4	....	8	....	....
3	4 ⊢ 6	5	....	30	0,18
4	....	7	27	....	0,27
5	8 ⊢ 10	....	15	72	....
6	10 ⊢ 12	....	....	83	....
7	....	13	10	93	0,10
8	14 ⊢ 16	....	....	....	0,07
			$\Sigma = ...$		$\Sigma = ...$

5.7 Representação gráfica de uma distribuição

Uma distribuição de frequência pode ser representada graficamente pelo **histograma**, pelo **polígono de frequência** e pelo **polígono de frequência acumulada**[6].

Construímos qualquer um dos gráficos mencionados utilizando o primeiro quadrante do sistema de eixos coordenados cartesianos ortogonais. Na linha horizontal (eixo das abscissas) colocamos os valores da variável e na linha vertical (eixo das ordenadas), as frequências.

5.7.1 Histograma

O **histograma** é formado por um conjunto de retângulos justapostos, cujas bases se localizam sobre o eixo horizontal, de tal modo que seus pontos médios coincidam com os pontos médios dos intervalos de classe.

[6] Alguns autores preferem designá-lo por **ogiva de Galton**.

As larguras dos retângulos são iguais às amplitudes dos intervalos de classe.

As alturas dos retângulos devem ser proporcionais às frequências das classes, sendo a amplitude dos intervalos igual. Isso nos permite tomar as alturas numericamente iguais às frequências.

À distribuição da Tabela 5.6 (p. 52) corresponde o seguinte histograma:

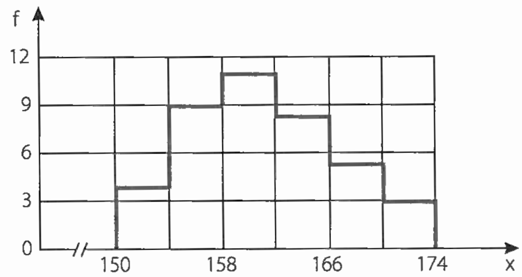


FIGURA 5.1

NOTAS:

- O histograma goza de uma propriedade da qual faremos considerável uso: **a área de um histograma é proporcional à soma das frequências.**
- No caso de usarmos as frequências relativas, obtemos um gráfico de área unitária.
- Quando queremos comparar duas distribuições, o ideal é fazê-lo pelo histograma de frequências relativas.

5.7.2 Polígono de frequência

O **polígono de frequência** é um gráfico em linha, sendo as frequências marcadas sobre perpendiculares ao eixo horizontal, levantadas pelos pontos médios dos intervalos de classe.

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.

À distribuição da Tabela 5.6 corresponde o seguinte polígono de frequência:

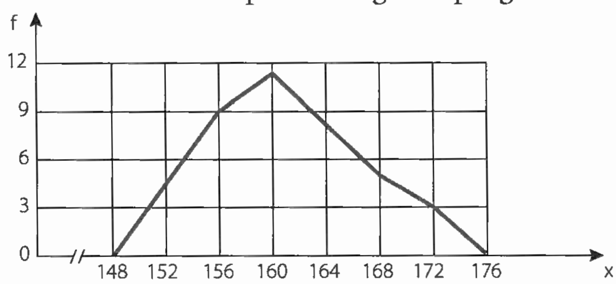


FIGURA 5.2

NOTA:

- No caso de termos uma variável essencialmente positiva, cuja distribuição se inicie no valor zero, devemos considerar um intervalo anterior localizado no semieixo negativo. Consideraremos, porém, apenas a parte positiva do segmento que liga o ponto médio desse intervalo com a frequência do intervalo 0 ⊢ ..... .

Exemplo:

i	CLASSES	f_i
1	0 ⊢ 2	1
2	2 ⊢ 4	2
3	4 ⊢ 6	4
4	6 ⊢ 8	3
5	8 ⊢ 10	1
		$\Sigma = 11$

TABELA 5.12

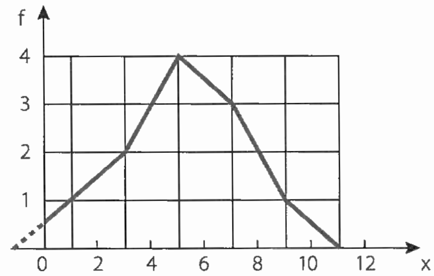


FIGURA 5.3

5.7.3 Polígono de frequência acumulada

O **polígono de frequência acumulada** é traçado marcando-se as frequências acumuladas sobre perpendiculares ao eixo horizontal, levantadas nos pontos correspondentes aos limites superiores dos intervalos de classe.

Assim, à distribuição da Tabela 5.6 corresponde o seguinte polígono de frequência acumulada:

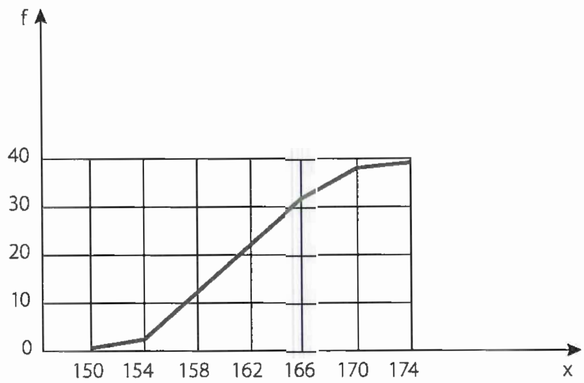


FIGURA 5.4

Uma distribuição de frequência sem intervalos de classe é representada graficamente por um diagrama onde cada valor da variável é representado por um segmento de reta vertical e de comprimento proporcional à respectiva frequência. Assim, para a distribuição da Tabela 5.13, temos:

i	x_i	f_i	F_i
1	2	4	4
2	3	7	11
3	4	5	16
4	5	2	18
5	6	1	19
6	7	1	20
		$\Sigma = 20$	

TABELA 5.13

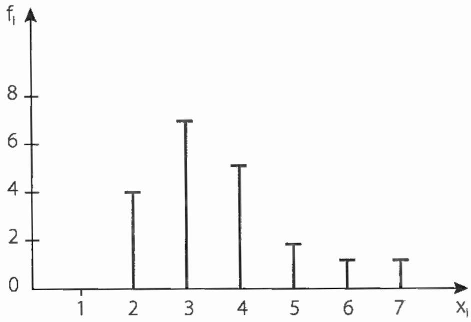


FIGURA 5.5

Também podemos representar a distribuição pelo gráfico da frequência acumulada, o qual se apresentará com pontos de descontinuidade nos valores observados da variável:

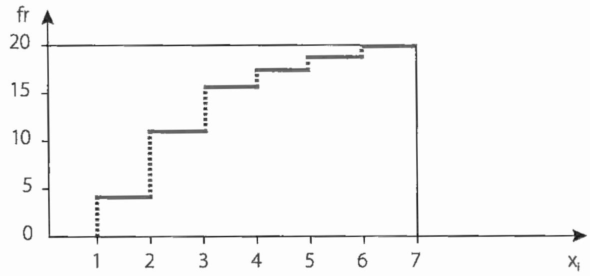


FIGURA 5.6

5.8 A curva de frequência

5.8.1 A curva de frequência. Curva polida

Como, em geral, os dados coletados pertencem a uma amostra extraída de uma população, podemos imaginar as amostras tornando-se cada vez mais amplas e a amplitude das classes ficando cada vez menor, o que nos permite concluir que a linha poligonal (contorno do polígono de frequência) tende a se transformar numa curva — **a curva de frequência** —, mostrando, de modo mais evidente, a verdadeira natureza da distribuição da população.

Podemos dizer, então, que, enquanto o polígono de frequência nos dá a **imagem real** do fenômeno estudado, a curva de frequência nos dá a **imagem tendencial**.

Assim, após o traçado de um polígono de frequência, é desejável, muitas vezes, que se lhe faça um **polimento**, de modo a mostrar o que seria tal polígono com um número maior de dados.

Esse procedimento, é claro, não nos dará uma certeza absoluta de que a curva obtida — **curva polida** — seja tal qual a curva resultante de um grande número de dados. Podemos, porém, afirmar que ela se assemelha mais à curva de frequência do que ao polígono de frequência obtido de uma amostra limitada.

O polimento, geometricamente, corresponde à eliminação dos vértices da linha poligonal. Consegue-se isso com o emprego de uma fórmula bastante simples, a qual, a partir das frequências reais, nos fornece novas frequências — **frequências calculadas** — que se localizarão, como no polígono de frequência, nos pontos médios.

A fórmula que nos dá a **frequência calculada** (fc_i) é:

$$fc_i = \frac{f_{i-1} + 2f_i + f_{i+1}}{4}$$

onde:

fc_i é a frequência calculada da classe considerada;

f_i é a frequência simples da classe considerada;

f_{i-1} é a frequência simples da classe anterior à classe considerada;

f_{i+1} é a frequência simples da classe posterior à classe considerada.

Quando fazemos uso da curva polida, convém mostrar as frequências realmente observadas por meio de pontos ou pequenos círculos, de modo que qualquer interessado possa, por si mesmo, julgar até que ponto os dados originais foram polidos.

Para a distribuição da Tabela 5.6, temos:

$$fc_1 = \frac{0 + 2 \times 4 + 9}{4} = \frac{17}{4} = 4{,}25 \qquad fc_4 = \frac{11 + 2 \times 8 + 5}{4} = \frac{32}{4} = 8$$

$$fc_2 = \frac{4 + 2 \times 9 + 11}{4} = \frac{33}{4} = 8{,}25 \qquad fc_5 = \frac{8 + 2 \times 5 + 3}{4} = \frac{21}{4} = 5{,}25$$

$$fc_3 = \frac{9 + 2 \times 11 + 8}{4} = \frac{39}{4} = 9{,}75 \qquad fc_6 = \frac{5 + 2 \times 3 + 0}{4} = \frac{11}{4} = 2{,}75$$

i	ESTATURAS (cm)	f_i	fc_i
1	150 ⊢ 154	4	4,2
2	154 ⊢ 158	9	8,2
3	158 ⊢ 162	11	9,8
4	162 ⊢ 166	8	8,0
5	166 ⊢ 170	5	5,2
6	170 ⊢ 174	3	2,8
		$\Sigma = 40$	

TABELA 5.14

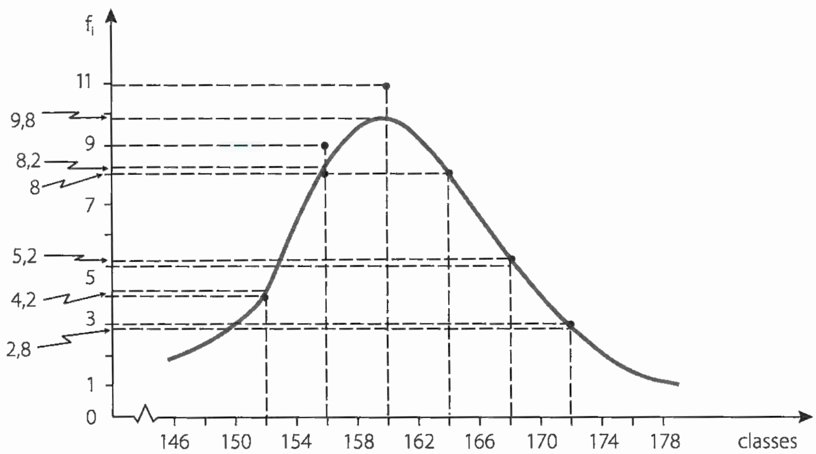


FIGURA 5.7

5.8.2 As formas das curvas de frequência

As curvas de frequência assumem as seguintes formas características:

Curvas em forma de sino

As **curvas em forma de sino** caracterizam-se pelo fato de apresentarem um valor máximo na região central.

São muitos os fenômenos que oferecem distribuições em forma de sino: a estatura de adultos, o peso de adultos, a inteligência medida em testes mentais, os preços relativos.

Distinguimos a curva em forma de sino **simétrica** e a **assimétrica.**

- **Curva simétrica**

Esta curva caracteriza-se por apresentar o valor máximo no ponto central e os pontos equidistantes desse ponto terem a mesma frequência (Figura 5.8).

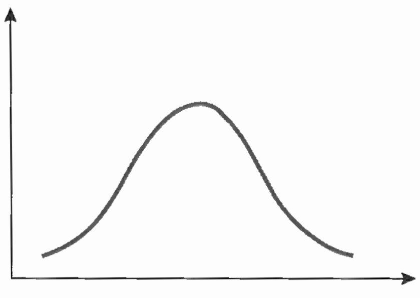


FIGURA 5.8

- **Curva assimétrica**

Na prática, não encontramos distribuições perfeitamente simétricas. As distribuições obtidas de medições reais são mais ou menos assimétricas, em relação à frequência máxima. Assim, as curvas correspondentes a tais distribuições apresentam a **cauda** de um lado da ordenada máxima mais longa do que do outro. Se a cauda mais alongada fica à direita, a curva é chamada **assimétrica positiva** ou **enviesada à direita** (Figura 5.9). Se a cauda se alonga à esquerda, a curva é chamada **assimétrica negativa** ou **enviesada à esquerda** (Figura 5.10).

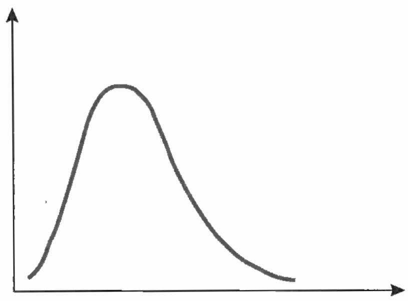


FIGURA 5.9

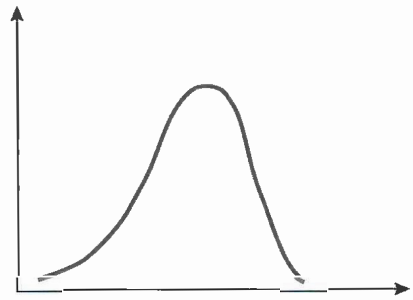


FIGURA 5.10

Curvas em forma de jota

As **curvas em forma de jota** são relativas a distribuições extremamente assimétricas, caracterizadas por apresentarem o ponto de ordenada máxima em uma das extremidades.

São curvas comuns aos fenômenos econômicos e financeiros: distribuição de vencimentos ou rendas pessoais (Figuras 5.11 e 5.12).

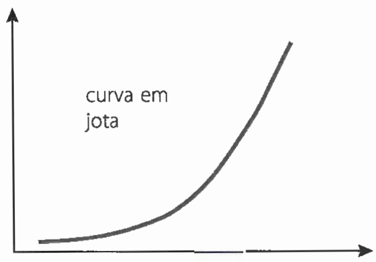


FIGURA 5.11

FIGURA 5.12

Curvas em forma de U

As **curvas em forma de U** são caracterizadas por apresentarem ordenadas máximas em ambas as extremidades.

Como exemplo de distribuição que dá origem a esse tipo de curva podemos citar a de mortalidade por idade (Figura 5.13).

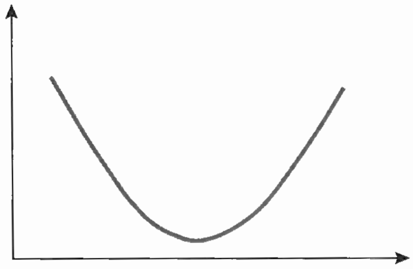


FIGURA 5.13

Distribuição retangular

Essa distribuição, muito rara na verdade, apresenta todas as classes com a mesma frequência. Tal distribuição seria representada por um histograma em que todas as colunas teriam a mesma altura (Figura 5.14) ou por um polígono de frequência reduzido a um segmento de reta horizontal (Figura 5.15).

FIGURA 5.14

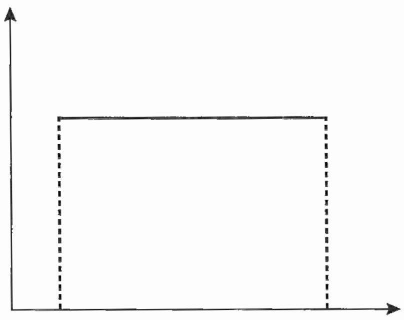


FIGURA 5.15

Exercícios

1. Considerando as distribuições de frequência seguintes, confeccione, para cada uma:
   a. o histograma;
   b. o polígono de frequência;
   c. o polígono de frequência acumulada.

I.

i	PESOS (kg)	f_i
1	40 ⊢ 44	2
2	44 ⊢ 48	5
3	48 ⊢ 52	9
4	52 ⊢ 56	6
5	56 ⊢ 60	4
		$\Sigma = 26$

II.

i	ESTATURAS (cm)	f_i
1	150 ⊢ 156	1
2	156 ⊢ 162	5
3	162 ⊢ 168	8
4	168 ⊢ 174	13
5	174 ⊢ 180	3
		$\Sigma = 30$

III.

i	SALÁRIOS (R$)	f_i
1	500 ⊢ 700	8
2	700 ⊢ 900	20
3	900 ⊢ 1.100	7
4	1.100 ⊢ 1.300	5
5	1.300 ⊢ 1.500	2
6	1.500 ⊢ 1.700	1
7	1.700 ⊢ 1.900	1
		$\Sigma = 44$

2. Confeccione o gráfico da distribuição:

ÁREAS (m^2)	300 ⊢ 400	400 ⊢ 500	500 ⊢ 600	600 ⊢ 700	700 ⊢ 800	800 ⊢ 900	900 ⊢ 1.000	1.000 ⊢ 1.100	1.100 ⊢ 1.200
Nº DE LOTES	14	46	58	76	68	62	48	22	6

3. Confeccione a curva polida relativa à distribuição de frequência:

i	CLASSES	f_i
1	4 ⊢ 8	2
2	8 ⊢ 12	5
3	12 ⊢ 16	9
4	16 ⊢ 20	6
5	20 ⊢ 24	2
6	24 ⊢ 28	1
		$\Sigma = 25$

4. Examinando o histograma abaixo, que corresponde às notas relativas à aplicação de um teste de inteligência a um grupo de alunos, responda:

a. Qual é o intervalo de classe que tem maior frequência?

b. Qual a amplitude total da distribuição?

c. Qual o número total de alunos?

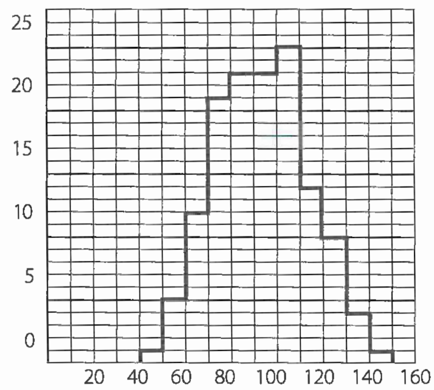
d. Qual é a frequência do intervalo de classe 110 ⊢ 120?

e. Quais os dois intervalos de classe que têm a mesma frequência?

f. Quais são os dois intervalos de classe tais que a frequência de um é o dobro da frequência do outro?

g. Quantos alunos receberam notas de teste entre 90 (inclusive) e 110?

h. Quantos alunos receberam notas não inferiores a 100?



5. Cite o tipo de curva correspondente a cada distribuição a seguir:

a. Número de mulheres de 15 a 30 anos, em uma dada população, casadas, classificadas segundo o número de vezes que hajam contraído matrimônio.

b. Notas de alunos que cursam a última série do 2º grau, em uma dada população.

c. Coeficientes de mortalidade por acidente, por grupo de idade.

d. Tempo de estacionamento de veículos motorizados em uma área de congestionamento.

e. Número de homens capacitados, por grupo de idade, que estão desempregados em uma determinada época.

6. Conhecidas as notas de 50 alunos:

68	85	33	52	65	77	84	65	74	57
71	35	81	50	35	64	74	47	54	68
80	61	41	91	55	73	59	53	77	45
41	55	78	48	69	85	67	39	60	76
94	98	66	66	73	42	65	94	88	89

determine:

a. a distribuição de frequência começando por 30 e adotando o intervalo de classe de amplitude igual a 10;

b. as frequências acumuladas;

c. as frequências relativas;

d. o histograma e o polígono de frequência.

7. A tabela abaixo apresenta os coeficientes de liquidez obtidos da análise de balanço em 50 indústrias:

3,9	7,4	10,0	11,8	2,3	4,5	10,5	8,4	15,6	7,6
18,8	2,9	2,3	0,4	5,0	9,0	5,5	9,2	12,4	8,7
4,5	4,4	10,6	5,6	8,5	2,4	17,8	11,6	0,8	4,4
7,1	3,2	2,7	16,2	2,7	9,5	13,1	3,8	6,3	7,9
4,8	5,3	12,9	6,9	6,3	7,5	2,6	3,3	4,6	16,0

a. Forme com esses dados uma distribuição com intervalos de classe iguais a 3, tais que os limites inferiores sejam múltiplos de 3.

b. Confeccione o histograma e o polígono de frequência correspondentes.

8. Um grau de nebulosidade, registrado em décimos, ocorre de acordo com a distribuição abaixo:

[redacted]	0 ⊢ 0,5	0,5 ⊢ 1,5	1,5 ⊢ 2,5	2,5 ⊢ 3,5	3,5 ⊢ 4,5	4,5 ⊢ 5,5	5,5 ⊢ 6,5	6,5 ⊢ 7,5	7,5 ⊢ 8,5	8,5 ⊢ 9,5	9,5 ⊢ 10,0
	320	125	75	65	45	45	55	65	90	145	676

Construa o histograma correspondente.

9. Considerando a distribuição abaixo:

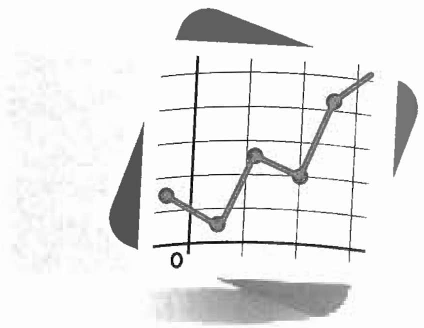
[redacted]	1 ⊢ 2	2 ⊢ 3	3 ⊢ 4	4 ⊢ 5	5 ⊢ 6	6 ⊢ 7	7 ⊢ 8	8 ⊢ 9	9 ⊢ 10	10 ⊢ 11	11 ⊢ 12	12 ⊢ 13	13 ⊢ 14	14 ⊢ 15	15 ⊢ 16	16 ⊢ 17	17 ⊢ 18
	7	3	10	11	12	37	35	45	39	30	25	7	10	4	6	1	4

confeccione:

a. um histograma;

b. um polígono de frequência;

c. a curva polida, indicando as frequências reais por meio de pequenos círculos.



6 MEDIDAS DE POSIÇÃO[1]

6.1 Introdução

O estudo que fizemos sobre distribuições de frequência, até agora, permite-nos descrever, de modo geral, os grupos dos valores que uma variável pode assumir. Dessa forma, podemos localizar a **maior concentração** de valores de uma dada distribuição, isto é, se ela se localiza no início, no meio ou no final, ou, ainda, se há uma distribuição por igual.

Para ressaltar, porém, as **tendências características** de cada distribuição, isoladamente, ou em confronto com outras, necessitamos introduzir conceitos que se expressem através de números, que nos permitam traduzir essas tendências. Esses conceitos são denominados **elementos típicos da distribuição** e são as:

a. medidas de posição;

b. medidas de variabilidade ou dispersão;

c. medidas de assimetria;

d. medidas de curtose.

---

[1] Consulte o **Apêndice — Instrumental Matemático**, para uma revisão do assunto **Média Aritmética** (p. 187).

Dentre os elementos típicos, destacamos, neste capítulo, as **medidas de posição** — estatísticas que representam uma série de dados orientando-nos quanto à posição da distribuição em relação ao eixo horizontal (eixo das abscissas).

As medidas de posição mais importantes são as **medidas de tendência central**, que recebem tal denominação pelo fato de os dados observados tenderem, em geral, a se agrupar em torno dos valores centrais. Dentre as medidas de tendência central, destacamos:

a. a média aritmética;

b. a mediana;

c. a moda.

As outras medidas de posição são as **separatrizes**, que englobam:

a. a própria mediana;

b. os quartis;

c. os percentis.

6.2 Média aritmética ($\bar{x}$)

Em um conjunto de dados, podemos definir vários tipos de médias. No entanto, em nossos estudos iremos nos limitar à mais importante: a **média aritmética**.

Média aritmética é o quociente da divisão da soma dos valores da variável pelo número deles:

$$\bar{x} = \frac{\sum x_i}{n}$$

sendo:
$\bar{x}$ a média aritmética;
x_i os valores da variável;
n o número de valores.

6.2.1 Dados não agrupados

Quando desejamos conhecer a média dos dados não agrupados, determinamos a **média aritmética simples**.

Exemplo:

Sabendo-se que a produção leiteira diária da vaca **A**, durante uma semana, foi de 10, 14, 13, 15, 16, 18 e 12 litros, temos, para produção média da semana:

$$\bar{x} = \frac{10+14+13+15+16+18+12}{7} = \frac{98}{7} = 14$$

Logo:

$$\bar{x} = 14 \text{ litros}$$

Às vezes, a média pode ser um número diferente de todos os da série de dados que ela representa. É o que acontece quando temos os valores 2, 4, 8 e 9, para os quais a média é **5**. Esse será o número representativo dessa série de valores, embora não esteja representado nos dados originais. Neste caso, costumamos dizer que a média **não tem existência concreta**.

6.2.2 Desvio em relação à média

Denominamos **desvio em relação à média** a diferença entre cada elemento de um conjunto de valores e a média aritmética.

Designando o desvio por $\mathbf{d_i}$, temos:

$$d_i = x_i - \bar{x}$$

Para o exemplo dado, temos:

$d_1 = x_1 - \bar{x} \Rightarrow d_1 = 10 - 14 = -4$

$d_2 = x_2 - \bar{x} \Rightarrow d_2 = 14 - 14 = 0$

$d_3 = x_3 - \bar{x} \Rightarrow d_3 = 13 - 14 = -1$

$d_4 = x_4 - \bar{x} \Rightarrow d_4 = 15 - 14 = 1$

$d_5 = x_5 - \bar{x} \Rightarrow d_5 = 16 - 14 = 2$

$d_6 = x_6 - \bar{x} \Rightarrow d_6 = 18 - 14 = 4$

$d_7 = x_7 - \bar{x} \Rightarrow d_7 = 12 - 14 = -2$

6.2.3 Propriedades da média

1ª propriedade

A soma algébrica dos desvios tomados em relação à média é nula:

$$\sum_{i=1}^{k} d_i = 0$$

No exemplo anterior, temos:

$$\sum_{i=1}^{7} d_i = (-4) + 0 + (-1) + 1 + 2 + 4 + (-2) = (-7) + 7 = 0 \Rightarrow$$

$$\Rightarrow \sum_{i=1}^{7} d_i = 0$$

2ª propriedade

Somando-se (ou subtraindo-se) uma constante (**c**) de todos os valores de uma variável, a média do conjunto fica aumentada (ou diminuída) dessa constante:

$$y_i = x_i \pm c \Rightarrow \bar{y} = \bar{x} \pm c$$

Somando 2 a cada um dos valores da variável do exemplo dado, temos:

$$y_1 = 12, y_2 = 16, y_3 = 15, y_4 = 17, y_5 = 18, y_6 = 20 \text{ e } y_7 = 14$$

Daí:

$$\sum_{i=1}^{7} y_1 = 12 + 16 + 15 + 17 + 18 + 20 + 14 = 112$$

Como $n = 7$, vem:

$$\bar{y} = \frac{112}{7} = 16 \Rightarrow \bar{y} = 16 = 14 + 2 \Rightarrow \bar{y} = \bar{x} + 2$$

3ª propriedade

Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante (**c**), a média do conjunto fica multiplicada (ou dividida) por essa constante:

$$y_i = x_i \times c \Rightarrow \bar{y} = \bar{x} \times c \quad \text{ou} \quad y_i = \frac{x_i}{c} \Rightarrow \bar{y} = \frac{\bar{x}}{c}$$

Multiplicando por 3 cada um dos valores da variável do exemplo dado, obtemos:

$$y_1 = 30, y_2 = 42, y_3 = 39, y_4 = 45, y_5 = 48, y_6 = 54 \text{ e } y_7 = 36$$

Daí:

$$\sum_{i=1}^{7} y_i = 30 + 42 + 39 + 45 + 48 + 54 + 36 = 294$$

Como n = 7, temos:

$$\bar{y} = \frac{294}{7} = 42 \Rightarrow \bar{y} = 42 = 14 \times 3 \Rightarrow \bar{y} = \bar{x} \times 3$$

6.2.4 Dados agrupados

Sem intervalos de classe

Consideremos a distribuição relativa a 34 famílias de quatro filhos, tomando para variável o número de filhos do sexo masculino:

Nº DE MENINOS	f_i
0	2
1	6
2	10
3	12
4	4
	$\Sigma = 34$

TABELA 6.1

Neste caso, como as frequências são números indicadores da intensidade de cada valor da variável, elas funcionam como fatores de ponderação, o que nos leva a calcular a **média aritmética ponderada**, dada pela fórmula:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$$

O modo mais prático de obtenção da média ponderada é abrir, na tabela, uma coluna correspondente aos produtos $x_i f_i$:

x_i	f_i	$x_i f_i$
0	2	0
1	6	6
2	10	20
3	12	36
4	4	16
	$\Sigma = 34$	$\Sigma = 78$

TABELA 6.2

Temos, então:

$$\sum x_i f_i = 78 \quad \text{e} \quad \sum f_i = 34$$

Logo:

$$\overline{x} = \frac{\sum x_i f_i}{\sum f_i} = \overline{x} = \frac{78}{34} = 2,29 \Rightarrow \overline{x} = 2,3$$

isto é:

$$\overline{x} = 2,3 \text{ meninos}$$

NOTA:

- Sendo **x** uma variável discreta, como interpretar o resultado obtido, 2 meninos e 3 décimos de menino?

  O valor médio 2,3 meninos sugere, neste caso, que o maior número de famílias tem 2 meninos e 2 meninas, sendo, porém, a tendência geral de uma leve superioridade numérica em relação ao número de meninos.

Resolva

1. Complete o esquema para o cálculo da média aritmética da distribuição:

	1	2	3	4	5	6
	2	4	6	8	3	1

Temos:

x_i	f_i	$x_i f_i$
1	2	2
2	4	....
3	6	....
4	8	....
5	3	....
6	1	....
	$\Sigma = ...$	$\Sigma = ...$

Como:

$$\sum f_i = ..., \sum x_i f_i = ...$$

e

$$\overline{x} = \frac{\sum x_i f_i}{\sum f_i},$$

temos:

$$\overline{x} = \frac{....}{....} = ... \Rightarrow \overline{x} = 3,4$$

Com intervalos de classe

Neste caso, convencionamos que **todos os valores incluídos em um determinado intervalo de classe coincidem com o seu ponto médio**, e determinamos a média aritmética ponderada por meio da fórmula:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$$

onde $\mathbf{x_i}$ é o ponto médio da classe.

Consideremos a distribuição:

i	ESTATURAS (cm)	f_i
1	150 ⊢ 154	4
2	154 ⊢ 158	9
3	158 ⊢ 162	11
4	162 ⊢ 166	8
5	166 ⊢ 170	5
6	170 ⊢ 174	3
		$\Sigma = 40$

TABELA 6.3

Pela mesma razão do caso anterior, vamos, inicialmente, abrir uma coluna para os pontos médios e outra para os produtos $x_i f_i$:

i	ESTATURAS (cm)	f_i	x_i	$x_i f_i$
1	150 ⊢ 154	4	152	608
2	154 ⊢ 158	9	156	1.404
3	158 ⊢ 162	11	160	1.760
4	162 ⊢ 166	8	164	1.312
5	166 ⊢ 170	5	168	840
6	170 ⊢ 174	3	172	516
		$\Sigma = 40$		$\Sigma = 6.440$

TABELA 6.4

Como, neste caso:

$$\sum x_i f_i = 6.440,\ \sum f_i = 40 \text{ e } \bar{x} = \frac{\sum x_i f_i}{\sum f_i},$$

temos:

$$\overline{x} = \frac{6.440}{40} = 161 \Rightarrow \overline{x} = 161 \text{ cm}$$

Resolva

1. Complete o esquema para o cálculo da média aritmética da distribuição de frequência:

CUSTO (R$)	450 ⊢ 550	550 ⊢ 650	650 ⊢ 750	750 ⊢ 850	850 ⊢ 950	950 ⊢ 1.050	1.050 ⊢ 1.150
f_i	8	10	11	16	13	5	1

Temos:

i	x_i	f_i	$x_i f_i$
1	500	8	4.000
2	....	10	....
3	....	11	....
4	....	16	....
5	....	13	....
6	....	5	....
7	1.100	1	....
		Σ = ...	Σ = ...

Logo:

$$x = \frac{....}{....} = ...,$$

donde:

$$x = R\$\ 755$$

Processo breve

Com o intuito de eliminarmos o grande número de cálculos que às vezes se apresentam na determinação da média, empregamos o que denominamos **processo breve** (em oposição ao processo usado anteriormente — **processo longo**), baseado em uma mudança da variável **x** por outra **y**, tal que:

$$y_i = \frac{x_i - x_0}{h}$$

onde $\mathbf{x_0}$ é uma constante arbitrária escolhida convenientemente dentre os pontos médios da distribuição — de preferência o de maior frequência.

h → AMPLITUDE DO INTERVALO DA CLASSE.

Fazendo essa mudança de variável, de acordo com a segunda e a terceira propriedades da média, ela resulta diminuída de $\mathbf{x_0}$ e dividida por **h**; mas isso pode ser compensado somando $\mathbf{x_0}$ à média da nova variável e, ao mesmo tempo, multiplicando-a por **h**. Resulta, então, a fórmula modificada:

$$\bar{x} = x_0 + \frac{\left(\sum y_i f_i\right) \times h}{\sum f_i}$$

Assim, para a distribuição da Tabela 6.3, tomando para o valor de x_0 o ponto médio de maior frequência (se bem que podemos tomar qualquer dos valores do ponto médio), isto é:

$$x_0 = 160$$

como h = 4, temos para valores da nova variável:

$$y_1 = \frac{152 - 160}{4} = \frac{-8}{4} = -2 \qquad y_4 = \frac{164 - 160}{4} = \frac{4}{4} = 1$$

$$y_2 = \frac{156 - 160}{4} = \frac{-4}{4} = -1 \qquad y_5 = \frac{168 - 160}{4} = \frac{8}{4} = 2$$

$$y_3 = \frac{160 - 160}{4} = \frac{0}{4} = 0 \qquad y_6 = \frac{172 - 160}{4} = \frac{12}{4} = 3$$

Vamos, então, calcular a média da distribuição da Tabela 6.3 pelo **processo breve**.

Começamos por completar a tabela dada com as colunas correspondentes aos pontos médios $(\mathbf{x_i})$, aos valores da nova variável $(\mathbf{y_i})$ e aos produtos $\mathbf{y_i f_i}$:

i	ESTATURAS (cm)	f_i	x_i	y_i	$y_i f_i$	
1	150 ⊢ 154	4	152	−2	−8	
2	154 ⊢ 158	9	156	−1	−9	−17
3	158 ⊢ 162	11	160	0	0	
4	162 ⊢ 166	8	164	1	8	
5	166 ⊢ 170	5	168	2	10	
6	170 ⊢ 174	3	172	3	9	27
	$x_0 = 160$	$\Sigma = 40$			$\Sigma = 10$	

TABELA 6.5

Temos, então, $x_0 = 160$, $\sum y_i f_i = 10$, $\sum f_i = 40$ e h = 4.

Substituindo esses valores na fórmula:

$$\bar{x} = x_0 + \frac{\left(\sum y_i f_i\right) h}{\sum f_i}$$

vem:

$$\bar{x} = 160 + \frac{10 \times 4}{40} = 160 + 1 \Rightarrow x = 161,$$

donde:

$$\bar{x} = 161 \text{ cm}$$

NOTAS:

- O processo breve, com a nova variável definida por nós, só pode ser usado em distribuições que apresentam intervalos de classe de mesma amplitude.
- O processo breve pode, também, ser aplicado para as distribuições sem intervalos de classe, bastando fazer **h = 1**.

Fases para o cálculo da média pelo processo breve:

1ª) Abrimos uma coluna para os valores x_i.

2ª) Escolhemos um dos pontos médios (de preferência o de maior frequência) para o valor de x_0.

3ª) Abrimos uma coluna para os valores de y_i e escrevemos **zero** na linha correspondente à classe onde se encontra o valor de x_0; a sequência **–1, –2, –3, ...,** logo acima do zero, e a sequência **1, 2, 3, ...,** logo abaixo.

4ª) Abrimos uma coluna para os valores do produto $y_i f_i$, conservando os sinais + ou –, e, em seguida, somamos algebricamente esses produtos.

5ª) Aplicamos a fórmula.

Exercício resolvido

1. Calcule a média aritmética, pelo processo breve, da distribuição:

CUSTOS (R$)	450 ⊢ 550	550 ⊢ 650	650 ⊢ 750	750 ⊢ 850	850 ⊢ 950	950 ⊢ 1.050	1.050 ⊢ 1.150
f_i	8	10	11	16	13	5	1

Temos:

i	x_i	f_i	y_i	y_if_i	
1	500	8	–3	–24	
2	600	10	–2	–20	
3	700	11	–1	–11	–55
4	800	16	0	0	
5	900	13	1	13	
6	1.000	5	2	10	
7	1.100	1	3	3	26
$x_0 = 800$		$\Sigma = 64$		$\Sigma = -29$	

Como:

$$h = 100$$

vem:

$$\bar{x} = 800 + \frac{(-29)\,100}{64} = 800 - \frac{2.900}{64} = 800 - 45{,}31 = 754{,}69$$

$\bar{x}$ = R$ 755 é a resposta.

Resolva

1. Complete o esquema para o cálculo da média aritmética da distribuição:

	30 ⊢ 50	50 ⊢ 70	70 ⊢ 90	90 ⊢ 110	110 ⊢ 130
	2	8	12	10	5

Temos:

i	x_i	f_i	y_i	y_if_i
1	40	....	....	....
2	....	....	....	....
3	....	12	....	....
4	....	....	....	....
5	....	....	2	....
$x_0 = ...$		$\Sigma = ...$		$\Sigma = ...$

Como:

h = ...

vem:

$$\bar{x} = ... + \frac{.... \times ....}{....} = ... + ... = ... \Rightarrow \bar{x} = 84{,}3$$

6.2.5 Emprego da média

A **média** é utilizada quando:

a. desejamos obter a medida de posição que possui a maior estabilidade;
b. houver necessidade de um tratamento algébrico ulterior.

6.3 A moda (Mo)

Denominamos **moda** o valor que ocorre com maior frequência em uma série de valores.

Desse modo, o **salário modal** dos empregados de uma indústria é o salário mais comum, isto é, o salário recebido pelo maior número de empregados dessa indústria.

6.3.1 Dados não agrupados

Quando lidamos com valores não agrupados, a moda é facilmente reconhecida: basta, de acordo com a definição, procurar o valor que mais se repete.

A série de dados:

7, 8, 9, 10, 10, 10, 11, 12, 13, 15

tem moda igual a 10.

Podemos, entretanto, encontrar séries nas quais não exista valor modal, isto é, nas quais nenhum valor apareça mais vezes que outros. É o caso da série:

3, 5, 8, 10, 12, 13,

que não apresenta moda (**amodal**).

Em outros casos, ao contrário, pode haver dois ou mais valores de concentração. Dizemos, então, que a série tem dois ou mais valores modais. Na série:

2, 3, 4, 4, 4, 5, 6, 7, 7, 7, 8, 9

temos duas modas: 4 e 7 (**bimodal**).

6.3.2 Dados agrupados

Sem intervalos de classe

Uma vez agrupados os dados, é possível determinar imediatamente a moda: basta fixar o valor da variável de maior frequência.

Na distribuição da Tabela 6.1, à frequência máxima (**12**) corresponde o valor **3** da variável. Logo:

$$Mo = 3$$

Com intervalos de classe

A classe que apresenta a maior frequência é denominada **classe modal**. Pela definição, podemos afirmar que a moda, neste caso, é o valor dominante que está compreendido entre os limites da classe modal.

O método mais simples para o cálculo da moda consiste em tomar o ponto médio da classe modal.

Damos a esse valor a denominação de **moda bruta**.

Temos, então:

$$Mo = \frac{\ell^* + L^*}{2}$$

onde:

ℓ^* é o limite inferior da classe modal;

L* é o limite superior da classe modal.

Assim, para a distribuição:

i	ESTATURAS (cm)	f_i
1	150 ⊢ 154	4
2	154 ⊢ 158	9
3	158 ⊢ 162	11 ←
4	162 ⊢ 166	8
5	166 ⊢ 170	5
6	170 ⊢ 174	3
		$\Sigma = 40$

TABELA 6.6

temos que a classe modal é i = 3, $\ell^* = 158$ e $L^* = 162$.

Como:

$$Mo = \frac{\ell^* + L^*}{2}$$

vem:

$$Mo = \frac{158 + 162}{2} = \frac{320}{2} = 160$$

Logo:

$$Mo = 160 \text{ cm}$$

NOTA:

- Há, para o cálculo da moda, outros métodos mais elaborados, como, por exemplo, o que faz uso da **fórmula de Czuber**:

$$Mo = \ell^* + \frac{D_1}{D_1 + D_2} \times h^*$$

na qual:

ℓ^* é o limite inferior da classe modal;

h* é a amplitude da classe modal;

$D_1 = f^* - f(ant)$;

$D_2 = f^* - f(post)$,

sendo:

f* a frequência simples da classe modal;

f(ant) a frequência simples da classe anterior à classe modal;

f(post) a frequência simples da classe posterior à classe modal.

Assim, para a distribuição da Tabela 6.6, temos:

$$D_1 = 11 - 9 = 2 \quad \text{e} \quad D_2 = 11 - 8 = 3$$

donde:

$$Mo = 158 + \frac{2}{2+3} \times 4 = 158 + \frac{2 \times 4}{2+3} = 158 + \frac{8}{5} = 158 + 1{,}6 = 159{,}6$$

Logo:

$$Mo = 159{,}6 \text{ cm}$$

Resolva

1. Complete o esquema para o cálculo da moda da distribuição de frequência:

i	CUSTOS (R$)	f_i
1	450 ⊢ 550	8
2	550 ⊢ 650	10
3	650 ⊢ 750	11
4	750 ⊢ 850	16
5	850 ⊢ 950	13
6	1.950 ⊢ 1.050	5
7	1.050 ⊢ 1.150	1
		$\Sigma = 64$

A classe modal é a de ordem...

Logo:

$$\ell^* = \ldots \text{ e } L^* = \ldots$$

Temos, pois:

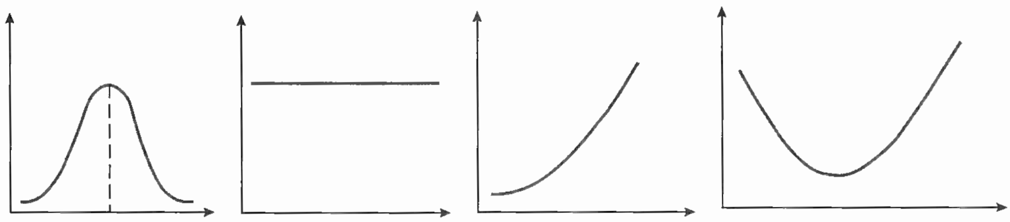
$$Mo = \frac{\ldots. + \ldots.}{2} = \frac{\ldots.}{2} = \ldots.,$$

isto é:

$$Mo = R\$\ 800$$

6.3.3 As expressões gráficas da moda

Na curva de frequência, a **moda** é o valor que corresponde, no eixo das abscissas, ao ponto de ordenada máxima. Assim, podemos ter:

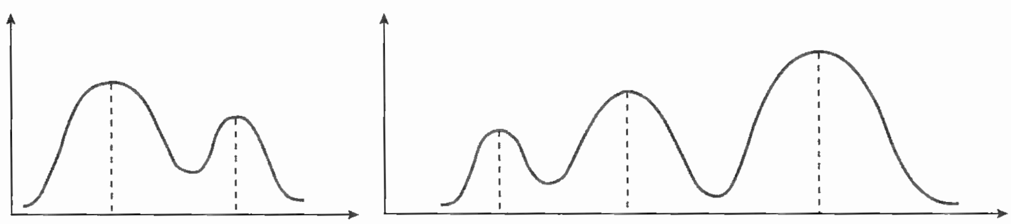


Mo
CURVA MODAL

CURVA NÃO MODAL

CURVA AMODAL

CURVA ANTIMODAL



Mo_1 Mo_2
CURVA BIMODAL

Mo_1 Mo_2 Mo_3
CURVA TRIMODAL

6.3.4 Emprego da moda

A **moda** é utilizada:

a. quando desejamos obter uma medida rápida e aproximada de posição;

b. quando a medida de posição deve ser o valor mais típico da distribuição.

6.4 A mediana (Md)

A **mediana** é outra medida de posição definida como o número que se encontra no centro de uma série de números, estando estes dispostos segundo uma ordem. Em outras palavras, a mediana de um conjunto de valores, ordenados segundo uma ordem de grandeza, é o valor situado de tal forma no conjunto que o separa em dois subconjuntos de mesmo número de elementos.

6.4.1 Dados não agrupados

Dada uma série de valores, como, por exemplo:

$$5, 13, 10, 2, 18, 15, 6, 16, 9,$$

de acordo com a definição de mediana, o primeiro passo a ser dado é o da ordenação (crescente ou decrescente) dos valores:

$$2, 5, 6, 9, 10, 13, 15, 16, 18.$$

Em seguida, tomamos aquele valor central que apresenta o mesmo número de elementos à direita e à esquerda. Em nosso exemplo, esse valor é o **10**, já que, nessa série, há quatro elementos acima dele e quatro abaixo.

Temos, então:

$$\text{Md} = 10$$

Se, porém, a série dada tiver um número par de termos, a mediana será, por definição, qualquer dos números compreendidos entre os dois valores centrais da série. Convencionou-se utilizar o **ponto médio**.

Assim, a série de valores:

$$2, 6, 7, 10, 12, 13, 18, 21$$

tem para mediana a média aritmética entre **10** e **12**.

Logo:

$$\text{Md} = \frac{10+12}{2} = \frac{22}{2} = 11$$

donde:

$$Md = 11$$

Verificamos que, estando ordenados os valores de uma série e sendo **n** o número de elementos da série, o valor mediano será:

- o termo de ordem $\frac{n+1}{2}$, se **n** for ímpar;
- a média aritmética dos termos de ordem $\frac{n}{2}$ e $\frac{n}{2} + 1$, se **n** for par.
- Podemos comprovar tal fato nas séries dadas:
- para n = 9, temos $\frac{9+1}{2} = 5$. Logo, a mediana é o 5º termo da série, isto é:

$$Md = 10$$

- para n = 8, temos $\frac{8}{2} = 4$ e $\frac{8}{2} + 1 = 5$. Logo, a mediana é a média aritmética do 4º e 5º termos da série, isto é:

$$Md = \frac{10+12}{2} = \frac{22}{2} = 11$$

Logo:

$$Md = 11$$

NOTAS:

- O valor da mediana pode coincidir ou não com um elemento da série, como vimos. Quando o número de elementos da série é ímpar, há coincidência. O mesmo não acontece, porém, quando esse número é par.
- A mediana e a média aritmética não têm, necessariamente, o mesmo valor. Na primeira série apresentada, por exemplo, temos:

$$\bar{x} = 10{,}4 \text{ e } Md = 10$$

- A mediana, como vimos, depende da posição e não dos valores dos elementos na série ordenada. Essa é uma das diferenças marcantes entre a mediana e a média (que se deixa influenciar, e muito, pelos valores extremos). Esta propriedade da mediana pode ser constatada através dos exemplos a seguir:

$$5, 7, 10, 13, 15 \Rightarrow \bar{x} = 10 \text{ e } Md = 10$$

$$5, 7, 10, 13, 65 \Rightarrow \bar{x} = 20 \text{ e } Md = 10$$

isto é, a média do segundo conjunto de valores é maior do que a do primeiro, por influência dos valores extremos, ao passo que a mediana permanece a mesma.

- A mediana é designada, muitas vezes, por **valor mediano**.

6.4.2 Dados agrupados

Se os dados se agrupam em uma distribuição de frequência, o cálculo da mediana se processa de modo muito semelhante àquele dos dados não agrupados, implicando, porém, a determinação prévia das frequências acumuladas. Ainda aqui, temos de determinar um valor tal que divida a distribuição em dois grupos que contenham o mesmo número de elementos.

Para o caso de uma distribuição, porém, a ordem, a partir de qualquer um dos extremos, é dada por:

$$\frac{\sum f_i}{2}$$

Sem intervalos de classe

Neste caso, é o bastante identificar a frequência acumulada imediatamente superior à metade da soma das frequências. A mediana será aquele valor da variável que corresponde a tal frequência acumulada.

Tomemos a distribuição relativa à Tabela 6.1, completando-a com a coluna correspondente à frequência acumulada:

Nº DE MENINOS	f_i	F_i
0	2	2
1	6	8
2	10	18
3	12	30
4	4	34
	$\Sigma = 34$	

TABELA 6.7

Sendo:

$$\frac{\sum f_i}{2} = \frac{34}{2} = 17$$

a menor frequência acumulada que supera esse valor é **18**, que corresponde ao valor **2** da variável, sendo este o valor mediano. Logo:

$$\text{Md} = 2 \text{ meninos}$$

NOTA:

- No caso de existir uma frequência acumulada (F_i), tal que:

$$F_i = \frac{\sum f_i}{2},$$

a mediana será dada por:

$$Md = \frac{x_i + x_{i+1}}{2},$$

isto é, a mediana será a média aritmética entre o valor da variável correspondente a essa frequência acumulada e o seguinte.

Exemplo:

x_i	f_i	F_i
12	1	1
14	2	3
15	1	4
16	2	6
17	1	7
20	1	8
	$\Sigma = 8$	

TABELA 6.8

Temos:

$$\frac{8}{2} = 4 = F_3$$

Logo:

$$Md = \frac{15 + 16}{2} = \frac{31}{2} = 15,5$$

donde:

$$Md = 15,5$$

Resolva

1. Complete o esquema para o cálculo da mediana das distribuições:

a.

	2	4	6	8	10
	3	7	12	8	4

Temos:

x_i	f_i	F_i
2	3	....
4	7	10
6	12	....
8	8	30
10	4	....
	$\Sigma = ...$	

Como:

$$\frac{\sum f_i}{2} = \frac{....}{2} = ...$$

vem:

$$Md = ...$$

b.

	0	1	2	3	4	5
	2	5	9	7	6	3

Temos:

x_i	f_i	F_i
0	2	2
....	....	....
....	9	....
....	....	....
4	....	....
....	....	....
	$\Sigma = ...$	

Como:

$$\frac{\sum f_i}{2} = \frac{....}{2} = ...$$

vem:

$$Md = \frac{....}{....}$$

isto é:

$$Md = ...$$

Com intervalos de classe

Neste caso, o problema consiste em determinar o ponto do intervalo em que está compreendida a mediana.

Para tanto, temos inicialmente que determinar a classe na qual se acha a mediana — **classe mediana.** Tal classe será, evidentemente, aquela correspondente à frequência acumulada imediatamente superior a $\frac{\sum f_i}{2}$.

Feito isto, um problema de interpolação[2] resolve a questão, admitindo-se, agora, que os valores se distribuam uniformemente em todo o intervalo de classe.

Assim, considerando a distribuição da Tabela 6.3, acrescida das frequências acumuladas:

[2] **Interpolação** é a inserção de uma determinada quantidade de valores entre dois números dados.

i	ESTATURAS (cm)	f_i	F_i	
1	150 ⊢ 154	4	4	
2	154 ⊢ 158	9	13	
3	158 ⊢ 162	11	24	← classe mediana
4	162 ⊢ 166	8	32	
5	166 ⊢ 170	5	37	
6	170 ⊢ 174	3	40	
		$\Sigma = 40$		

TABELA 6.9

temos:

$$\frac{\sum f_i}{2} = \frac{40}{2} = 20$$

Como há **24** valores incluídos nas três primeiras classes da distribuição e como pretendemos determinar o valor que ocupa o **20º** lugar, a partir do início da série, vemos que este deve estar localizado na terceira classe (**i = 3**), supondo que as frequências dessas classes estejam uniformemente distribuídas.

Como há **11** elementos nessa classe e o intervalo de classe é igual a **4**, devemos tomar, a partir do limite inferior, a distância:

$$\frac{20-13}{11} \times 4 = \frac{7}{11} \times 4$$

e a mediana será dada por:

$$\text{Md} = 158 + \frac{7}{11} \times 4 = 158 + \frac{28}{11} = 158 + 2{,}54 = 160{,}54$$

Logo:

$$\text{Md} = 160{,}5 \text{ cm}$$

Na prática, executamos os seguintes passos:

1º) Determinamos as frequências acumuladas.

2º) Calculamos $\frac{\sum f_i}{2}$.

3º) Marcamos a classe correspondente à frequência acumulada imediatamente superior à $\frac{\sum f_i}{2}$ — **classe mediana** — e, em seguida, empregamos a fórmula:

$$\text{Md} = \ell^* + \frac{\left[\frac{\sum f_i}{2} - F(ant)\right] h^*}{f^*}$$

na qual:

ℓ^* é o limite inferior da classe mediana;

F (ant) é a frequência acumulada da classe anterior à classe mediana;

f* é a frequência simples da classe mediana;

h* é a amplitude do intervalo da classe mediana.

Tomando como exemplo a distribuição anterior, temos:

$$\frac{\sum f_i}{2} = \frac{40}{2} = 20$$

Logo, a classe mediana é a de ordem 3. Então:

$$\ell^* = 158,\ F(\text{ant}) = 13,\ f^* = 11 \text{ e } h^* = 4$$

Substituindo esses valores na fórmula, obtemos:

$$Md = 158 + \frac{(20-13)4}{11} = 158 + \frac{28}{11} = 158 + 2{,}54 = 160{,}54,$$

isto é:

$$Md = 160{,}5 \text{ cm}$$

Resolva

1. Complete o esquema para o cálculo da mediana da distribuição de frequência:

CUSTOS (R$)	450 ⊢ 550	550 ⊢ 650	650 ⊢ 750	750 ⊢ 850	850 ⊢ 950	950 ⊢ 1.050	1.050 ⊢ 1.150
f_i	8	10	11	16	13	5	1

Temos:

i	CUSTOS (R$)	f_i	F_i
1	450 ⊢ 550	8	8
2	550 ⊢ 650	....	18
3	650 ⊢ 750	....	....
4	750 ⊢ 850	....	....
5	850 ⊢ 950	....	....
6	950 ⊢ 1.050	....	....
7	1.050 ⊢ 1.150	....	....
		Σ = ...	

$$\frac{\sum f_i}{2} = \frac{....}{2} = ...$$

ℓ^* = ..., F(ant) = ..., f* = ... e h* = ...

Logo:

$$Md = ... + \frac{(.... - ....)....}{....} = ... + \frac{....}{....} = ... + ... = ...,$$

isto é:

Md = R$ 769

NOTA:

- No caso de existir uma frequência acumulada exatamente igual a $\frac{\sum f_i}{2}$, **a mediana será o limite superior da classe correspondente.**

Exemplo:

i	CLASSES	f_i	F_i
1	0 ⊢ 10	1	1
2	10 ⊢ 20	3	4
3	20 ⊢ 30	9	13
4	30 ⊢ 40	7	20
5	40 ⊢ 50	4	24
6	50 ⊢ 60	2	26
		$\sum = 26$	

TABELA 6.10

Temos:

$$\frac{\sum f_i}{2} = \frac{26}{2} = 13$$

Logo:

$$Md = L^* \Rightarrow Md = 30$$

6.4.3 Emprego da mediana

Empregamos a mediana quando:

a. desejamos obter o ponto que divide a distribuição em partes iguais;

b. há valores extremos que afetam de uma maneira acentuada a média;

c. a variável em estudo é salário.

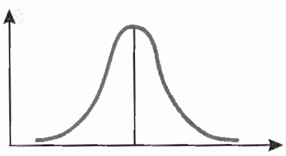
6.5 Posição relativa da média, mediana e moda

Quando uma distribuição é simétrica, as três medidas coincidem. A assimetria, porém, torna-as diferentes e essa diferença é tanto maior quanto maior é a assimetria. Assim, em uma distribuição em forma de sino, temos:

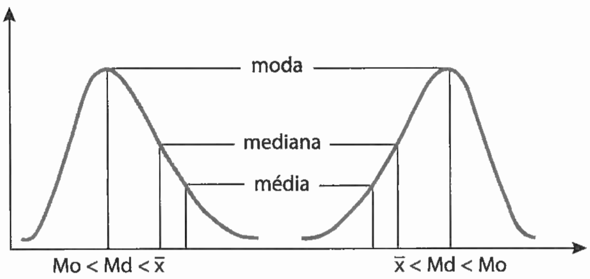
$\bar{x}$ = Md = Mo, no caso da **curva simétrica**;

Mo < Md < $\bar{x}$, no caso da **curva assimétrica positiva**;

$\bar{x}$ < Md < Mo, no caso da **curva assimétrica negativa**.



$\bar{x}$ = Md = Mo



6.6 As separatrizes

Como vimos, a mediana caracteriza uma série de valores devido à sua posição central. No entanto, ela apresenta uma outra característica, tão importante quanto a primeira: **ela separa a série em dois grupos que apresentam o mesmo número de valores**.

Assim, além das medidas de posição que estudamos, há outras que, consideradas individualmente, não são medidas de tendência central, mas estão ligadas à mediana relativamente à sua segunda característica, já que se baseiam em sua posição na série. Essas medidas — os **quartis**, os **percentis** e os **decis** — são, juntamente com a **mediana**, conhecidas pelo nome genérico de **separatrizes**.

6.6.1 Os quartis

Denominamos **quartis** os valores de uma série que a dividem em quatro partes iguais.

Há, portanto, três quartis:

a. O **primeiro quartil (Q_1)** — valor situado de tal modo na série que uma quarta parte **(25%)** dos dados é menor que ele e as três quartas partes restantes **(75%)** são maiores.

b. O **segundo quartil (Q_2)** — evidentemente, coincide com a mediana **(Q_2 = Md)**.

c. O **terceiro quartil (Q_3)** — valor situado de tal modo que as três quartas partes **(75%)** dos termos são menores que ele e uma quarta parte **(25%)** é maior.

Quando os dados são agrupados, para determinar os quartis usamos a mesma técnica do cálculo da mediana, bastando substituir, na fórmula da mediana, $\frac{\sum f_i}{2}$ por:

$$\frac{k\sum f_i}{4}$$

sendo **k** o número de ordem do quartil.

Assim, temos:

$$Q_1 = \ell\star + \frac{\left[\frac{\sum f_i}{4} - F(ant)\right] h\star}{f\star}$$

e

$$Q_3 = \ell\star + \frac{\left[\frac{3\sum f_i}{4} - F(ant)\right] h\star}{f\star}$$

Exemplo:

ESTATURAS (cm)	f_i	F_i
150 ⊢ 154	4	4
154 ⊢ 158	9	13 ← (Q_1)
158 ⊢ 162	11	24
162 ⊢ 166	8	32 ← (Q_3)
166 ⊢ 170	5	37
170 ⊢ 174	3	40
	$\Sigma = 40$	

TABELA 6.11

Primeiro quartil

Temos:

$$\frac{\sum f_i}{4} = \frac{40}{4} = 10$$

$$Q_1 = 154 + \frac{(10-4)\ 4}{9} =$$

$$= 154 + \frac{24}{9} = 154 + 2,66 = 156,66$$

$$Q_1 = 156,7 \text{ cm}$$

Terceiro quartil

Temos:

$$\frac{3\sum f_i}{4} = \frac{3 \times 40}{4} = 30$$

$$Q_3 = 162 + \frac{(30-24)\ 4}{8} =$$

$$= 162 + \frac{24}{8} = 162 + 3 = 165$$

$$Q_3 = 165 \text{ cm}$$

Resolva

1. Complete os esquemas para o cálculo do primeiro e do terceiro quartis da distribuição de frequência:

CUSTOS (R\$)	450 ⊢ 550	550 ⊢ 650	650 ⊢ 750	750 ⊢ 850	850 ⊢ 950	950 ⊢ 1.050	1.050 ⊢ 1.150
f_i	8	10	11	16	13	5	1

Temos:

i	CUSTOS (R\$)	f_i	F_i
1	450 ⊢ 550	8	8
2	550 ⊢ 650	10	18 ← (Q_1)
3	650 ⊢ 750	11	29
4	750 ⊢ 850	16	45
5	850 ⊢ 950	13	58 ← (Q_3)
6	950 ⊢ 1.050	5	63
7	1.050 ⊢ 1.150	1	64
		$\Sigma = 64$	

Primeiro quartil

$$k = 1 \Rightarrow \frac{\sum f_i}{4} = \frac{....}{4} = ...$$

$\ell^* = ...$, F(ant) = ..., $f^* = ...$, $h^* = ...$

$$Q_1 = ... + \frac{(.... - ....)\ ....}{....} =$$

$$= .... + \frac{.... \times ....}{....} =$$

$$= .... + ... = ...$$

Q_1 = R\$ 630

Terceiro quartil

$$k = 3 \Rightarrow \frac{3\sum f_i}{4} = \frac{3 \times ....}{4} = \frac{....}{4} = ...$$

$\ell^* = ...$, F(ant) = ..., $f^* = ...$, $h^* = ...$

$$Q_3 = ... + \frac{(.... - ....)\ ....}{....} =$$

$$= .... + \frac{.... \times ....}{....} =$$

$$= .... + ... = ...$$

Q_3 = R\$ 873

6.6.2 Os percentis

Denominamos **percentis** os 99 valores que separam uma série em 100 partes iguais.

Indicamos:

$$P_1, P_2, ..., P_{32}, ..., P_{99}$$

É evidente que:

$$P_{50} = Md, P_{25} = Q_1 \text{ e } P_{75} = Q_3$$

O cálculo de um percentil segue a mesma técnica do cálculo da mediana, porém, a fórmula $\frac{\sum f_i}{2}$ será substituída por:

$$\frac{k\sum f_i}{100}$$

sendo **k** o número de ordem do percentil.

Assim, para o 27º percentil, temos:

$$k = 27 \Rightarrow P_{27} = \ell\star + \frac{\left(\frac{27\sum f_i}{100} - F(ant)\right) h\star}{f\star}$$

Exemplo:

Considerando a Tabela 6.11, temos, para o oitavo percentil:

$$k = 8 \Rightarrow \frac{8\sum f_i}{100} = \frac{8 \times 40}{100} = 3{,}2$$

Logo:

$$P_8 = 150 + \frac{(3{,}2 - 0)\ 4}{4} = 150 + \frac{12{,}8}{4} = 150 + 3{,}2 = 153{,}2$$

donde:

$$P_8 = 153{,}2 \text{ cm}$$

Resolva

1. Complete o esquema para o cálculo do vigésimo percentil da distribuição:

CUSTOS (R$)	450 ⊢ 550	550 ⊢ 650	650 ⊢ 750	750 ⊢ 850	850 ⊢ 950	950 ⊢ 1.050	1.050 ⊢ 1.150
f_i	8	10	11	16	13	5	1

Temos:

i	CUSTOS (R$)	f_i	F_i
1	450 ⊢ 550	8	8
2	550 ⊢ 650	10	18 ← (P_{20})
3	650 ⊢ 750	11	29
4	750 ⊢ 850	16	45
5	850 ⊢ 950	13	58
6	950 ⊢ 1.050	5	63
7	1.050 ⊢ 1.150	1	64
		$\Sigma = 64$	

$$k = 20 \Rightarrow \frac{20\sum f_i}{100} = \frac{20 \times ....}{100} = \frac{....}{100} = ...$$

$\ell^* = ...$, F(ant) $= ....$, $f^* = ...$, $h^* = ...$

$$P_{20} = ... + \frac{(.... - ....)\ ....}{....} =$$

$$= .... + \frac{.... \times ....}{....} =$$

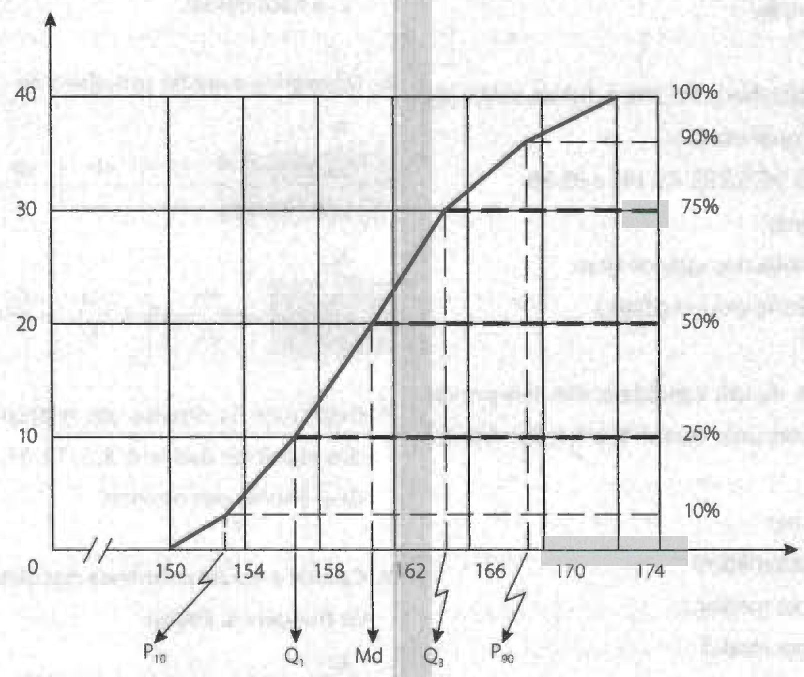
$$= .... + ... = ...$$

isto é:

$$P_{20} = \text{R\$ } 598$$

NOTA:

- **Construindo o polígono de frequência acumulada percentual, podemos determinar, geometricamente, as separatrizes:**



Exercícios

1. Considerando os conjuntos de dados:
   a. 3, 5, 2, 6, 5, 9, 5, 2, 8, 6
   b. 20, 9, 7, 2, 12, 7, 20, 15, 7
   c. 51,6; 48,7; 50,3; 49,5; 48,9
   d. 15, 18, 20, 13, 10, 16, 14

   calcule:
   I. a média;
   II. a mediana;
   III. a moda.

2. Os salários-hora de cinco funcionários de uma companhia são:
   R$ 75, R$ 90, R$ 83, R$ 142 e R$ 88.
   Determine:
   a. a média dos salários-hora;
   b. o salário-hora mediano.

3. As notas de um candidato, em seis provas de um concurso, foram: 8,4; 9,1; 7,2; 6,8; 8,7 e 7,2.
   Determine:
   a. a nota média;
   b. a nota mediana;
   c. a nota modal.

4. Considerando a distribuição abaixo:

	3	4	5	6	7	8
	4	8	11	10	8	3

   calcule:
   a. a média; b. a mediana; c. a moda.

5. Em uma classe de 50 alunos, as notas obtidas formaram a seguinte distribuição:

NOTAS	2	3	4	5	6	7	8	9	10
Nº DE ALUNOS	1	3	6	10	13	8	5	3	1

   calcule:
   a. a nota média;
   b. a nota mediana;
   c. a nota modal.

6. Determine a média aritmética de:

   a.

	50	60	80	90
	8	5	4	3

   b.

	50	58	66
	20	50	30

7. Determine os desvios em relação à média dos seguintes dados: 6, 8, 5, 12, 11, 7, 4, 15. Qual a soma dos desvios?

8. Calcule a média aritmética das distribuições de frequência abaixo:

   a.

NOTAS	f_i
0 ⊢ 2	5
2 ⊢ 4	8
4 ⊢ 6	14
6 ⊢ 8	10
8 ⊢ 10	7
	$\Sigma = 44$

b.

ESTATURAS (cm)	f_i
150 ⊢ 158	5
158 ⊢ 166	12
166 ⊢ 174	18
174 ⊢ 182	27
182 ⊢ 190	8
	$\Sigma = 70$

c.

SALÁRIOS (R$)	f_i
500 ⊢ 700	18
700 ⊢ 900	31
1.900 ⊢ 1.100	15
1.100 ⊢ 1.300	3
1.300 ⊢ 1.500	1
1.500 ⊢ 1.700	1
1.700 ⊢ 1.900	1
	$\Sigma = 70$

d.

PESOS (kg)	f_i
145 ⊢ 151	10
151 ⊢ 157	9
157 ⊢ 163	8
163 ⊢ 169	6
169 ⊢ 175	3
175 ⊢ 181	3
181 ⊢ 187	1
	$\Sigma = 40$

9. Calcule a mediana de cada uma das distribuições do exercício **8**.

10. Calcule a moda de cada uma das distribuições do exercício **8**.

11. Calcule o primeiro e o terceiro quartis das distribuições do exercício **8**.

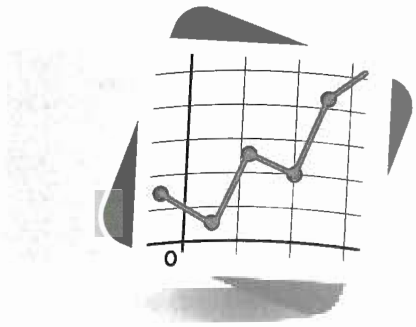
12. Calcule o 10º, o 1º, o 23º, o 15º e o 90º percentis da distribuição **b** do exercício **8**.

13. A curva de frequência acumulada serve para determinar:

a. a lei do acaso.
b. a média.
c. a mediana.
d. a moda.
e. o desvio padrão.

14. Uma curva simétrica se caracteriza pelo seguinte atributo:

a. É assimétrica à esquerda.
b. A moda é maior que a mediana e a média.
c. A moda, a mediana e a média são iguais.
d. O desvio padrão é maior que a mediana e a moda.
e. Os decis são equivalentes à média.



7 MEDIDAS DE DISPERSÃO OU DE VARIABILIDADE

7.1 Dispersão ou variabilidade

Vimos anteriormente que um conjunto de valores pode ser convenientemente sintetizado, por meio de procedimentos matemáticos, em poucos valores representativos — **média aritmética**, **mediana** e **moda**. Tais valores podem servir de comparação para dar a posição de qualquer elemento do conjunto.

No entanto, quando se trata de interpretar dados estatísticos, mesmo aqueles já convenientemente simplificados, é necessário ter-se uma ideia retrospectiva de como se apresentavam esses mesmos dados nas tabelas.

Assim, não é o bastante dar uma das medidas de posição para caracterizar perfeitamente um conjunto de valores, pois, mesmo sabendo, por exemplo, que a temperatura média de duas cidades é a mesma, e igual a 24 °C, ainda assim somos levados a pensar a respeito do clima dessas cidades. Em uma delas poderá a temperatura variar entre limites de muito calor e de muito frio e haver, ainda, uma temperatura média de 24 °C. A outra poderá ter uma variação pequena de temperatura e possuir, portanto, no que se refere à temperatura, um clima mais favorável.

Vemos, então, que a média — ainda que considerada como um número que tem a faculdade de representar uma série de valores — não pode, por si mesma, destacar o grau de homogeneidade ou heterogeneidade que existe entre os valores que compõem o conjunto.

Consideremos os seguintes conjuntos de valores das variáveis **x**, **y** e **z**:

X: 70, 70, 70, 70, 70.

Y: 68, 69, 70, 71, 72.

Z: 5, 15, 50, 120, 160.

Calculando a média aritmética de cada um desses conjuntos, obtemos:

$$\bar{x} = \frac{\sum x_i}{n} \Rightarrow \bar{x} = \frac{350}{5} = 70$$

$$\bar{y} = \frac{\sum y_i}{n} \Rightarrow \bar{y} = \frac{350}{5} = 70$$

$$\bar{z} = \frac{\sum z_i}{n} \Rightarrow \bar{z} = \frac{350}{5} = 70$$

Vemos, então, que os três conjuntos apresentam a mesma média aritmética: **70**.

Entretanto, é fácil notar que o conjunto **X** é **mais homogêneo** que os conjuntos **Y** e **Z**, já que todos os valores são iguais à média.

O conjunto **Y**, por sua vez, é **mais homogêneo** que o conjunto **Z**, pois há menor diversificação entre cada um de seus valores e a média representativa.

Chamando de **dispersão** ou **variabilidade** a maior ou menor diversificação dos valores de uma variável em torno de um valor de tendência central tomado como ponto de comparação, podemos dizer que o conjunto **X** apresenta **dispersão** ou **variabilidade nula** e que o conjunto **Y** apresenta uma **dispersão** ou **variabilidade menor** que o conjunto **Z**.

Portanto, para qualificar os valores de uma dada variável, ressaltando a maior ou menor dispersão ou variabilidade entre esses valores e a sua medida de posição, a Estatística recorre às **medidas de dispersão** ou **de variabilidade**.

Dessas medidas, estudaremos a **amplitude total**, a **variância**, o **desvio padrão** e o **coeficiente de variação**.

7.2 Amplitude total

7.2.1 Dados não agrupados

A **amplitude total** é a diferença entre o maior e o menor valor observado:

$$AT = x(\text{máx.}) - x(\text{mín.})$$

Exemplo:

Para os valores:

40, 45, 48, 52, 54, 62 e 70

temos:

$$AT = 70 - 40 = 30$$

Logo:

$$AT = 30$$

Quando dizemos que a amplitude total dos valores é **30**, estamos afirmando alguma coisa do grau de sua concentração. É evidente que, quanto maior a amplitude total, maior a dispersão ou variabilidade dos valores da variável.

Relativamente aos três conjuntos de valores mencionados no início deste capítulo, temos:

$$AT_x = 70 - 70 = 0, \text{(dispersão nula)}$$
$$AT_y = 72 - 68 = 4$$
$$AT_z = 160 - 5 = 155$$

7.2.2 Dados agrupados

Sem intervalos de classe

Neste caso, ainda temos:

$$AT = x(\text{máx.}) - x(\text{mín.})$$

Exemplo:

Considerando a tabela abaixo:

	0	1	2	3	4
	2	6	12	7	3

TABELA 7.1

temos:

$$AT = 4 - 0 = 4$$

Logo:

$$AT = 4$$

Com intervalos de classe

Neste caso, a **amplitude total** é a diferença entre o limite superior da última classe e o limite inferior da primeira classe:

$$AT = L(\text{máx.}) - \ell(\text{mín.})$$

Exemplo:

Considerando a distribuição abaixo:

i	ESTATURAS (cm)	f_i
1	150 ⊢ 154	4
2	154 ⊢ 158	9
3	158 ⊢ 162	11
4	162 ⊢ 166	8
5	166 ⊢ 170	5
6	170 ⊢ 174	3
		$\Sigma = 40$

TABELA 7.2

temos:

$$AT = 174 - 150 = 24$$

Logo:

$$AT = 24 \text{ cm}$$

A amplitude total tem o inconveniente de só levar em conta os dois valores extremos da série, descuidando do conjunto de valores intermediários, o que quase sempre invalida a idoneidade do resultado. Ela é apenas uma indicação aproximada da dispersão ou variabilidade.

Faz-se uso da amplitude total quando se quer determinar a amplitude da temperatura em um dia ou no ano, no controle de qualidade ou como uma medida de cálculo rápido, e quando a compreensão popular é mais importante que a exatidão e a estabilidade.

7.3 Variância
Desvio padrão

7.3.1 Introdução

Como vimos, a amplitude total é instável, por se deixar influenciar pelos valores extremos, que são, na sua maioria, devidos ao acaso.

A **variância** e o **desvio padrão** são medidas que fogem a essa falha, pois levam em consideração a totalidade dos valores da variável em estudo, o que faz delas índices de variabilidade bastante estáveis e, por isso mesmo, os mais geralmente empregados.

A **variância** baseia-se nos desvios em torno da média aritmética, porém determinando a **média aritmética dos quadrados dos desvios**[1]. Assim, representando a variância por s^2, temos:

---

[1] Lembremos que $\sum d_i = \sum (x_i - \bar{x}) = 0$.

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{\sum f_i}$$

Ou, lembrando que $\sum f_i = n$:

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

NOTA:

- Quando nosso interesse não se restringe à descrição dos dados mas, partindo da amostra, visamos a tirar inferências válidas para a respectiva população, convém efetuar uma modificação, que consiste em usar o divisor **n – 1** em lugar de **n**.
  Podemos, ainda, com o intuito de conservar a definição, calcular a variância usando o divisor **n** e, em seguida, multiplicar o resultado por $\frac{n}{n-1}$.

Sendo a variância calculada a partir dos quadrados dos desvios, ela é um número em unidade quadrada em relação à variável em questão, o que, sob o ponto de vista prático, é um inconveniente.

Por isso mesmo, imaginou-se uma nova medida que tem utilidade e interpretação práticas, denominada **desvio padrão**, definida como a **raiz quadrada da variância** e representada por **s**:

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

Assim:

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

NOTA:

- Tanto o **desvio padrão** como a **variância** são usados como medidas de dispersão ou variabilidade. O uso de uma ou de outra dependerá da finalidade que se tenha em vista.
  A **variância** é uma medida que tem pouca utilidade como estatística descritiva, porém é extremamente importante na inferência estatística e em combinações de amostras.

Se bem que a fórmula dada para o cálculo do desvio seja a que torna mais fácil a sua compreensão, ela não é uma boa fórmula para fins de computação, pois, em geral, a média aritmética $(\bar{x})$ é um número fracionário, o que torna pouco prático o cálculo das quantidades $(x_i - \bar{x})^2$.

Podemos simplificar os cálculos fazendo uso da igualdade:

$$\sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}$$

Assim, substituindo $\sum (x_1 - \bar{x})^2$ por seu equivalente em ①, obtemos:

$$s = \sqrt{\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}},$$

que pode ser escrita do seguinte modo:

$$s = \sqrt{\frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2} \qquad ②$$

Não apenas este método é usualmente mais prático, como também mais preciso. Quando a média não é exata e tem de ser arredondada, cada desvio fica afetado ligeiramente do erro, devido a esse arredondamento. O mesmo acontece com os quadrados, podendo os resultados do cálculo ser menos exatos do que quando a fórmula ② é usada.

O desvio padrão goza de algumas propriedades, dentre as quais destacamos:

1ª) Somando-se (ou subtraindo-se) uma constante a (de) todos os valores de uma variável, o desvio padrão não se altera:

$$y_i = x_i \pm c \Rightarrow s_y = s_x$$

2ª) Multiplicando-se todos os valores de uma variável por uma constante (diferente de zero), o desvio padrão fica multiplicado por essa constante:

$$y_i = c \times x_i \Rightarrow s_y = c \times s_x$$

Essas propriedades nos permitem introduzir, no cálculo do desvio padrão, simplificações úteis, como veremos mais adiante.

Para o cálculo do desvio padrão, consideremos os seguintes casos:

7.3.2 Dados não agrupados

Tomemos, como exemplo, o conjunto de valores da variável **x**:

40, 45, 48, 52, 54, 62, 70.

O modo mais prático para se obter o desvio padrão é formar uma tabela com duas colunas: uma para x_i e outra para x_i^2. Assim:

x_i	x_i^2
40	1.600
45	2.025
48	2.304
52	2.704
54	2.916
62	3.844
70	4.900
$\Sigma = 371$	$\Sigma = 20.293$

TABELA 7.3

Como n = 7, temos:

$$s = \sqrt{\frac{20.293}{7} - \left(\frac{371}{7}\right)^2} = \sqrt{2.899 - 53^2} = \sqrt{2.899 - 2.809} = \sqrt{90} = 9,486$$

Logo:

$$s = 9,49$$

Resolva

1. Complete o esquema para o cálculo do desvio padrão, dados os valores da variável:

8, 10, 11, 15, 16, 18

Temos:

	x_i	x_i^2
	8	64
	....	....
	....	....
	....	....
	....	....
	....	....
n = ...	Σ = ...	Σ = ...

Logo:

$$s = \sqrt{\frac{....}{....} - \left(\frac{....}{....}\right)^2} = \sqrt{.... - ....^2} =$$

$$= \sqrt{.... - ....} =$$

$$= \sqrt{....} = ...,$$

isto é:

$$s = 3,56$$

2. Comprove a primeira propriedade do desvio padrão somando 5 a cada valor da variável do exercício anterior.

3. Comprove a segunda propriedade do desvio padrão multiplicando por 2 cada valor da variável do exercício 1.

7.3.3 Dados agrupados

Sem intervalos de classe

Como, neste caso, temos a presença de frequências, devemos levá-las em consideração, resultando a fórmula:

$$s = \sqrt{\frac{\sum f_i x_i^2}{n} - \left(\frac{\sum f_i x_i}{n}\right)^2}$$

Consideremos, como exemplo, a distribuição da Tabela 7.1.

O modo mais prático para se obter o desvio padrão é abrir, na tabela dada, uma coluna para os produtos $f_i x_i$ e outra para $f_i x_i^2$, lembrando que para obter $f_i x_i^2$ basta multiplicar cada $f_i x_i$ pelo seu respectivo x_i. Assim:

x_i	f_i	$f_i x_i$	$f_i x_i^2$
0	2	0	0
1	6	6	6
2	12	24	48
3	7	21	63
4	3	12	48
	$\Sigma = 30$	$\Sigma = 63$	$\Sigma = 165$

TABELA 7.4

Logo:

$$s = \sqrt{\frac{165}{30} - \left(\frac{63}{30}\right)^2} = \sqrt{5,5 - 4,41} = \sqrt{1,09} = 1,044$$

Daí:

$$s = 1,04$$

Resolva

1. Complete o esquema para o cálculo do desvio padrão da distribuição:

1	2	3	4	5	6
2	5	8	6	3	1

Temos:

x_i	f_i	f_ix_i	$f_ix_i^2$
1	2	2	2
2	....	....	....
3	....	....	....
4	....	....	....
5	....	....	....
6	....	....	....
	$\Sigma =$	$\Sigma =$	$\Sigma =$ ....

Logo:

$$s = \sqrt{\frac{....}{....} - \left(\frac{....}{....}\right)^2} = \sqrt{.... - (....)^2} =$$

$$= \sqrt{.... - ....} =$$

$$= \sqrt{....} = ...,$$

isto é:

$$s = 1{,}24$$

Com intervalos de classe

Tomemos como exemplo a distribuição da Tabela 7.2.

Começamos por abrir as colunas para x_i (ponto médio), para f_ix_i e para $f_ix_i^2$. Assim:

i	ESTATURAS (cm)	f_i	x_i	f_ix_i	$f_ix_i^2$
1	150 ⊢ 154	4	152	608	92.416
2	154 ⊢ 158	9	156	1.404	219.024
3	158 ⊢ 162	11	160	1.760	281.600
4	162 ⊢ 166	8	164	1.312	215.168
5	166 ⊢ 170	5	168	840	141.120
6	170 ⊢ 174	3	172	516	88.752
		$\Sigma = 40$		$\Sigma = 6.440$	$\Sigma = 1.038.080$

TABELA 7.5

Logo:

$$s = \sqrt{\frac{1.038.080}{40} - \left(\frac{6.440}{40}\right)^2} = \sqrt{25.952 - 25.921} = \sqrt{31} = 5{,}567$$

Daí:

$$s = 5{,}57 \text{ cm}$$

7.3.4 Processo breve

Baseados na mudança da variável **x** por outra **y**, tal que:

$$y_i = \frac{x_i - x_0}{h},$$

e pelas mesmas razões expostas para o cálculo da média, podemos obter um processo breve de cálculo, com a aplicação da seguinte fórmula:

$$s = h\sqrt{\frac{\sum f_i y_i^2}{n} - \left(\frac{\sum f_i y_i}{n}\right)^2}$$

Assim, para a distribuição da Tabela 7.2, temos, completando com as colunas para x_i, y_i, $f_i y_i$ e $f_i y_i^2$:

i	ESTATURAS (cm)	f_i	x_i	y_i	$f_i y_i$	$f_i y_i^2$
1	150 ⊢ 154	4	152	−2	−8	16
2	154 ⊢ 158	9	156	−1	−9	9
3	158 ⊢ 162	11	160	0	0	0
4	162 ⊢ 166	8	164	1	8	8
5	166 ⊢ 170	5	168	2	10	20
6	170 ⊢ 174	3	172	3	9	27
	h = 4	$\Sigma = 40$			$\Sigma = 10$	$\Sigma = 80$

TABELA 7.6

Logo:

$$s = 4\sqrt{\frac{80}{40} - \left(\frac{10}{40}\right)^2} = 4\sqrt{2 - 0{,}0625} = 4\sqrt{1{,}9375} = 4 \leftrightarrow 1{,}3919 = 5{,}5676$$

Daí:

$$s = 5{,}57 \text{ cm}$$

NOTA:

- **Valem as mesmas observações que fizemos para a média aritmética (p. 81).**

Fases para o cálculo do desvio padrão pelo processo breve:

1ª) Abrimos uma coluna para os valores x_i (ponto médio).

2ª) Escolhemos um dos pontos médios (de preferência o de maior frequência) para valor de x_0.

3ª) Abrimos uma coluna para os valores de y_i e escrevemos **zero** na linha correspondente à classe onde se encontra o valor de x_0; a sequência **−1, −2, −3, ...**, logo acima de zero, e a sequência **1, 2, 3, ...**, logo abaixo.

4ª) Abrimos uma coluna para os valores do produto f_iy_i, conservando os sinais + ou −, e, em seguida, somamos algebricamente esses produtos.

5ª) Abrimos uma coluna para os valores do produto $f_iy_i^2$, obtidos multiplicando cada f_iy_i pelo seu respectivo y_i, e, em seguida, somamos esses produtos.

6ª) Aplicamos a fórmula.

Exercício resolvido

1. Calcule o desvio padrão da distribuição, pelo processo breve.

CUSTOS (R$)	450 ⊢ 550	550 ⊢ 650	650 ⊢ 750	750 ⊢ 850	850 ⊢ 950	950 ⊢ 1.050	1.050 ⊢ 1.150
f_i	8	10	11	16	13	5	1

Temos:

i	x_i	f_i	y_i	f_iy_i	$f_iy_i^2$
1	500	8	−3	−24	72
2	600	10	−2	−20	40
3	700	11	−1	−11 −55	11
4	800	16	0	0	0
5	900	13	1	13	13
6	1.000	5	2	10	20
7	1.100	1	3	3 26	9
	h = 100	$\Sigma = 64$		$\Sigma = -29$	$\Sigma = 165$

Como h = 100, vem:

$$s = 100\sqrt{\frac{165}{64} - \left(\frac{-29}{64}\right)^2} = 100\sqrt{2{,}5781 - (0{,}4531)^2} = 100\sqrt{2{,}5781 - 0{,}2052} =$$

$$100\sqrt{2{,}3729} = 100 \times 1{,}54042 = 154{,}042 \Rightarrow$$

$$\Rightarrow s = \text{R\$ } 154$$

Resolva

1. Complete o esquema para o cálculo do desvio padrão da distribuição, pelo processo breve:

	30 ⊢ 50	50 ⊢ 70	70 ⊢ 90	90 ⊢ 110	110 ⊢ 130
	2	8	12	10	5

Temos:

i	x_i	f_i	y_i	f_iy_i	$f_iy_i^2$
1	40	2	....	....	....
2	....	....	....	....	....
3	....	....	....	....	....
4	....	....	....	....	....
5	....	....	....	....	....
	h = ...	Σ = ...		Σ = ...	Σ = ...

Logo:

$$s = \ldots\sqrt{\frac{\ldots}{\ldots} - \left(\frac{\ldots}{\ldots}\right)^2} = \ldots\sqrt{\ldots - (\ldots)^2} = \ldots\sqrt{\ldots - \ldots} = \ldots\sqrt{\ldots} =$$

$$= \ldots \times \ldots = \ldots,$$

isto é:

$$s = 21{,}88$$

7.4 Coeficiente de variação

O desvio padrão por si só não nos diz muita coisa. Assim, um desvio padrão de duas unidades pode ser considerado pequeno para uma série de valores cujo valor médio é **200**; no entanto, se a média for igual a **20**, o mesmo não pode ser dito. Além disso, o fato de o desvio padrão ser expresso na mesma unidade dos dados limita o seu emprego quando desejamos comparar duas ou mais séries de valores, relativamente à sua dispersão ou variabilidade, quando expressas em unidades diferentes.

Para contornar essas dificuldades e limitações, podemos caracterizar a dispersão ou variabilidade dos dados em termos relativos a seu valor médio, medida essa denominada **coeficiente de variação (CV)**:

$$CV = \frac{s}{\bar{x}} \times 100$$

Para a distribuição da Tabela 7.6, onde $\bar{x}$ = 161 cm e s = 5,57 cm, temos:

$$CV = \frac{5,57}{161} \times 100 = 0,03459 \times 100 = 3,459$$

Daí:

$$CV = 3,5\%$$

Exemplo:

Tomemos os resultados das medidas das estaturas e dos pesos de um mesmo grupo de indivíduos:

	$\bar{x}$	s
ESTATURAS	175 cm	5,0 cm
PESOS	68 kg	2,0 kg

Temos:

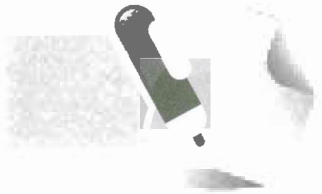
$$CV_E = \frac{5}{175} \times 100 = 0,0285 \times 100 = 2,85\%$$

$$CV_P = \frac{2}{68} \times 100 = 0,0294 \times 100 = 2,94\%$$

Logo, nesse grupo de indivíduos, os pesos apresentam maior grau de dispersão que as estaturas.

NOTA:

- Se bem que, para qualificar a dispersão de uma distribuição, seja mais proveitoso o coeficiente de variação, não devemos deduzir daí que a variância e o desvio padrão careçam de utilidade. Pelo contrário, são medidas muito úteis no tratamento de assuntos relativos à inferência estatística, como já dissemos.



Exercícios

1. Calcule a amplitude total dos conjuntos de dados:
   a. 1, 3, 5, 9
   b. 20, 14, 15, 19, 21, 22, 20
   c. 17,9; 22,5; 13,3; 16,8; 15,4; 14,2
   d. –10, –6, 2, 3, 7, 9, 10

2. Calcule a amplitude total das distribuições:

   a.

	2	3	4	5	6	7	8
	1	3	5	8	5	4	2

   b.

	1,5 ⊢ 1,6	1,6 ⊢ 1,7	1,7 ⊢ 1,8	1,8 ⊢ 1,9	1,9 ⊢ 2,0	2,0 ⊢ 2,1	2,1 ⊢ 2,2
	4	8	12	15	12	8	4

3. Calcule os desvios padrões dos conjuntos de dados do exercício **1**.

4. Calcule os desvios padrões das distribuições do exercício **2**.

5. Dada a distribuição relativa a cem lançamentos de cinco moedas simultaneamente:

	0	1	2	3	4	5
	4	14	34	29	16	3

   calcule o desvio padrão.

6. Calcule o desvio padrão da distribuição:

	2 ⊢ 6	6 ⊢ 10	10 ⊢ 14	14 ⊢ 18	18 ⊢ 22
	5	12	21	15	7

7. Calcule os desvios padrões das distribuições do exercício **8**, cap. 6, p. 100-101.

8. Sabendo que um conjunto de dados apresenta para média aritmética e para desvio padrão, respectivamente, 18,3 e 1,47, calcule o coeficiente de variação.

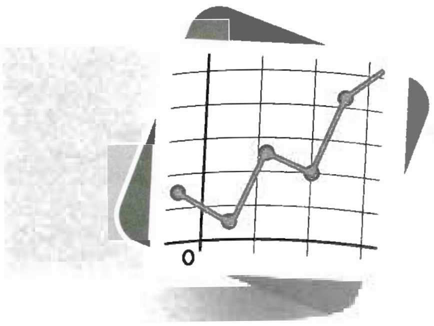
9. Em um exame final de Matemática, o grau médio de um grupo de 150 alunos foi 7,8 e o desvio padrão, 0,80. Em Estatística, entretanto, o grau médio final foi 7,3 e o desvio padrão, 0,76. Em que disciplina foi maior a dispersão?

10. Medidas as estaturas de 1.017 indivíduos, obtivemos $\bar{x} = 162{,}2$ cm e $s = 8{,}01$ cm. O peso médio desses mesmos indivíduos é 52 kg, com um desvio padrão de 2,3 kg. Esses indivíduos apresentam maior variabilidade em estatura ou em peso?

11. Um grupo de 85 moças tem estatura média de 160,6 cm, com um desvio padrão igual a 5,97 cm. Outro grupo de 125 moças tem uma estatura média de 161,9 cm, sendo o desvio padrão igual a 6,01 cm. Qual é o coeficiente de variação de cada um dos grupos? Qual o grupo mais homogêneo?

12. Um grupo de cem estudantes tem uma estatura média de 163,8 cm, com um coeficiente de variação de 3,3%. Qual o desvio padrão desse grupo?

13. Uma distribuição apresenta as seguintes estatísticas: $s = 1{,}5$ e $CV = 2{,}9\%$. Determine a média da distribuição.

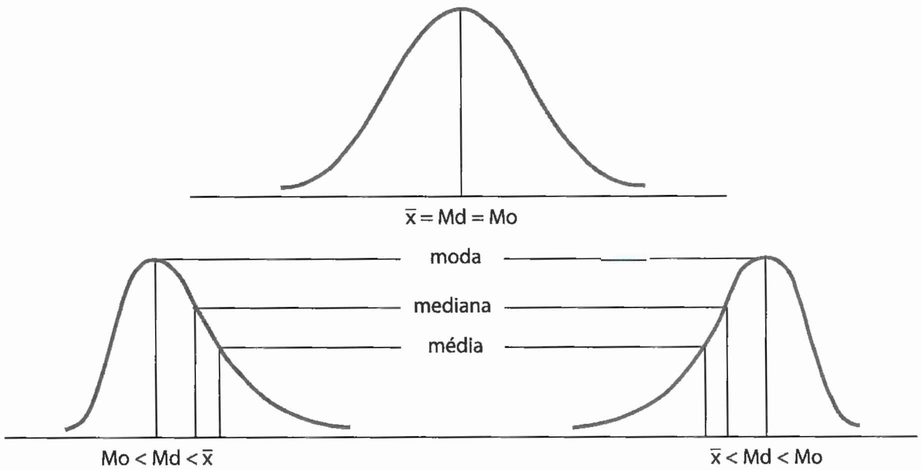


8 MEDIDAS DE ASSIMETRIA MEDIDAS DE CURTOSE

8.1 Assimetria

8.1.1 Introdução

A natureza da **assimetria** já foi estudada no capítulo **6**, item 6.5, quando vimos que, sendo a distribuição **simétrica**, a média e a moda coincidem; sendo a distribuição **assimétrica à esquerda ou negativa**, a média é menor que a moda; e sendo **assimétrica à direita ou positiva**, a média é maior que a moda:



Baseando-nos nessas relações entre a média e a moda, podemos empregá-las para determinar o tipo de assimetria. Assim, calculando o valor da diferença:

$$\bar{x} - Mo$$

se:

$\bar{\mathbf{x}}$ **– Mo = 0** $\Rightarrow$ assimetria nula ou distribuição simétrica;

$\bar{\mathbf{x}}$ **– Mo < 0** $\Rightarrow$ assimetria negativa ou à esquerda;

$\bar{\mathbf{x}}$ **– Mo > 0** $\Rightarrow$ assimetria positiva ou à direita.

Exemplo:

DISTRIBUIÇÃO A

PESOS(kg)	f_i
2 ⊢ 6	6
6 ⊢ 10	12
10 ⊢ 14	24
14 ⊢ 18	12
18 ⊢ 22	6
	$\Sigma = 60$

DISTRIBUIÇÃO B

PESOS(kg)	f_i
2 ⊢ 6	6
6 ⊢ 10	12
10 ⊢ 14	24
14 ⊢ 18	30
18 ⊢ 22	6
	$\Sigma = 78$

DISTRIBUIÇÃO C

PESOS(kg)	f_i
2 ⊢ 6	6
6 ⊢ 10	30
10 ⊢ 14	24
14 ⊢ 18	12
18 ⊢ 22	6
	$\Sigma = 78$

Temos:

A	B	C
$\bar{x}$ = 12 kg	$\bar{x}$ = 12,9 kg	$\bar{x}$ = 11,1 kg
Md = 12 kg	Md = 13,5 kg	Md = 10,5 kg
Mo = 12 kg	Mo = 16 kg	Mo = 8 kg
s = 4,42 kg	s = 4,20 kg	s = 4,20 kg

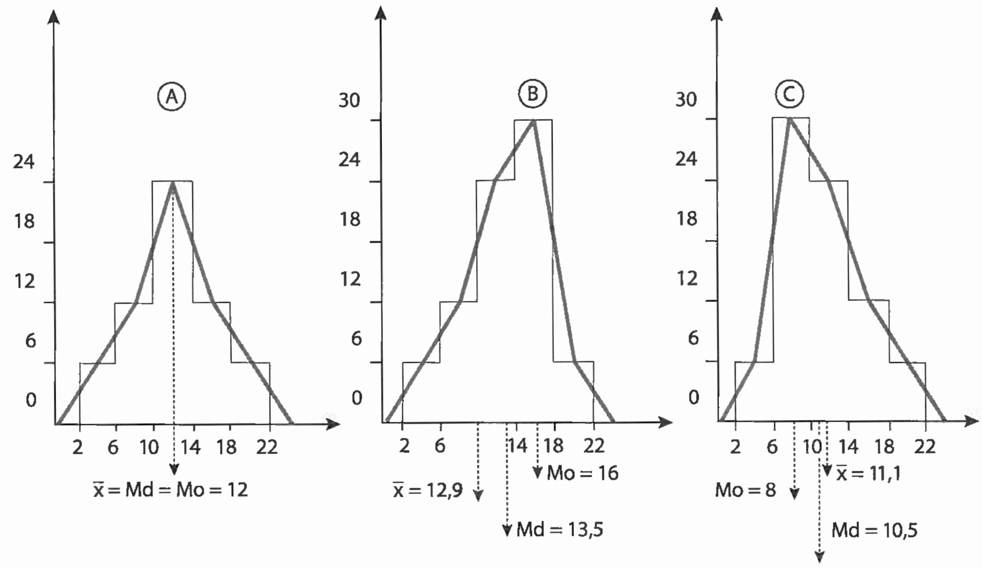
Logo:

A. 12 – 12 = 0 $\Rightarrow$ a distribuição é simétrica.

B. 12,9 – 16 = –3,1 kg $\Rightarrow$ a distribuição é assimétrica negativa.

C. 11,1 – 8 = 3,1 kg $\Rightarrow$ a distribuição é assimétrica positiva.

Considerando os gráficos das distribuições anteriores, temos:



8.1.2 Coeficiente de assimetria

A medida anterior, por ser absoluta, apresenta a mesma deficiência do desvio padrão, isto é, não permite a possibilidade de comparação entre as medidas de duas distribuições. Por esse motivo, daremos preferência ao **coeficiente de assimetria de Pearson**, dado por:

$$As = \frac{3(\bar{x} - Md)}{s}$$

Se $0{,}15 < |As| < 1$, a assimetria é considerada **moderada**; se $|As| > 1$, é **forte**.

Exemplo:

Considerando as distribuições **A**, **B** e **C** dadas anteriormente, temos:

$$As_A = \frac{3(12-12)}{4{,}42} = 0 \Rightarrow \text{simetria}$$

$$As_B = \frac{3(12{,}9-13{,}5)}{4{,}20} = -\ 0{,}429 \Rightarrow \text{assimetria negativa}$$

$$As_C = \frac{3(11{,}1-10{,}5)}{4{,}20} = 0{,}429 \Rightarrow \text{assimetria positiva}$$

Exercícios

1. Considere os seguintes resultados relativos a três distribuições de frequência:

DISTRIBUIÇÕES	$\bar{x}$	Mo
A	52	52
B	45	50
C	48	46

Determine o tipo de assimetria de cada uma delas.

2. Uma distribuição de frequência apresenta as seguintes medidas: $\bar{x} = 48{,}1$, Md $= 47{,}9$ e s $= 2{,}12$. Calcule o coeficiente de assimetria.

3. Em uma distribuição de frequência foram encontradas as seguintes medidas: $\bar{x} = 33{,}18$, Mo $= 27{,}50$, Md $= 31{,}67$ e s $= 12{,}45$.
   a. Classifique o tipo de assimetria.
   b. Calcule o coeficiente de assimetria.

4. Considerando a distribuição de frequência relativa aos pesos de cem operários de uma fábrica:

PESOS (kg)	50 ⊢ 58	58 ⊢ 66	66 ⊢ 74	74 ⊢ 82	82 ⊢ 90	90 ⊢ 98
Nº DE OPERÁRIOS	10	15	25	24	16	10

determine o grau de assimetria.

8.2 Curtose

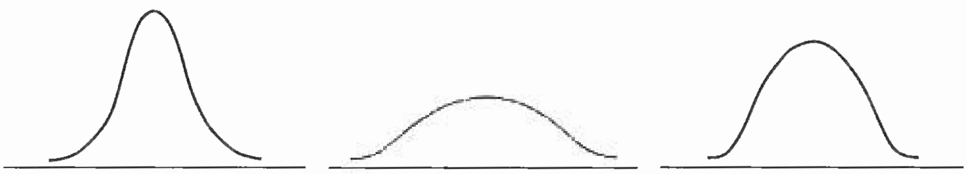
8.2.1 Introdução

Denominamos **curtose** o grau de achatamento de uma distribuição em relação a uma distribuição padrão, denominada **curva normal** (curva correspondente a uma distribuição teórica de probabilidade).

Quando a distribuição apresenta uma curva de frequência mais fechada que a normal (ou mais aguda em sua parte superior), ela recebe o nome de **leptocúrtica**.

Quando a distribuição apresenta uma curva de frequência mais aberta que a normal (ou mais achatada na sua parte superior), ela é chamada **platicúrtica**.

A curva normal, que é a nossa base referencial, recebe o nome de **mesocúrtica**.



8.2.2 Coeficiente de curtose

Uma fórmula para a medida da **curtose** é:

$$C = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}$$

Essa fórmula é conhecidda como **coeficiente percentílico de curtose**.

Relativamente à curva normal, temos:

$$\mathbf{C = 0,263}$$

Assim:

$\mathbf{C = 0,263} \Rightarrow$ curva mesocúrtica

$\mathbf{C < 0,263} \Rightarrow$ curva leptocúrtica

$\mathbf{C > 0,263} \Rightarrow$ curva platicúrtica

Exemplo:

Sabendo-se que uma distribuição apresenta as seguintes medidas:

$Q_1 = 24,4$ cm, $Q_3 = 41,2$ cm, $P_{10} = 20,2$ cm e $P_{90} = 49,5$ cm,

temos:

$$C = \frac{41,2 - 24,4}{2(49,5 - 20,2)} = \frac{16,8}{58,6} = 0,2866 \Rightarrow C = 0,287$$

Como:

$$0,287 > 0,263,$$

concluímos que a distribuição é **platicúrtica**, em relação à normal.

Exercícios

1. Considere as seguintes medidas, relativas a três distribuições de frequência:

DISTRIBUIÇÕES	Q_1	Q_3	P_{10}	P_{90}
A	814	935	772	1.012
B	63,7	80,3	55,0	86,6
C	28,8	45,6	20,5	49,8

a. Calcule os respectivos graus de curtose.

b. Classifique cada uma das distribuições em relação à curva normal.

2. Determine o grau de curtose e classifique a distribuição em relação à curva normal:

PESOS (kg)	50 ⊢ 58	58 ⊢ 66	66 ⊢ 74	74 ⊢ 82	82 ⊢ 90	90 ⊢ 98
Nº DE OPERÁRIOS	10	15	25	24	16	10

9 PROBABILIDADE

9.1 Introdução

Embora o cálculo das probabilidades pertença ao campo da Matemática, sua inclusão neste livro se justifica pelo fato de a maioria dos fenômenos de que trata a Estatística ser de natureza aleatória ou probabilística. Consequentemente, o conhecimento dos aspectos fundamentais do cálculo de probabilidades é uma necessidade essencial para o estudo da **Estatística Indutiva** ou **Inferencial**.

Procuramos resumir aqui os conhecimentos que julgamos necessários para termos um ponto de apoio em nossos primeiros passos no caminho da Estatística Inferencial. Esses passos serão apresentados no capítulo seguinte, que trata da conceituação de **variável aleatória** e das duas principais distribuições de probabilidades de **variáveis discretas e contínuas**.

9.2 Experimento aleatório

Em quase tudo, em maior ou menor grau, vislumbramos o **acaso**. Assim, da afirmação "é provável que o meu time ganhe a partida de hoje" pode resultar:

a. que, apesar do favoritismo, ele perca;

b. que, como pensamos, ele ganhe;

c. que empate.

Como vimos, o resultado final depende do **acaso**. Fenômenos como esse são chamados **fenômenos aleatórios** ou **experimentos aleatórios**.

Experimentos ou **fenômenos aleatórios** são aqueles que, mesmo repetidos várias vezes sob condições semelhantes, apresentam resultados imprevisíveis.

9.3 Espaço amostral

A cada experimento correspondem, em geral, vários resultados possíveis. Assim, ao lançarmos uma moeda, há dois resultados possíveis: ocorrer **cara** ou ocorrer **coroa**. Já ao lançarmos um dado há seis resultados possíveis: **1**, **2**, **3**, **4**, **5** ou **6**.

Ao conjunto desses resultados possíveis damos o nome de **espaço amostral** ou **conjunto universo**, representado por **S**.

Os dois experimentos citados anteriormente têm os seguintes espaços amostrais:

- lançamento de uma moeda: $S = \{Ca, Co\}$;
- lançamento de um dado: $S = \{1, 2, 3, 4, 5, 6\}$.

Do mesmo modo, como em **dois** lançamentos sucessivos de uma moeda podemos obter **cara** nos dois lançamentos, ou **cara** no primeiro e **coroa** no segundo, ou **coroa** no primeiro e **cara** no segundo, ou **coroa** nos dois lançamentos, o espaço amostral é:

$$S = \{(Ca, Ca), (Ca, Co), (Co, Ca), (Co, Co)\}.$$

Cada um dos elementos de **S** que corresponde a um resultado recebe o nome de **ponto amostral**. Assim:

$2 \in S \Rightarrow 2$ é um ponto amostral de **S**.

9.4 Eventos

Chamamos de **evento** qualquer subconjunto do espaço amostral **S** de um experimento aleatório.

Assim, qualquer que seja **E**, se **E** $\subset$ **S** (**E** está contido em **S**), então **E** é um evento de **S**.

Se **E** = **S**, **E** é chamado **evento certo**.

Se **E** $\subset$ **S** e **E** é um conjunto unitário, **E** é chamado **evento elementar.**

Se **E** = $\varnothing$, **E** é chamado **evento impossível**.

Exemplo:

No lançamento de um dado, onde S = $\{1, 2, 3, 4, 5, 6\}$, temos:

A = $\{2, 4, 6\} \subset$ S; logo, **A** é um evento de **S**.

B = $\{1, 2, 3, 4, 5, 6\} \subset$ S; logo, **B** é um evento certo de **S** (B = S).

C = $\{4\} \subset$ S; logo, **C** é um evento elementar de **S**.

D = $\varnothing \subset$ S; logo, **D** é um evento impossível de **S**.

Um evento é sempre definido por uma sentença. Assim, os eventos acima podem ser definidos pelas sentenças:

"Obter um número par na face superior."

"Obter um número menor ou igual a 6 na face superior."

"Obter o número 4 na face superior."

"Obter um número maior que 6 na face superior."

9.5 Probabilidade

Dado um experimento aleatório, sendo **S** o seu espaço amostral, vamos admitir que todos os elementos de **S** tenham a mesma chance de acontecer, ou seja, que **S** é um **conjunto equiprovável**.

Chamamos de **probabilidade de um evento A**(A $\subset$ S) o número real P(A), tal que:

$$P(A) = \frac{n(A)}{n(S)}$$

onde:

n(A) é o número de elementos de **A**;

n(S) é o número de elementos de **S**.

Exemplos:

a. Considerando o lançamento de uma moeda e o evento **A** "obter cara", temos:

$$S = \{Ca, Co\} \Rightarrow n(S) = 2$$

$$A = \{Ca\} \Rightarrow n(A) = 1$$

Logo:

$$P(A)=\frac{1}{2}$$

O resultado acima nos permite afirmar que, ao lançarmos uma moeda equilibrada, temos 50% de chance de que apareça **cara** na face superior.

b. Considerando o lançamento de um dado, vamos calcular:

- a probabilidade do evento **A** "obter um número par na face superior".

  Temos:

  $$S = \{1, 2, 3, 4, 5, 6\} \Rightarrow n(S) = 6$$
  $$A = \{2, 4, 6\} \Rightarrow n(A) = 3$$

  Logo:

  $$P(A)=\frac{3}{6}=\frac{1}{2}$$

- a probabilidade do evento **B** "obter um número menor ou igual a 6 na face superior".

  Temos:

  $$S = \{1, 2, 3, 4, 5, 6\} \Rightarrow n(S) = 6$$
  $$B = \{1, 2, 3, 4, 5, 6\} \Rightarrow n(B) = 6$$

  Logo:

  $$P(B)=\frac{6}{6}=1$$

- a probabilidade do evento **C** "obter um número 4 na face superior".

  Temos:

  $$S = \{1, 2, 3, 4, 5, 6\} \Rightarrow n(S) = 6$$
  $$C = \{4\} \Rightarrow n(C) = 1$$

  Logo:

  $$P(C)=\frac{1}{6}$$

- a probabilidade do evento **D** "obter um número maior que 6 na face superior".

  Temos:

  $$S = \{1, 2, 3, 4, 5, 6\} \Rightarrow n(S) = 6$$
  $$D = \varnothing \Rightarrow n(D) = 0$$

  Logo:

  $$P(D)=\frac{0}{6}=0$$

Pelos exemplos que acabamos de ver, podemos concluir que, sendo **n(S) = n**:

a. a probabilidade do **evento certo** é igual a 1:

$$P(S) = 1$$

b. a probabilidade do **evento impossível** é igual a zero:

$$P(\varnothing) = 0$$

c. a probabilidade de um **evento E qualquer ($E \subset S$)** é um número real P(E), tal que:

$$0 \leq P(E) \leq 1$$

d. a probabilidade de um **evento elementar E qualquer** é, lembrando que n(E) = 1:

$$P(E) = \frac{1}{n}$$

9.6 Eventos complementares

Sabemos que um evento pode ocorrer ou não. Sendo **p** a probabilidade de que ele ocorra (sucesso) e **q** a probabilidade de que ele não ocorra (insucesso), para um mesmo evento existe sempre a relação:

$$p + q = 1 \Rightarrow q = 1 - p$$

Assim, se a probabilidade de se realizar um evento é $p = \frac{1}{5}$, a probabilidade de que ele não ocorra é:

$$q = 1 - p \Rightarrow q = 1 - \frac{1}{5} = \frac{4}{5}$$

Sabemos que a probabilidade de tirar o 4 no lançamento de um dado é $p = \frac{1}{6}$. Logo, a probabilidade de não tirar o 4 no lançamento de um dado é:

$$q = 1 - \frac{1}{6} = \frac{5}{6}$$

9.7 Eventos independentes

Dizemos que **dois eventos** são **independentes** quando a realização ou a não realização de um dos eventos não afeta a probabilidade da realização do outro e vice-versa.

Por exemplo, quando lançamos dois dados, o resultado obtido em um deles independe do resultado obtido no outro.

Se dois eventos são independentes, a probabilidade de que eles se realizem **simultaneamente** é igual ao produto das probabilidades de realização dos dois eventos.

Assim, sendo $\mathbf{p_1}$ a probabilidade de realização do primeiro evento e $\mathbf{p_2}$ a probabilidade de realização do segundo evento, a probabilidade de que tais eventos se realizem simultaneamente é dada por:

$$p = p_1 \times p_2$$

Exemplo:

Lançamos dois dados. A probabilidade de obtermos 1 no primeiro dado é:

$$p_1 = \frac{1}{6}$$

A probabilidade de obtermos 5 no segundo dado é:

$$p_2 = \frac{1}{6}$$

Logo, a probabilidade de obtermos, simultaneamente, 1 no primeiro e 5 no segundo é:

$$p = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

9.8 Eventos mutuamente exclusivos

Dizemos que **dois ou mais eventos** são **mutuamente exclusivos** quando a realização de um exclui a realização do(s) outro(s).

Assim, no lançamento de uma moeda, o evento "tirar cara" e o evento "tirar coroa" são mutuamente exclusivos, já que, ao se realizar um deles, o outro não se realiza.

Se dois eventos são mutuamente exclusivos, a probabilidade de que um **ou** outro se realize é igual à soma das probabilidades de que cada um deles se realize:

$$p = p_1 + p_2$$

Exemplo:

Lançamos um dado. A probabilidade de se tirar o 3 **ou** o 5 é:

$$p = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3},$$

pois, como vimos, os dois eventos são mutuamente exclusivos.

Exercícios resolvidos

1. Qual a probabilidade de sair o ás de ouros quando retiramos uma carta de um baralho de 52 cartas?

   Como só há um ás de ouros, o número de elementos do evento é 1; logo:

$$p = \frac{1}{52}$$

2. Qual a probabilidade de sair um rei quando retiramos uma carta de um baralho de 52 cartas?

   Como há 4 reis, o número de elementos do evento é 4; logo:

$$p = \frac{4}{52} = \frac{1}{13}$$

3. Em um lote de 12 peças, quatro são defeituosas. Sendo retirada uma peça, calcule:

   a. a probabilidade de essa peça ser defeituosa.

   Temos:

$$p = \frac{4}{12} = \frac{1}{3}$$

   b. a probabilidade de essa peça não ser defeituosa.

   Sendo este evento e o anterior complementares, temos:

$$p = 1 - \frac{1}{3} = \frac{2}{3}$$

4. No lançamento de dois dados, calcule a probabilidade de se obter soma igual a 5.

   O evento é formado pelos elementos (1, 4), (2, 3), (3, 2) e (4, 1). Como o número de elementos de **S** é 36, temos:

$$p = \frac{4}{36} = \frac{1}{9}$$

5. De dois baralhos de 52 cartas retiram-se, simultaneamente, uma carta do primeiro baralho e uma carta do segundo. Qual a probabilidade de a carta do primeiro baralho ser um rei e a do segundo ser o 5 de paus?

   Temos:

$$p_1 = \frac{4}{52} = \frac{1}{13}$$

   e

$$p_2 = \frac{1}{52}$$

   Como esses dois acontecimentos são independentes e simultâneos, vem:

$$p_1 = \frac{1}{13} \times \frac{1}{52} \times \frac{1}{676}$$

6. Uma urna **A** contém: três bolas brancas, quatro pretas, duas verdes; uma urna **B** contém: cinco bolas brancas, duas pretas, uma verde; uma urna **C** contém: duas bolas brancas, três pretas, quatro verdes. Uma bola é retirada de cada urna. Qual é a probabilidade de as três bolas retiradas da primeira, segunda e ter-

ceira urnas serem, respectivamente, branca, preta e verde?

Temos:

$$p_1 = \frac{3}{9} = \frac{1}{3},\ p_2 = \frac{2}{8} = \frac{1}{4},\ p_3 = \frac{4}{9}$$

Como os três eventos são independentes e simultâneos, vem:

$$p = \frac{1}{3} \times \frac{1}{4} \times \frac{4}{9} = \frac{1}{27}$$

7. De um baralho de 52 cartas retiram-se, ao acaso, duas cartas sem reposição. Qual é a probabilidade de a primeira carta ser o ás de paus e a segunda ser o rei de paus?

A probabilidade de sair o ás de paus na primeira carta é:

$$p_1 = \frac{1}{52}$$

Após a retirada da primeira carta, restam 51 cartas no baralho, já que a carta retirada não foi reposta. Assim, a probabilidade de a segunda carta ser o rei de paus é:

$$p_2 = \frac{1}{51}$$

Como esses eventos são independentes, temos:

$$p = \frac{1}{52} \times \frac{1}{51} \times \frac{1}{2.652}$$

8. Qual a probabilidade de sair uma figura quando retiramos uma carta de um baralho de 52 cartas?

Temos:

$$p_r = \frac{4}{52} = \frac{1}{13},\ p_d = \frac{1}{13},\ p_v = \frac{1}{13}$$

Como os eventos são mutuamente exclusivos, vem:

$$p = \frac{1}{13} + \frac{1}{13} + \frac{1}{13} = \frac{3}{13}$$

> **NOTA:**
> - Este problema pode ser resolvido, ainda, com o seguinte raciocínio: como em um baralho temos 12 figuras (quatro damas, quatro valetes, quatro reis), vem:
>
> $$p = \frac{12}{52} = \frac{3}{13}$$

9. Qual a probabilidade de sair uma carta de copas ou de ouros quando retiramos uma carta de um baralho de 52 cartas?

Temos:

$$p_c = \frac{13}{52} = \frac{1}{4},\ p_o = \frac{13}{52} = \frac{1}{4}$$

Como os eventos são mutuamente exclusivos, vem:

$$p = \frac{1}{4} + \frac{1}{4} = \frac{2}{4} = \frac{1}{2}$$

10. No lançamento de um dado, qual a probabilidade de se obter um número não inferior a 5?

A probabilidade de se ter um número não inferior a 5 é a probabilidade de se obter 5 ou 6. Assim:

$$p = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

11. São dados dois baralhos de 52 cartas. Tiramos, ao mesmo tempo, uma carta do primeiro baralho e uma carta do segundo. Qual é a probabilidade de tirarmos uma dama e um rei, não necessariamente nessa ordem?

A probabilidade de tirarmos uma dama do primeiro baralho $\left(\frac{4}{52}\right)$ e um rei do segundo $\left(\frac{4}{52}\right)$ é, de acordo com o problema 7:

$$p_1 = \frac{4}{52} \times \frac{4}{52} = \frac{1}{13} \times \frac{1}{13} = \frac{1}{169}$$

A probabilidade de tirarmos um rei do primeiro baralho e uma dama do segundo é:

$$p_2 = \frac{4}{52} \times \frac{4}{52} = \frac{1}{169}$$

Como esses dois eventos são mutuamente exclusivos, temos:

$$p = \frac{1}{169} + \frac{1}{169} = \frac{2}{169}$$

12. Dois dados são lançados conjuntamente. Determine a probabilidade de a soma ser 10 ou maior que 10.

A soma deverá ser, então, 10, 11 ou 12.

Para que a soma seja 10, a probabilidade é:

$$\left.\begin{matrix}(4, 6)\\(5, 5)\\(6, 4)\end{matrix}\right\} \Rightarrow n(10) = 3 \Rightarrow p_{10} = \frac{3}{36}$$

Para que a soma seja 11, a probabilidade é:

$$\left.\begin{matrix}(5, 6)\\(6, 5)\end{matrix}\right\} \Rightarrow n(11) = 2 \Rightarrow p_{11} = \frac{2}{36}$$

Para que a soma seja 12, a probabilidade é:

$$(6, 6)\} \Rightarrow n(12) = 1 \Rightarrow p_{12} = \frac{1}{36}$$

Como esses três eventos são mutuamente exclusivos, temos:

$$p = \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{6}{36} = \frac{1}{6}$$

Exercícios

1. Determine a probabilidade de cada evento:

a. Um número par aparece no lançamento de um dado.

b. Uma figura aparece ao se extrair uma carta de um baralho de 52 cartas.

c. Uma carta de ouros aparece ao se extrair uma carta de um baralho de 52 cartas.

d. Uma só coroa aparece no lançamento de três moedas.

2. Um número inteiro é escolhido aleatoriamente dentre os números 1, 2, 3, ..., 49, 50. Determine a probabilidade de:

a. o número ser divisível por 5;

b. o número terminar em 3;

c. o número ser divisível por 6 ou por 8;

d. o número ser divisível por 4 e por 6.

3. Dois dados são lançados simultaneamente. Determine a probabilidade de:

a. a soma ser menor que 4;

b. a soma ser 9;

c. o primeiro resultado ser maior que o segundo;

d. a soma ser menor ou igual a 5.

4. Uma moeda é lançada duas vezes. Calcule a probabilidade de:

a. não ocorrer cara nenhuma vez;

b. obter-se cara na primeira ou na segunda jogada.

5. Um inteiro entre 3 e 11 será escolhido ao acaso.
   a. Qual é a probabilidade de que este número seja ímpar?
   b. Qual é a probabilidade de que este número seja ímpar e divisível por 3?

6. Uma carta é retirada ao acaso de um baralho de 52 cartas. Qual a probabilidade de que a carta retirada seja uma dama ou uma carta de copas?

7. No lançamento de dois dados, qual é a probabilidade de se obter um par de pontos iguais?

8. Em um lote de 12 peças, quatro são defeituosas. Sendo retiradas aleatoriamente duas peças, calcule:
   a. a probabilidade de ambas serem defeituosas;
   b. a probabilidade de ambas não serem defeituosas;
   c. a probabilidade de ao menos uma ser defeituosa.

9. No lançamento de um dado, qual é a probabilidade de sair o numero 6 ou um número ímpar?

10. Duas cartas são retiradas ao acaso de um baralho de 52 cartas. Calcule a probabilidade de se obter:
    a. dois valetes;
    b. um valete e uma dama.

11. Um casal planeja ter três filhos. Determine a probabilidade de nascer:
    a. três homens;
    b. dois homens e uma mulher.

12. Uma moeda é lançada três vezes. Calcule a probabilidade de obtermos:
    a. três caras;
    b. duas caras e uma coroa;
    c. uma cara somente;
    d. nenhuma cara;
    e. pelo menos uma cara;
    f. no máximo uma cara.

13. Um dado é lançado duas vezes. Calcule a probabilidade de:
    a. sair um 6 no primeiro lançamento;
    b. sair um 6 no segundo lançamento;
    c. não sair 6 em nenhum lançamento;
    d. sair um 6 pelo menos.

14. Uma urna contém 50 bolas idênticas. Sendo as bolas numeradas de 1 a 50, determine a probabilidade de, em uma extração ao acaso:
    a. obtermos a bola de número 27;
    b. obtermos uma bola de número par;
    c. obtermos uma bola de número maior que 20;
    d. obtermos uma bola de número menor ou igual a 20.

15. Uma loja dispõe de 12 geladeiras do mesmo tipo, das quais quatro apresentam defeitos.
    a. Se um freguês vai comprar uma geladeira, qual a probabilidade de levar uma defeituosa?
    b. Se um freguês vai comprar duas geladeiras, qual a probabilidade de levar duas defeituosas?

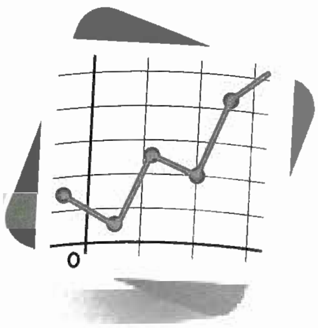
c. Se um freguês vai comprar duas geladeiras, qual a probabilidade de levar pelo menos uma defeituosa?

16. Um par de dados é atirado. Encontre a probabilidade de que a soma seja 10 ou maior que 10 se:
   a. um 5 aparece no primeiro dado;
   b. um 5 aparece pelo menos em um dos dados.

17. Lança-se um par de dados. Aparecendo dois números diferentes, encontre a probabilidade de que:
   a. a soma seja 6;
   b. o 1 apareça;
   c. a soma seja 4 ou menor que 4.

18. Um lote é formado por dez peças boas, quatro com defeitos e duas com defeitos graves. Uma peça é escolhida ao acaso. Calcule a probabilidade de que:
   a. ela não tenha defeitos graves;
   b. ela não tenha defeitos;
   c. ela seja boa ou tenha defeitos graves.

19. Considere o mesmo lote do problema anterior. Retiram-se duas peças ao acaso. Calcule a probabilidade de que:
   a. ambas sejam perfeitas;
   b. pelo menos uma seja perfeita;
   c. nenhuma tenha defeitos graves;
   d. nenhuma seja perfeita.



10 DISTRIBUIÇÕES BINOMIAL E NORMAL

O que pretendemos, neste capítulo, é apresentar dois modelos teóricos de distribuição de probabilidade, aos quais um experimento aleatório estudado possa ser adaptado, o que permitirá a solução de grande número de problemas práticos.

10.1 Variável aleatória

Suponhamos um espaço amostral **S** e que a cada ponto amostral seja atribuído um número. Fica, então, definida uma função chamada **variável aleatória**, indicada por uma letra maiúscula, sendo seus valores indicados por letras minúsculas.

Assim, se o espaço amostral relativo ao "lançamento simultâneo de duas moedas" é S = {(Ca, Ca), (Ca, Co), (Co, Ca), (Co, Co)} e se **X** representa "o número de caras" que aparecem, a cada ponto amostral podemos associar um número para **X**, de acordo com a Tabela 10.1:

PONTO AMOSTRAL	X
(Ca, Ca)	2
(Ca, Co)	1
(Co, Ca)	1
(Co, Co)	0

TABELA 10.1

10.2 Distribuição de probabilidade

Consideremos a distribuição de frequências relativa ao número de acidentes diários em um estacionamento:

NÚMERO DE ACIDENTES	FREQUÊNCIAS
0	22
1	5
2	2
3	1
	$\Sigma = 30$

TABELA 10.2

Em um dia, a probabilidade de:

- não ocorrer acidente é:

$$p = \frac{22}{30} = 0,73$$

- ocorrer um acidente é:

$$p = \frac{5}{30} = 0,17$$

- ocorrerem dois acidentes é:

$$p = \frac{2}{30} = 0,07$$

- ocorrerem três acidentes é:

$$p = \frac{1}{30} = 0,03$$

Podemos, então, escrever:

NÚMERO DE ACIDENTES	PROBABILIDADES
0	0,73
1	0,17
2	0,07
3	0,03
	$\Sigma = 1,00$

TABELA 10.3

Essa tabela é denominada **distribuição de probabilidade**.

Seja **X** uma variável aleatória que pode assumir os valores $x_1, x_2, x_3, ..., x_n$. A cada valor x_i correspondem pontos do espaço amostral. Associamos, então, a cada valor x_i a probabilidade p_i de ocorrência de tais pontos no espaço amostral.

Assim, temos:

$$\sum p_i = 1$$

Os valores $x_1, x_2, ..., x_n$ e seus correspondentes $p_1, p_2, ..., p_n$ definem uma **distribuição de probabilidade.**

Assim, voltando à Tabela 10.1, temos:

PONTO AMOSTRAL	X	P(X)
(Ca, Ca)	2	$1/2 \times 1/2 = 1/4$
(Ca, Co)	1	$1/2 \times 1/2 = 1/4$ } $1/4 + 1/4 = 2/4$
(Co, Ca)	1	$1/2 \times 1/2 = 1/4$ }
(Co, Co)	0	$1/2 \times 1/2 = 1/4$

TABELA 10.4

Logo, podemos escrever:

NÚMERO DE CARAS (X)	P(X)
2	1/4
1	2/4
0	1/4
	$\Sigma = 1$

TABELA 10.5

Ao definir a distribuição de probabilidade, estabelecemos uma correspondência unívoca entre os valores da variável aleatória **X** e os valores da variável **P**. Esta correspondência define uma **função**; os valores x_i (i = 1, 2, ..., n) formam o **domínio da função** e os valores p_i (i = 1, 2, 3, ..., n), o seu **conjunto imagem**.

Essa função, assim definida, é denominada **função probabilidade** e representada por:

$$f(x) = P(X = x_i)$$

A função $P(X = x_i)$ determina a **distribuição de probabilidade da variável aleatória X**.

Assim, ao lançarmos um dado, a variável aleatória **X**, definida por "pontos de um dado", pode tomar os valores 1, 2, 3, ..., 6.

Como a cada um destes valores está associada uma e uma só probabilidade de realização e $\sum P(x_i) = 1$, fica definida uma função de probabilidade, da qual resulta a seguinte **distribuição de probabilidade**:

X	P(X)
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6
	$\Sigma = 1$

TABELA 10.6

10.3 Distribuição binomial

Vamos, neste item, considerar experimentos que satisfaçam as seguintes condições:

a. O experimento deve ser repetido, nas mesmas condições, um número finito de vezes (**n**).

b. As provas repetidas devem ser independentes, isto é, o resultado de uma não deve afetar os resultados das sucessivas.

c. Em cada prova deve aparecer um dos dois possíveis resultados: **sucesso** e **insucesso**.

d. No decorrer do experimento, a probabilidade **p** do sucesso e a probabilidade **q** (q = 1 – p) do insucesso manter-se-ão constantes.

Resolveremos problemas do tipo: determinar a probabilidade de se obter **k** sucessos em **n** tentativas.

O experimento "obtenção de caras em cinco lançamentos sucessivos e independentes de uma moeda" satisfaz essas condições.

Sabemos que, quando da realização de um experimento qualquer em uma única tentativa, se a probabilidade de realização de um evento (sucesso) é **p**, a probabilidade de não realização desse mesmo evento (insucesso) é $\mathbf{1 - p = q}$.

Suponhamos, agora, que realizemos a mesma prova **n** vezes sucessivas e independentes. A probabilidade de que um evento se realize **k** vezes nas provas é dada pela função:

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

na qual:

$\mathbf{P(X = k)}$ é a probabilidade de que o evento se realize **k** vezes em **n** provas;

p é a probabilidade de que o evento se realize em uma só prova — **sucesso**;

q é a probabilidade de que o evento não se realize no decurso dessa prova — **insucesso**;

$\binom{\mathbf{n}}{\mathbf{k}}$ é o coeficiente binomial de **n** sobre **k**, igual a $\frac{n!}{k!(n-k)!}$[1].

Essa função, denominada **lei binomial**, define a **distribuição binomial**.

NOTA:

- O nome **binomial** vem do fato de $\binom{n}{k} p^k q^{n-k}$ ser o termo geral do desenvolvimento do **binômio de Newton**.

[1] n! é o fatorial de **n**. Consulte o **Apêndice — Instrumental Matemático**, para revisão do assunto **Fatorial** (p. 190).

Exercícios resolvidos

1. Uma moeda é lançada cinco vezes seguidas e independentes. Calcule a probabilidade de serem obtidas três caras nessas cinco provas.

   Temos:

   $$n = 5 \text{ e } k = 3$$

   Pela lei binomial, podemos escrever:

   $$P(X = 3) = \binom{5}{3} p^3 q^{5-3} = \binom{5}{3} p^3 q^2$$

   Se a probabilidade de obtermos "cara" numa só prova (sucesso) é $p = \frac{1}{2}$ e a probabilidade de não obtermos "cara" numa só prova (insucesso) é $q = 1 - \frac{1}{2} = \frac{1}{2}$, então:

   $$P(X = 3) = \binom{5}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 = \frac{5!}{3!2!} \times \frac{1}{8} \times \frac{1}{4} =$$

   $$= \frac{5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1 \times 2 \times 1} \times \frac{1}{8} \times \frac{1}{4} =$$

   $$= \frac{5}{16}$$

   Logo:

   $$P(X = 3) = \frac{5}{16}$$

2. Dois times de futebol, **A** e **B**, jogam entre si seis vezes. Encontre a probabilidade de o time **A** ganhar quatro jogos.

   Temos:

   $$n = 6,\ k = 4,\ p = \frac{1}{3} \text{ e } q = 1 - \frac{1}{3} = \frac{2}{3}$$

   Então:

   $$P(X = 4) = \binom{6}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2 =$$

   $$= \cancel{15}^{5} \times \frac{1}{81} \times \frac{4}{\cancel{9}_{3}} = \frac{20}{243}$$

   Logo:

   $$P(X = 4) = \frac{20}{243}$$

Exercícios

1. Determine a probabilidade de obtermos exatamente três caras em seis lances de uma moeda.
2. Jogando-se um dado três vezes, determine a probabilidade de se obter um múltiplo de 3 duas vezes.
3. Dois times de futebol, **A** e **B**, jogam entre si seis vezes. Encontre a probabilidade de o time **A**:
   a. ganhar dois ou três jogos;
   b. ganhar pelo menos um jogo.

4. A probabilidade de um atirador acertar o alvo é $\frac{2}{3}$. Se ele atirar cinco vezes, qual a probabilidade de acertar exatamente dois tiros?

5. Seis parafusos são escolhidos ao acaso da produção de certa máquina, que apresenta 10% de peças defeituosas. Qual a probabilidade de serem defeituosos dois deles?

10.4 Distribuição normal. Curva normal

Entre as distribuições teóricas de variável aleatória contínua, uma das mais empregadas é a **distribuição normal**.

Muitas das variáveis analisadas na pesquisa socioeconômica correspondem à distribuição normal ou dela se aproximam.

O aspecto gráfico de uma distribuição normal é o da Figura 10.1:

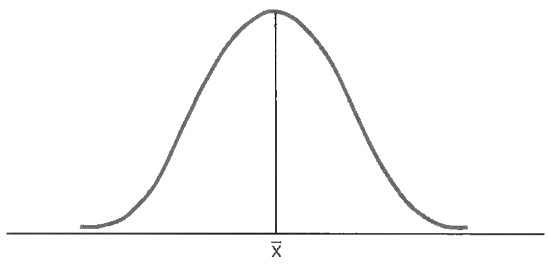


FIGURA 10.1

Para uma perfeita compreensão da distribuição normal, observe a Figura 10.1 e procure visualizar as seguintes propriedades:

1ª) A variavel aleatória **X** pode assumir todo e qualquer valor real.

2ª) A representação gráfica da distribuição normal é uma curva em forma de sino, simétrica em torno da média ($\bar{x}$), que recebe o nome de **curva normal** ou **de Gauss**.

3ª) A área total limitada pela curva e pelo eixo das abscissas é igual a 1, já que essa área corresponde à probabilidade de a variável aleatória **X** assumir qualquer valor real.

4ª) A curva normal é assintótica em relação ao eixo das abscissas, isto é, aproxima-se indefinidamente do eixo das abscissas sem, contudo, alcançá-lo.

5ª) Como a curva é simétrica em torno de $\bar{x}$, a probabilidade de ocorrer valor maior do que a média é igual à probabilidade de ocorrer valor menor do que a média, isto é, ambas as probabilidades são iguais a 0,5. Escrevemos: $P(X > \bar{x}) = P(X < \bar{x}) = 0{,}5$.

Quando temos em mãos uma variável aleatória com distribuição normal, nosso principal interesse é obter a probabilidade de essa variável aleatória assumir um valor em um determinado intervalo. Vejamos como proceder, por meio de um exemplo concreto.

Seja **X** a variável aleatória que representa os diâmetros dos parafusos produzidos por certa máquina. Vamos supor que essa variável tenha distribuição normal com média $\bar{x}$ = 2 cm e desvio padrão s = 0,04 cm.

Pode haver interesse em conhecer a probabilidade de um parafuso ter um diâmetro com valor entre 2 e 2,05 cm.

É fácil notar que essa probabilidade, indicada por:

$$P(2 < X < 2{,}05),$$

corresponde à área hachurada na Figura 10.2:

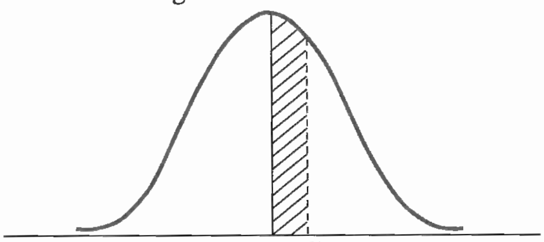


FIGURA 10.2

O cálculo direto dessa probabilidade exige um conhecimento de Matemática mais avançado do que aquele que dispomos no curso de 2º grau. Entretanto, podemos contornar facilmente esse problema. Basta aceitar, sem demonstração, que, se **X** é uma variável aleatória com distribuição normal de média $\bar{\mathbf{x}}$ e desvio padrão **s**, então a variável:

$$z = \frac{x - \bar{x}}{s}$$

tem **distribuição normal reduzida**, isto é, tem distribuição normal de média 0 e desvio padrão 1.

As probabilidades associadas à distribuição normal padronizada são encontradas em tabelas, não havendo necessidade de serem calculadas.

O **anexo II** (p. 218) é uma tabela de distribuição normal reduzida, que nos dá a probabilidade de **Z** tomar qualquer valor entre a média 0 e um dado valor **z**, isto é:

$$P(0 < Z < z).$$

Temos, então, que se **X** é uma variável aleatória com distribuição normal de média $\bar{\mathbf{x}}$ e desvio padrão **s**, podemos escrever:

$$P(\bar{x} < X < x) = P(0 < Z < z),$$

com $z = \dfrac{x - \bar{x}}{s}$.

Voltemos, então, ao nosso problema.

Queremos calcular P(2 < X < 2,05). Para obter essa probabilidade, precisamos, em primeiro lugar, calcular o valor de **z** que corresponde a x = 2,05 (x = 2 ⇒ z = 0, pois $\overline{x} = 2$). Temos, então:

$$z = \frac{x - \overline{x}}{s} = \frac{2{,}05 - 2}{0{,}04} = \frac{0{,}05}{0{,}04} = 1{,}25,$$

donde:

$$P(2 < X < 2{,}05) = P(0 < X < 1{,}25).$$

Procuremos, agora, no **anexo II** o valor de z = 1,25.

Na primeira coluna encontramos o valor **1,2**. Em seguida, encontramos, na primeira linha, o valor **5**, que corresponde ao último algarismo do número **1,25**. Na intersecção da linha e coluna correspondentes encontramos o valor **0,3944**, o que nos permite escrever:

$$P(0 < Z < 1{,}25) = 0{,}3944.$$

Assim, a probabilidade de um parafuso fabricado por essa máquina apresentar um diâmetro entre a média $\overline{x} = 2$ e o valor x = 2,05 é 0,3944.

Escrevemos, então:

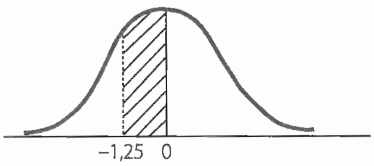
$$P(2 < X < 2{,}05) = P(0 < Z < 1{,}25) = 0{,}3944 \text{ ou } 39{,}44\%.$$

Exercícios resolvidos

1. Determine as probabilidades:

   a. P(–1,25 < Z < 0)

   A probabilidade procurada correspon-de à parte hachurada da figura:



   Sabemos que:

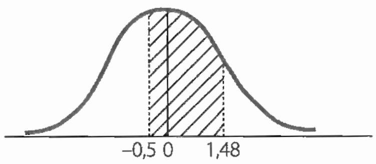
   P(0 < Z < 1,25) = 0,3944

   Pela simetria da curva, temos:

   P(– 1,25 < Z < 0) = P(0 < Z < 1,25) = 0,3944

   b. P(–0,5 < Z < 1,48)

   A probabilidade procurada correspon-de à parte hachurada da figura:



   Temos:

   P(– 0,5 < Z < 1,48) = P(– 0,5 < Z < 0) + P(0 < Z < 1,48)

   Como:

   P(– 0,5 < Z < 0) = P(0 < Z < 0,5) = 0,1915 e

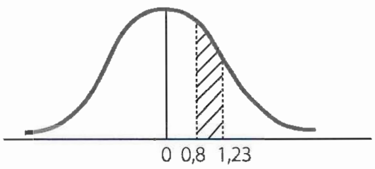
$P(0 < Z < 1{,}48) = 0{,}4306$,

obtemos:

$P(-0{,}5 < Z < 1{,}48) = 0{,}1915 + 0{,}4306 = 0{,}6221$

c. $P(0{,}8 < Z < 1{,}23)$

A probabilidade procurada corresponde à parte hachurada da figura:



Temos:

$P(0{,}8 < Z < 1{,}23) = P(0 < Z < 1{,}23) - P(0 < Z < 0{,}8)$

Como:

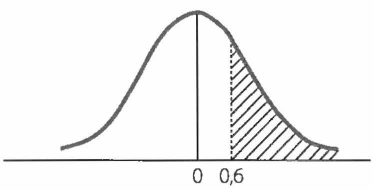
$P(0 < Z < 1{,}23) = 0{,}3907$ e $P(0 < Z < 0{,}8) = 0{,}2881$,

obtemos:

$P(0{,}8 < Z < 1{,}23) = 0{,}3907 - 0{,}2881 = 0{,}1026$

d. $P(Z > 0{,}6)$

A probabilidade procurada corresponde à parte hachurada da figura:



Temos:

$P(Z > 0{,}6) = P(Z > 0) - P(0 < Z < 0{,}6)$

Como:

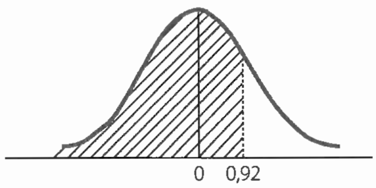
$P(Z > 0) = 0{,}5$ e $P(0 < Z < 0{,}6) = 0{,}2258$,

obtemos:

$P(Z > 0{,}6) = 0{,}5 - 0{,}2258 = 0{,}2742$

e. $P(Z < 0{,}92)$

A probabilidade procurada corresponde à parte hachurada da figura:



Temos:

$P(Z < 0{,}92) = P(Z < 0) + P(0 < Z < 0{,}92)$

Como:

$P(Z < 0) = 0{,}5$ e $P(0 < Z < 0{,}92) = 0{,}3212$,

obtemos:

$P(Z < 0{,}92) = 0{,}5 + 0{,}3212 = 0{,}8212$

2. Os salários semanais dos operários industriais são distribuídos normalmente, em torno da média de R$ 500, com desvio padrão de R$ 40. Calcule a probabilidade de um operário ter um salário semanal situado entre R$ 490 e R$ 520.

Devemos, inicialmente, determinar os valores da variável de distribuição normal reduzida.

Assim:

$$z_1 = \frac{490 - 500}{40} = -0{,}25$$

e

$$z_2 = \frac{520 - 500}{40} = 0{,}5$$

Logo, a probabilidade procurada é dada por:

$P(490 < X < 520) = P(-0{,}25 < Z < 0{,}5) = P(-0{,}25 < Z < 0) + P(0 < Z < 0{,}5) = 0{,}0987 + 0{,}1915 = 0{,}2902$

É, pois, de se esperar que, em média, 29,02% dos operários tenham salários entre R$ 490 e R$ 520.

Exercícios

1. Sendo **Z** uma variável com distribuição normal reduzida, calcule:

a. $P(0 < Z < 1{,}44)$
b. $P(-0{,}85 < Z < 0)$
c. $P(-1{,}48 < Z < 2{,}05)$
d. $P(0{,}72 < Z < 1{,}89)$
e. $P(Z > -2{,}03)$
f. $P(Z > 1{,}08)$
g. $P(Z < -0{,}66)$
h. $P(Z < 0{,}60)$

2. Um teste padronizado de escolaridade tem distribuição normal com média 100 e desvio padrão 10. Determine a probabilidade de um indivíduo submetido ao teste ter nota:

a. maior que 120;
b. maior que 80;
c. entre 85 e 115;
d. maior que 100.

3. Os pesos de 600 estudantes são normalmente distribuídos com média 65,3 kg e desvio padrão 5,5 kg. Determine o número de estudantes que pesam:

a. entre 60 e 70 kg;
b. mais que 63,2 kg;
c. menos que 68 kg.

4. A duração de um certo componente eletrônico tem média de 850 dias e desvio padrão de 40 dias. Sabendo que a duração é normalmente distribuída, calcule a probabilidade de esse componente durar:

a. entre 700 e 1.000 dias;
b. mais de 800 dias;
c. menos de 750 dias.

11 CORRELAÇÃO E REGRESSÃO

11.1 Introdução

Nos capítulos anteriores, nossa preocupação era descrever a distribuição de valores de uma única variável. Com esse objetivo, aprendemos a calcular medidas de tendência central e variabilidade.

Quando, porém, consideramos observações de duas ou mais variáveis, surge um novo problema: as **relações** que podem existir entre as variáveis estudadas. Nesse caso, as medidas estudadas não são eficientes.

Assim, quando consideramos variáveis como peso e altura de um grupo de pessoas, uso do cigarro e incidência do câncer, vocabulário e compreensão da leitura, dominância e submissão, procuramos verificar se existe alguma relação entre as variáveis de cada um dos pares e qual o grau dessa relação. Para isso, é necessário o conhecimento de novas medidas.

Sendo a relação entre as variáveis de natureza **quantitativa**, a **correlação** é o instrumento adequado para descobrir e medir essa relação.

Uma vez caracterizada a relação, procuramos descrevê-la através de uma função matemática. A **regressão** é o instrumento adequado para a determinação dos parâmetros dessa função.

NOTA:
- Ficaremos restritos às relações entre duas variáveis (correlação simples).

11.2 Correlação

11.2.1 Relação funcional e relação estatística

Como sabemos, o perímetro e o lado de um quadrado estão relacionados. A relação que os liga é perfeitamente definida e pode ser expressa por meio de uma sentença matemática:

$$2p = 4\ell,$$

onde 2p é o perímetro e ℓ é o lado.

Atribuindo-se, então, um valor qualquer a ℓ, é possível determinar **exatamente** o valor de 2p.

Consideremos, agora, a relação que existe entre o peso e a estatura de um grupo de pessoas. É evidente que essa relação não é do mesmo tipo da anterior; ela é bem menos precisa. Assim, pode acontecer que a estaturas diferentes correspondam pesos iguais ou que a estaturas iguais correspondam pesos diferentes. Contudo, em média, quanto maior a estatura, maior o peso.

As relações do tipo perímetro – lado são conhecidas como **relações funcionais** e as do tipo peso – estatura, como **relações estatísticas**.

Quando duas variáveis estão ligadas por uma **relação estatística**, dizemos que existe **correlação** entre elas.

NOTA:
- As relações funcionais são um caso limite das relações estatísticas.

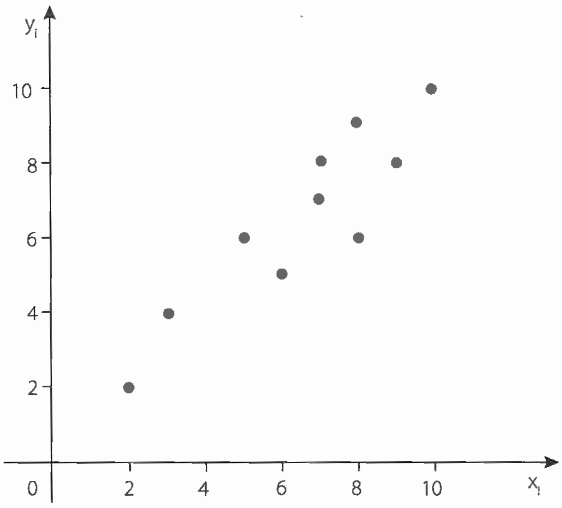
11.2.2 Diagrama de dispersão

Consideremos uma amostra aleatória, formada por dez dos 98 alunos de uma classe da faculdade **A** e pelas notas obtidas por eles em Matemática e Estatística:

Nºs	NOTAS	
	MATEMÁTICA (x_i)	ESTATÍSTICA (y_i)
01	5,0	6,0
08	8,0	9,0
24	7,0	8,0
38	10,0	10,0
44	6,0	5,0
58	7,0	7,0
59	9,0	8,0
72	3,0	4,0
80	8,0	6,0
92	2,0	2,0

TABELA 11.1

Representando, em um sistema coordenado cartesiano ortogonal, os pares ordenados (x_i, y_i), obtemos uma nuvem de pontos que denominamos **diagrama de dispersão**. Esse diagrama nos fornece uma ideia grosseira, porém útil, da correlação existente.

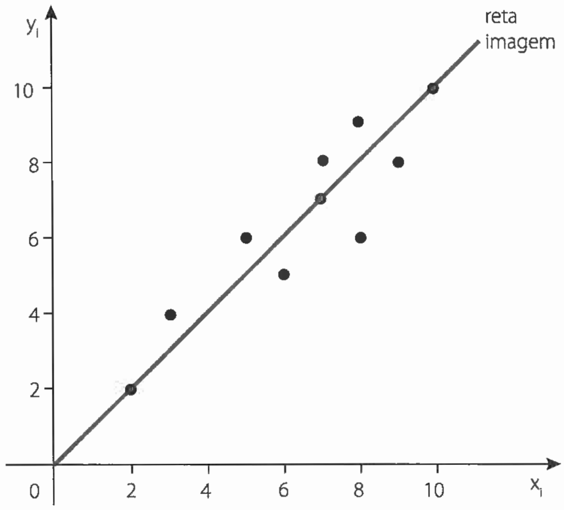


11.2.3 Correlação linear

Os pontos obtidos, vistos em conjunto, formam uma elipse em diagonal.

Podemos imaginar que, quanto mais fina for a elipse, mais ela se aproximará de uma reta. Dizemos, então, que a correlação de forma elíptica tem como "imagem" uma reta, sendo, por isso, denominada **correlação linear**.

É possível verificar que a cada correlação está associada como "imagem" uma relação funcional. Por esse motivo, as relações funcionais são chamadas **relações perfeitas**.



Como a correlação em estudo tem como "imagem" uma reta ascendente, ela é chamada **correlação linear positiva**.

Assim, uma correlação é:

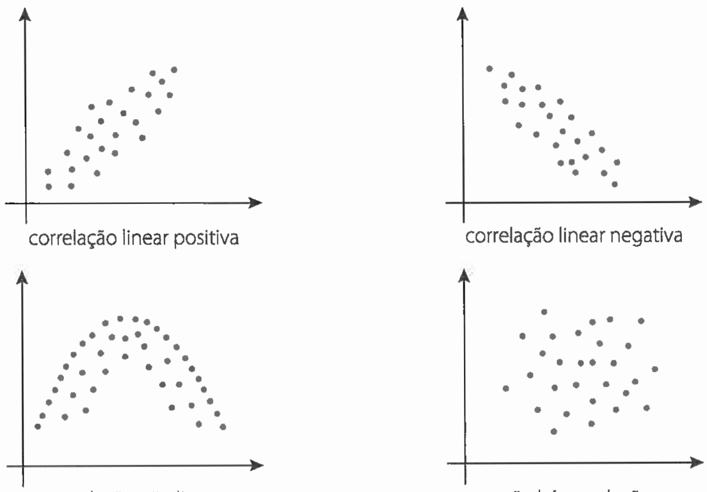
a. **linear positiva** se os pontos do diagrama têm como "imagem" uma reta ascendente;

b. **linear negativa** se os pontos têm como "imagem" uma reta descendente;

c. **não linear** se os pontos têm como "imagem" uma curva.

Se os pontos apresentam-se dispersos, não oferecendo uma "imagem" definida, concluímos que não há relação alguma entre as variáveis em estudo.

Temos, então:



correlação linear positiva

correlação linear negativa

correlação não linear

não há correlação

11.2.4 Coeficiente de correlação linear

O instrumento empregado para a medida da correlação linear é o **coeficiente de correlação.** Esse coeficiente deve indicar o grau de intensidade da correlação entre duas variáveis e, ainda, o sentido dessa correlação (positivo ou negativo).

Faremos uso do **coeficiente de correlação de Pearson**, que é dado por:

$$r = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{\left[n \sum x_i^2 - (\sum x_i)^2\right]\left[n \sum y_i^2 - (\sum y_i)^2\right]}}$$

onde **n** é o número de observações.

Os valores limites de **r** são -1 e $+1$, isto é, o valor de **r** pertence ao intervalo $[-1, +1]$.

Assim:

a. se a correlação entre duas variáveis é perfeita e positiva, então $\mathbf{r = +1}$;

b. se a correlação é perfeita e negativa, então $\mathbf{r = -1}$;

c. se não há correlação entre as variáveis, então $\mathbf{r = 0}$.

Logicamente:

a. se $r = +1$, há uma correlação perfeita e positiva entre as variáveis;

b. se $r = -1$, há uma correlação perfeita e negativa entre as variáveis;

c. se $r = 0$, ou **não há correlação** entre as variáveis, ou a relação que porventura exista **não é linear**.

NOTAS:

- Para que uma relação possa ser descrita por meio do **coeficiente de correlação de Pearson** é imprescindível que ela se aproxime de uma função linear. Uma maneira prática de verificarmos a linearidade da relação é a inspeção do diagrama de dispersão: se a elipse apresenta saliências ou reentrâncias muito acentuadas, provavelmente trata-se de **correlação curvilínea**.
- Para podermos tirar algumas conclusões significativas sobre o comportamento simultâneo das variáveis analisadas, é necessário que:

  $$0{,}6 \leq |\ r\ | \leq 1.$$

  Se $0{,}3 \leq |\ r\ | < 0{,}6$, há uma correlação relativamente fraca entre as variáveis.

  Se $0 < |\ r\ | < 0{,}3$, a correlação é muito fraca e, praticamente, nada podemos concluir sobre a relação entre as variáveis em estudo.

Vamos, então, calcular o coeficiente de correlação relativo à Tabela 11.1. O modo mais prático para obtermos **r** é abrir, na tabela, colunas correspondentes aos valores de x_iy_i, x_i^2 e y_i^2. Assim:

n = 10

MATEMÁTICA (x_i)	ESTATÍSTICA (y_i)	x_iy_i	x_i^2	y_i^2
5	6	30	25	36
8	9	72	64	81
7	8	56	49	64
10	10	100	100	100
6	5	30	36	25
7	7	49	49	49
9	8	72	81	64
3	4	12	9	16
8	6	48	64	36
2	2	4	4	4
$\Sigma = 65$	$\Sigma = 65$	$\Sigma = 473$	$\Sigma = 481$	$\Sigma = 475$

TABELA 11.2

Logo:

$$r = \frac{10 \times 473 - 65 \times 65}{\sqrt{(10 \times 481 - 65^2)\ (10 \times 475 - 65^2)}} = \frac{4.730 - 4.225}{\sqrt{(4.810 - 4.225)\ (4.750 - 4.225)}}$$

$$= \frac{505}{\sqrt{585 \times 525}} = \frac{505}{554{,}18} = 0{,}911$$

Daí:

$$r = 0{,}91,$$

resultado que indica uma correlação linear positiva altamente significativa entre as duas variáveis.

Resolva

1. Complete o esquema de cálculo do coeficiente de correlação para os valores das variáveis x_i e y_i:

	4	6	8	10	12
	12	10	8	12	14

Temos:

	x_i	y_i	$x_i y_i$	x_i^2	y_i^2
n = 5	4	12	48	16	144
	....	....	....	....	....
	....	....	....	....	....
	....	....	....	....	....
	12	14	168	144	196
	$\Sigma = ...$	$\Sigma = ...$	$\Sigma = ...$	$\Sigma = ...$	$\Sigma = ...$

Logo:

$$r = \frac{.... \times .... - .... \times ....}{\sqrt{(.... \times .... - .... \times ....)(.... \times .... - .... \times ....)}} = \frac{.... - ....}{\sqrt{(.... - ....)(.... - ....)}} =,$$

$$= \frac{.....}{\sqrt{.... \times ....}} = \frac{.....}{\sqrt{....}} = \frac{.....}{.....} = ...,$$

donde r = 0,42.

A correlação linear entre as variáveis X e Y é positiva, porém fraca.

11.3 Regressão

11.3.1 Ajustamento da reta

Sempre que desejamos estudar determinada variável em função de outra[1], fazemos uma **análise de regressão.**

Podemos dizer que a **análise de regressão** tem por objetivo descrever, através de um modelo matemático, a relação entre duas variáveis, partindo de **n** observações das mesmas.

A variável sobre a qual desejamos fazer uma estimativa recebe o nome de **variável dependente** e a outra recebe o nome de **variável independente**.

Assim, supondo **X** a variável independente e **Y** a dependente, vamos procurar determinar o ajustamento de uma reta à relação entre essas variáveis, ou seja, vamos obter uma função definida por:

$$Y = aX + b,$$

onde **a** e **b** são os parâmetros.

Sejam duas variáveis **X** e **Y**, entre as quais exista uma correlação acentuada, embora não perfeita, como, por exemplo, as que formam a Tabela 11.2.

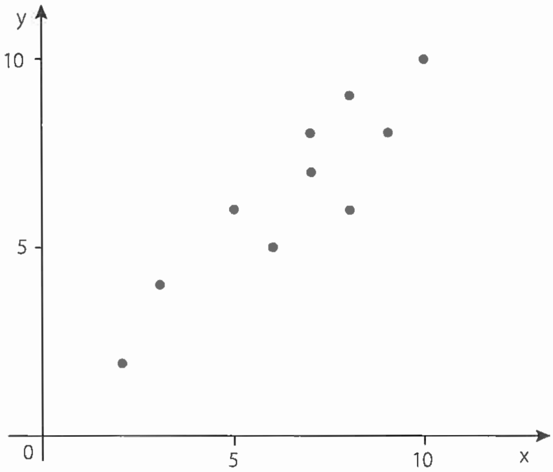
[1] Lembre-se de que estamos restritos à regressão linear simples.

Daí, temos:

	5	8	7	10	6	7	9	3	8	2
	6	9	8	10	5	7	8	4	6	2

TABELA 11.3

cujo diagrama de dispersão é dado por:



Podemos concluir, pela forma do diagrama, que se trata de uma correlação retilínea, de modo a permitir o ajustamento de uma reta, imagem da função definida por:

$$Y = aX + b$$

Vamos, então, calcular os valores dos parâmetros **a** e **b** com a ajuda das fórmulas:

$$a = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2}$$

e

$$b = \bar{y} - a\bar{x}$$

onde:

n é o número de observações;

$\bar{\mathbf{x}}$ é a média dos valores x_i $\left(\bar{x} = \frac{\sum x_i}{n}\right)$;

$\bar{\mathbf{y}}$ é a média dos valores y_i $\left(\bar{y} = \frac{\sum y_i}{n}\right)$.

NOTA:

- Como estamos fazendo uso de uma amostra para obtermos os valores dos parâmetros, o resultado, na realidade, é uma **estimativa** da verdadeira equação de regressão. Sendo assim, escrevemos:

$$\hat{Y} = aX + b,$$

onde $\hat{Y}$ é o **Y estimado.**

Formemos, então, a tabela de valores:

	x_i	y_i	$x_i y_i$	x_i^2
n = 10	5	6	30	25
	8	9	72	64
	7	8	56	49
	10	10	100	100
	6	5	30	36
	7	7	49	49
	9	8	72	81
	3	4	12	9
	8	6	48	64
	2	2	4	4
	$\Sigma = 65$	$\Sigma = 65$	$\Sigma = 473$	$\Sigma = 481$

TABELA 11.4

Temos, assim:

$$a = \frac{10 \times 473 - 65 \times 65}{10 \times 481 - (65)^2} = \frac{4.730 - 4.225}{4.810 - 4.225} = \frac{505}{585} = 0,8632$$

Como:

$$\bar{x} = \frac{65}{10} = 6,5 \text{ e } \bar{y} = \frac{65}{10} = 6,5,$$

vem:

$$b = 6,5 - 0,8632 \times 6,5 = 6,5 - 5,6108 = 0,8892,$$

donde:

$$a = 0,86 \text{ e } b = 0,89$$

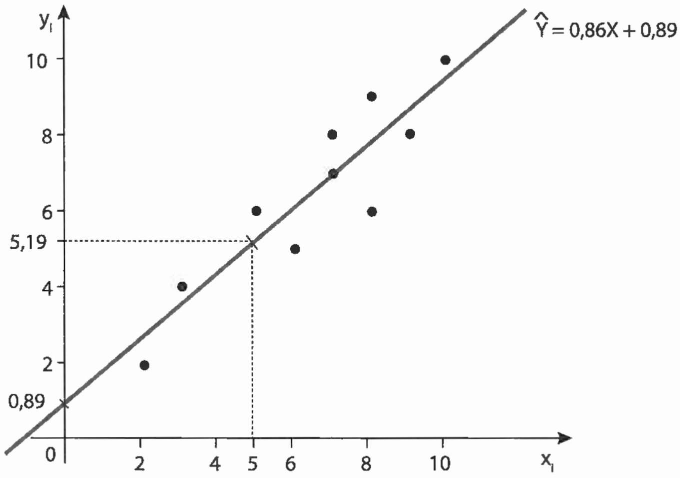
Logo:

$$\hat{Y} = 0,86X + 0,89$$

Para traçarmos a reta no gráfico, basta determinar dois de seus pontos:

$$X = 0 \Rightarrow \hat{Y} = 0,89$$
$$X = 5 \Rightarrow \hat{Y} = 0,86 \times 5 + 0,89 = 5,19$$

Assim, temos:



11.3.2 Interpolação e extrapolação

Voltando à Tabela 11.1, vemos que **4,0** não figura entre as notas de Matemática. Entretanto, podemos estimar a nota correspondente em Estatística fazendo **X = 4,0** na equação:

$$\hat{Y} = 0,86X + 0,89$$

Assim:

$$X = 4,0 \Rightarrow \hat{Y} = 0,86 \times 4,0 + 0,89 = 4,33$$

O mesmo acontece com a nota **1,0.** Repetindo o procedimento, temos:

$$X = 1,0 \Rightarrow \hat{Y} = 0,86 \times 1,0 + 0,89 = 1,75$$

Como $4 \in [2, 10]$, dizemos que foi feita uma **interpolação**; e como $1 \notin [2,10]$, dizemos que foi feita uma **extrapolação.**

NOTA:

- Uma norma fundamental no uso de equações de regressão é a de nunca extrapolar, exceto quando considerações teóricas ou experimentais demonstrem a possibilidade de extrapolação.

Resolva

1. Complete o esquema para o ajustamento de uma reta aos dados:

	2	4	6	8	10	12	14
	30	25	22	18	15	11	10

Temos:

x_i	y_i	x_iy_i	x_i^2
2	30	60	4
....	....	....	....
....	....	....	....
....	....	....	....
....	....	....	....
....	....	....	....
14	10	140	196
$\Sigma = ...$	$\Sigma = ...$	$\Sigma = ...$	$\Sigma = ...$

n = 7

Logo:

$$a = \frac{.... \times .... - .... \times ....}{.... \times .... - (....)^2} = \frac{.... - ....}{.... - ....} = \frac{....}{....} = ...$$

$$b = ... - (....) = ... + ... = ...$$

donde:

a = ... e b = ...,

isto é:

$$\hat{Y} = -1{,}7X + 32{,}3$$

Exercícios

1. Um grupo de pessoas fez uma avaliação do peso aparente de alguns objetos. Com o peso real e a média dos pesos aparentes, dados pelo grupo, obteve-se a tabela:

	18	30	42	62	73	97	120
	10	23	33	60	91	98	159

Calcule o índice de correlação.

2. Considere os resultados de dois testes, **X** e **Y**, obtidos por um grupo de alunos da escola **A**:

	11	14	19	19	22	28	30	31	34	37
	13	14	18	15	22	17	24	22	24	25

a. Verifique, pelo diagrama, se existe correlação retilínea.

b. Em caso afirmativo, calcule o coeficiente de correlação.

c. Escreva, em poucas linhas, as conclusões a que chegou sobre a relação entre essas variáveis.

3. A tabela abaixo apresenta a produção de uma indústria:

ANOS	1980	1981	1982	1983	1984	1985	1986	1987	1988
QUANTIDADES (t)	34	36	36	38	41	42	43	44	46

Calcule:

a. o coeficiente de correlação;

Sugestão: Para simplificar os cálculos, use para o tempo uma variável auxiliar, por exemplo: $x_i' = x_i - 1984$.

b. a reta ajustada;

c. a produção estimada para 1989.

NOTA:

- Lembre-se de que foi usada uma variável auxiliar.

4. A tabela abaixo apresenta valores que mostram como o comprimento de uma barra de aço varia conforme a temperatura:

TEMPERATURA (°C)	10	15	20	25	30
COMPRIMENTO (mm)	1.003	1.005	1.010	1.011	1.014

Determine:

a. o coeficiente de correlação;

b. a reta ajustada a essa correlação;

c. o valor estimado do comprimento da barra para a temperatura de 18°C;

d. o valor estimado do comprimento da barra para a temperatura de 35°C.

5. A variação do valor da UPC, relativamente a alguns meses de 1995, deu origem à tabela:

	mai.	jun.	jul.	ago.	set.	out.	nov.
	10,32	10,32	11,34	11,34	11,34	12,22	12,22

a. Calcule o grau de correlação.

b. Estabeleça a equação de regressão de **Y** sobre **X**.

c. Estime o valor da UPC para o mês de dezembro.

Sugestão: Substitua os meses, respectivamente, por 1, 2, ..., 7.

6. A partir da tabela:

	1	2	3	4	5	6
	70	50	40	30	20	10

a. calcule o coeficiente de correlação;

b. determine a reta ajustada;

c. estime o valor de **Y** para X = 0.

7. Certa empresa, estudando a variação da demanda de seu produto em relação à variação de preço de venda, obteve a tabela:

	38	42	50	56	59	63	70	80	95	110
	350	325	297	270	256	246	238	223	215	208

a. Determine o coeficiente de correlação.

b. Estabeleça a equação da reta ajustada.

c. Estime **Y** para X = 60 e X = 120.

8. Pretendendo-se estudar a relação entre as variáveis "consumo de energia elétrica" (x_i) e "volume de produção nas empresas industriais" (y_i), fez-se uma amostragem que inclui vinte empresas, computando-se os seguintes valores:

$$\sum x_i = 11{,}34,\ \sum y_i = 20{,}72,\ \sum x_i^2 = 12{,}16,\ \sum y_i^2 = 84{,}96 \text{ e } \sum x_i y_i = 22{,}13$$

Determine:

a. o cálculo do coeficiente de correlação;

b. a equação de regressão de **Y** para **X**;

c. a equação de regressão de **X** para **Y**.

12 NÚMEROS-ÍNDICES

12.1 Introdução

Um jornal, por ocasião de um pleito eleitoral, publicou uma tabela com os resultados da apuração na região:

CIDADES	CANDIDATO X	CANDIDATO Y	VOTOS BRANCOS	VOTOS NULOS	TOTAL
A	39.544	30.279	980	11.549	82.352
B	18.872	19.897	787	6.210	45.766
C	8.139	4.903	177	1.324	14.543
D	16.263	8.659	464	2.997	28.383
E	746	899	45	216	1.906
F	3.149	3.120	93	517	6.879

TABELA 12.1

Para um estudo comparativo das variações dos votos brancos, essa tabela, com números absolutos, em nada nos ajuda. Confeccionando, porém, uma nova tabela, com **números relativos**, obtemos:

CIDADES	VOTOS BRANCOS (%)
A	1,19
B	1,72
C	1,22
D	1,63
E	2,36
F	1,35

TABELA 12.2

o que nos leva a concluir, de imediato, que a cidade **E** foi a que apresentou maior índice de votos brancos.

Não são poucas as situações em que, para a descrição ou análise de um fenômeno quantitativo, o emprego dos **números relativos** revela-se mais pertinente do que o dos **números absolutos**. Isso acontece, naturalmente, quando pretendemos efetuar comparações dos valores tomados por uma mesma variável em épocas ou regiões diferentes.

Essas comparações representam o caso mais simples das medidas estatísticas, que denominamos números-índices, usados, principalmente, nos negócios e na economia.

12.2 Números-índices

Consideremos a tabela abaixo, relativa às matrículas efetivadas em certo estabelecimento de ensino durante o período de 1989 a 1994:

ANOS	1989	1990	1991	1992	1993	1994
MATRÍCULA	1.050	1.150	1.200	1.400	1.560	1.700
NÚMERO-ÍNDICE	100,0	109,5	114,3	133,3	148,6	161,9

TABELA 12.3

A vantagem dos números-índices é permitir uma rápida avaliação da variação relativa (percentual) sofrida pelo número de matrículas, e que se traduz, em relação a 1989, por um aumento de 9,5% em 1990, de 14,3% em 1991, de 33,3% em 1992, de 48,6% em 1993 e de 61,9% em 1994.

Assim, podemos dizer que:

Número-índice ou, simplesmente, **índice** é a relação entre dois estados de uma variável ou de um grupo de variáveis, suscetível de variar no tempo ou no espaço (ou de grupo de indivíduos para grupo de indivíduos).

O índice representa, portanto, o nível de um fenômeno em relação ao nível que ele tinha num dado período (ou numa dada região) tomado como **base**, e é geralmente expresso em percentagem.

Os índices mais utilizados relacionam, em geral, variações de preço, de quantidade ou de valor (preço × quantidade) ao longo do tempo.

NOTA:

- **Os índices não estão associados apenas aos negócios e à economia, mas são largamente utilizados em todos os ramos das ciências físicas, químicas, naturais e sociais. A Psicologia, por exemplo, emprega os índices para medir a inteligência (quociente de inteligência — QI).**

12.3 Relativos de preços

Quando queremos analisar a variação no preço (ou na quantidade ou no valor) de um só bem, basta expressar tal variação em termos percentuais, obtendo o que denominamos **relativo** de preço (de quantidade ou de valor).

Assim, representando por **o** a época-base ou base e por **t** a época atual, temos:

$\mathbf{p_o}$**:** preço na época-base;

$\mathbf{p_t}$**:** preço na época atual.

Atribuindo ao preço na época-base o valor 100, por meio de uma regra de três simples, calculamos o **relativo** correspondente ao preço atual:

$$\begin{array}{l} p_o \text{ — } 100 \\ p_t \text{ — } p_{o,t} \end{array} \Rightarrow \quad p_{o,t} = \frac{p_t}{p_o} \times 100 \qquad (p_{o,t} \text{ é o relativo de preço})$$

Do mesmo modo, obtemos:

$$q_{o,t} = \frac{q_t}{q_o} \times 100 \qquad \text{(relativo de quantidade)}$$

$$v_{o,t} = \frac{v_t}{v_o} \times 100 \qquad \text{(relativo de valor)}$$

Exercício resolvido

1. Sabendo que o preço de determinado produto era de R$ 50 em 1994 e de R$ 60 em 1995, determine o relativo de preço em 1995, tomando como base o ano de 1994. (É comum a notação 1994 = 100 para indicar que o ano de 1994 é tomado como base.)

Temos:

$$p_{94} = 50 \text{ e } p_{95} = 60$$

Logo:

$$p_{94,\,95} = \frac{p_{95}}{p_{94}} \times 100 = \frac{60}{50} \times 100 =$$

$$= 1{,}20 \times 100 = 120$$

Daí:

$$p_{94,\,95} = 120\%$$

Esse resultado nos permite afirmar que o preço do produto em 1995 corresponde a 120% de seu preço em 1994.

Concluímos, então, que o preço do produto entre 1994 e 1995 sofreu um aumento de:

$$120 - 100 = 20\%$$

12.4 Elos de relativos

Vários relativos formam **elos** quando cada um deles é calculado tomando como base o ano anterior; são os relativos de base móvel.

Assim, se um bem apresentou, no período de 1991 a 1994, respectivamente os preços de R$ 240, R$ 300, R$ 360 e R$ 540[1], os elos relativos são:

$$p_{91,\,92} = \frac{p_{92}}{p_{91}} \times 100 = \frac{300}{240} \times 100 = 1{,}25 \times 100 = 125$$

$$p_{92,\,93} = \frac{p_{93}}{p_{92}} \times 100 = \frac{360}{300} \times 100 = 1{,}2 \times 100 = 120$$

$$p_{93,\,94} = \frac{p_{94}}{p_{93}} \times 100 = \frac{540}{360} \times 100 = 1{,}5 \times 100 = 150$$

Com esses resultados, podemos formar a tabela de elos:

[1] No período de 1991 a 1994, a moeda circulante no Brasil não era o real. Por questões didáticas, estamos deixando de considerar esse detalhe.

	1991	1992	1993	1994
	—	125	120	150

TABELA 12.4

Fazemos uso dos elos de relativos quando queremos acompanhar os crescimentos (positivos ou negativos) anuais (ou mensais, ou diários).

12.5 Relativos em cadeia

O **relativo em cadeia** é o índice de base fixa: todos os relativos são calculados tomando-se uma determinada época como base.

Utilizando como exemplo os dados do item anterior e considerando 1991 como ano-base, obtemos:

$$p_{91,\,92} = \frac{p_{92}}{p_{91}} \times 100 = \frac{300}{240} \times 100 = 1{,}25 \times 100 = 125$$

$$p_{91,\,93} = \frac{p_{93}}{p_{91}} \times 100 = \frac{360}{240} \times 100 = 1{,}5 \times 100 = 150$$

$$p_{91,\,94} = \frac{p_{94}}{p_{91}} \times 100 = \frac{540}{240} \times 100 = 2{,}25 \times 100 = 225$$

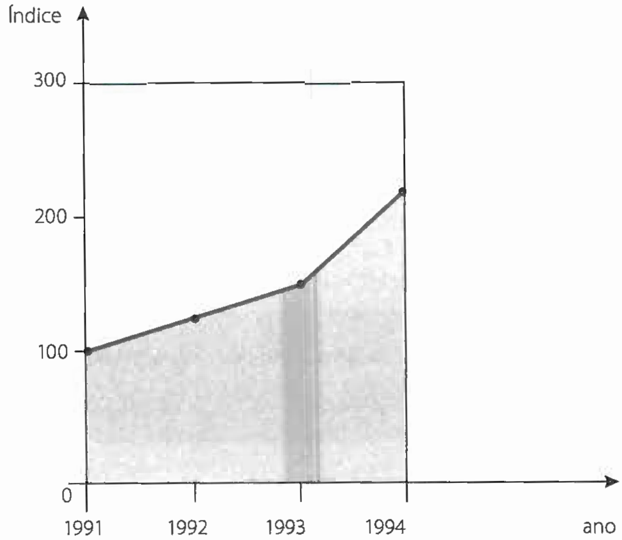
Esses resultados dão origem à tabela de relativos em cadeia:

	1991	1992	1993	1994
	100	125	150	225

TABELA 12.5

Fazemos uso dos relativos em cadeia quando desejamos comparar um determinado ano, considerado significativo, com os anos anteriores e os consecutivos.

O gráfico abaixo mostra a evolução do preço do bem em questão:



Exercício

1. Dada a tabela:

QUANTIDADE DE BENS (1991-94)

BENS	ANOS			
	1991	1992	1993	1994
Autoveículos (mil unid.)	1.128,0	1.165,2	780,9	859,3
Cimento (milhões de t)	24,9	27,2	26,1	25,4
Aço (milhões de t)	13,9	15,2	13,1	12,9
Petróleo bruto (milhões de m³)	9,6	10,6	12,4	15,1

Dados fictícios.

a. calcule os relativos para o bem autoveículos, tomando 1991 = 100;

b. forme a tabela dos elos de relativos para o cimento;

c. forme a tabela dos relativos em cadeia para o aço, tomando 1992 = 100;

d. verifique a igualdade: $q_{91,92} \times q_{92,93} \times q_{93,94} = q_{91,94}$ para o petróleo bruto;

e. represente a evolução dos índices das questões **a** e **c**, usando o gráfico em linha.

12.6 Índices agregativos

Os índices que estudamos até agora servem apenas para caracterizar a marcha do **preço de um só bem**. No entanto, a variação de preços exige um índice que sintetize a variação dos preços de um **conjunto de bens (agregado)**. Para atingir esse objetivo, lançamos mão de um novo tipo de índice: o **índice agregativo**.

Existem inúmeras maneiras de calcularmos os índices agregativos, embora os fundamentos básicos sejam constantes, variando apenas aspectos relacionados com o campo específico de aplicação do índice.

12.6.1 Índice agregativo simples

Um modo de determinar o índice agregativo simples é calcular a média aritmética dos relativos, obtendo o **índice médio de relativos**.

Assim, dada a tabela abaixo:

BENS	RELATIVOS DE PREÇOS	
	1994	1995
A (m)	100	150
B (kg)	100	125
C (ℓ)	100	160
	$\Sigma = 300$	$\Sigma = 435$

TABELA 12.6

temos, lembrando que n = 3:

$$I_{\bar{p}} = \frac{435}{3} \Rightarrow I_{\bar{p}} = 145\%.$$

12.6.2 Índice agregativo ponderado

No cálculo do índice simples, todos os itens do agregado são colocados em um mesmo nível. Sabemos, porém, que na prática isso não acontece; há bens de importância muito maior que outros, razão pela qual devemos considerar **coeficientes de ponderação**, atribuindo, a cada item, a importância que lhe cabe.

Para o cálculo do índice agregativo ponderado, há várias fórmulas: de Laspeyres, de Paasche, de Fisher etc.

Tomando como referência os relativos de preço, aplicaremos um dos métodos de ponderação para obtermos os índices mais usuais na investigação econômica.

Fórmula de Laspeyres ou método da época-base

Ponderando os relativos de preço $\frac{p_t}{p_o}$, onde p_t é o preço na época atual e p_o é o preço na época-base, pelos valores (preços × quantidades) do ano-base $p_o q_o$, obtemos a fórmula de Laspeyres:

$$Lp_{o,\,t} = \frac{\sum \frac{p_t}{p_o} \times p_o q_o}{\sum p_o q_o}$$

que, simplificada, nos dá:

$$Lp_{o,\,t} = \frac{\sum p_t q_o}{\sum p_o q_o}$$

Exercício resolvido

1. Considerando a tabela:

BENS	1993		1994	
	p	q	p	q
A	20	4	28	3
B	40	3	56	3
C	15	8	30	12

calcule o índice ponderado para preços, empregando a fórmula de Laspeyres e tomando 1993 = 100.

Lembrando que:

$$Lp_{93,\,94} = \frac{\sum p_{94}\, q_{93}}{\sum p_{93}\, q_{93}},$$

temos:

$$Lp_{93,\,94} = \frac{(28 \times 4) + (56 \times 3) + (30 \times 8)}{(20 \times 4) + (40 \times 3) + (15 \times 8)} =$$

$$= \frac{112 + 168 + 240}{80 + 120 + 120} = \frac{520}{320} =$$

$$= 1{,}625 \text{ ou } 162{,}5\%$$

12.6.3 Índices de preços

Para construir um índice de preços, qualquer que seja a sua finalidade, devemos inicialmente considerar os seguintes pontos:

a. Qual o objetivo do índice?

b. Que produtos devem ser incluídos no seu cálculo?

c. Quais os preços a serem incluídos no seu cálculo?

d. Qual o peso a ser atribuído a cada bem em particular?

e. Qual a fórmula adequada?

Embora não tendo uma resposta imediata para as questões acima, alguns pontos básicos devem ser observados sempre que pretendemos construir qualquer índice.

a. **Objetivo do índice**

   É fundamental qualificar, com toda a precisão, o objetivo do índice; determinar o que ele está medindo e a quem se refere. Dessa determinação dependerá a seleção dos produtos que comporão o índice.

b. **Produtos a serem incluídos**

   Devem ser incluídos os produtos julgados mais importantes e que sejam representativos do conjunto de bens que integram o setor para o qual se vai calcular o índice.

c. **Preços a serem incluídos**

   Deve-se identificar o setor para o qual vão ser determinados os preços (varejo, atacado etc.). Também é necessário decidir a forma de cotação e como deverão ser coletados os preços.

d. **Pesos a serem atribuídos**

   O sistema de pesos a ser atribuído deve depender essencialmente da finalidade ou da utilidade do índice. Os pesos, por isso mesmo, devem refletir a importância relativa de cada bem no conjunto tomado para a determinação do índice.

e. **Fórmula**

   Em geral, quando se trata de índices de preços, é usada a fórmula de Laspeyres, que emprega **pesos fixos**, permitindo a revisão periódica de seus valores. Resulta daí a possibilidade de termos sempre as mesmas comparações, feitas diretamente ou através de elos de relativos.

Índice de custo de vida

O **índice de custo de vida** ou **índice de preços ao consumidor** é um número-índice que procura medir a variação de preços de um conjunto de bens e serviços necessários à vida do consumidor final padrão (**família padrão**).

É evidente que devem ser considerados os preços dos bens consumidos em alimentação, vestuário, mobiliário, habitação, lazer, saúde, higiene etc., além, é claro, dos gastos com água, luz, transporte, educação e outros.

Sua metodologia está firmada em pesquisas, junto às famílias, que determinam a lista dos bens e serviços consumidos por elas e a percentagem dos gastos com cada grupo de bens e serviços.

A partir desses dados, fixamos um índice de preços (Laspeyres) para cada grupo.

Finalmente, calculamos a média aritmética ponderada dos índices de preços dos grupos, tomando para pesos os valores percentuais dos gastos com cada grupo na despesa total da família padrão.

IPC — Índice de Preços ao Consumidor

Esse índice reflete os gastos de famílias com renda entre um e oito salários mínimos, sendo o chefe da família assalariado em sua ocupação principal.

A coleta de preços é feita pelo IBGE, em dez regiões metropolitanas. O período pesquisado é do dia 16 de um mês ao dia 15 do mês seguinte.

ICB — Índice da Cesta Básica

Esse índice é empregado para corrigir o salário mínimo a cada bimestre.

Sua metodologia é semelhante à do IPC, porém representa os gastos de famílias com renda de até dois salários mínimos.

IGP — Índice Geral de Preços

O IGP, calculado pela Fundação Getulio Vargas, é a média ponderada dos seguintes índices: Índice de Preços por Atacado (60%), Índice de Custo de Vida (30%) e Índice de Custo da Construção Civil na Cidade do Rio de Janeiro (10%). O período de coleta de preços é de 1º a 30 do mês de referência.

IPC da FIPE

FIPE é a Fundação Instituto de Pesquisas Econômicas da USP, que pesquisa o custo de vida em São Paulo para famílias que ganham de dois a seis salários mínimos. A FIPE compara os preços médios de quatro semanas com os das quatro semanas imediatamente anteriores.

12.7 Deflacionamento de dados

Sabemos que os aumentos de preços implicam baixas no poder de compra ou no valor da moeda. Por isso mesmo, a manutenção do poder de compra dos salários é um problema que muito preocupa os assalariados de países onde o valor da moeda está continuamente se deteriorando.

Assim, embora os **salários nominais** estejam frequentemente **aumentando**, os **salários reais** podem estar diminuindo, devido ao aumento do custo de vida (inflação), e, consequentemente, tendo o seu **poder aquisitivo** reduzido.

Daí a importância dos índices de preços, pois a eles recorremos para responder a questões como esta:

Sabendo-se que um assalariado, em 1º de maio de 1993, ganhava **x** cruzeiros por mês, qual deveria ser seu salário mensal, em 1º de janeiro de 1994, para que ele se encontrasse em situação equivalente à anterior?

Esse problema trata da conversão de **salários nominais** em **salários reais**, de importância fundamental na época das negociações salariais, principalmente quando há inflação.

Para determinarmos os salários reais (SR), também denominados **salários deflacionados**, dividimos os salários nominais das várias épocas (S_t) pelo índice de preços das épocas correspondentes (IP_t) e multiplicamos o resultado por 100:

$$SR = \frac{S_t}{IP_t} \times 100 \qquad ①$$

Assim, se o salário de um professor, em dezembro de 1995, era de R$ 1.071 e o IP de dezembro de 1995, com base em novembro, era de 101,24%, o valor aquisitivo desse professor é dado por:

$$SR = \frac{1.071}{101,24} \times 100 = 1.057,88,$$

isto é:

R$ 1.058.

Esse procedimento é denominado **deflacionamento** de salários e o índice de preços usado na determinação do salário real é chamado **deflator**.

Processo semelhante pode ser empregado para **deflacionar** outras séries temporais.

Assim, substituindo em ① "salário" por "valor", obtemos:

$$VR = \frac{V_t}{IP_t} \times 100$$

Tomando como exemplo o faturamento de uma empresa no período de 1991 a 1994, dado pela Tabela 12.7, vamos determinar o seu faturamento real, relativamente:

a. ao período de 1990;

b. ao período de 1991.

ANOS	FATURAMENTO (R$)	IP 1990 = 100
1991	180.000	140,8
1992	220.000	291,1
1993	430.000	362,5
1994	480.000	410,3

TABELA 12.7

a. Para obtermos o faturamento da empresa relativamente ao ano de 1990, basta dividir cada valor constante da coluna referente ao faturamento pelo índice geral de preços do respectivo ano. Com isso, estamos **deflacionando** a série. Assim:

$$\frac{180.000}{140,8} \times 100 = 127.841 \qquad \frac{430.000}{362,5} \times 100 = 118.620$$

$$\frac{220.000}{291,1} \times 100 = 75.575 \qquad \frac{480.000}{410,3} \times 100 = 116.988$$

Logo:

ANOS	FATURAMENTO A PREÇOS DE 1990 (R$)
1991	127.841
1992	75.575
1993	118.620
1994	116.988

TABELA 12.8

b. A fim de obtermos o faturamento da empresa, em termos de preços de 1991, devemos, inicialmente, mudar a base do ano 1990 para o ano 1991 e, em seguida, operarmos como em **a.** Assim:

$$IP_{91,\ 92} = \frac{291,1}{140,8} \times 100 = 206,7$$

$$IP_{91,\ 93} = \frac{362,5}{140,8} \times 100 = 257,5$$

$$IP_{91,\ 94} = \frac{410,3}{140,8} \times 100 = 291,4$$

donde:

$$\frac{220.000}{206,7} \times 100 = 106.434$$

$$\frac{430.000}{257,5} \times 100 = 166.990$$

$$\frac{480.000}{291,4} \times 100 = 164.722$$

Logo:

ANOS	FATURAMENTO (R$)	IP 1991 = 100	FATURAMENTO A PREÇOS DE 1991 (R$)
1991	180.000	100,0	180.000
1992	220.000	206,7	106.434
1993	430.000	257,5	166.990
1994	480.000	291,4	164.722

TABELA 12.9

Pelo exame da tabela, vemos que o faturamento no ano de 1994 foi, em termos reais, inferior ao de 1991, embora, em termos nominais, tenha aumentado.

Exercícios

1. Dada a tabela abaixo:

ANOS	1989	1990	1991	1992	1993	1994
ÍNDICES 1989 = 100	100	152	203	321	415	580

calcule os índices, tomando 1991 como ano-base.

2. O salário médio horário de determinada classe operária, em 1994, foi de R$ 2.560. O IP, nesse mesmo ano, era igual a 1.575,7 e o de 1991 era igual a 387,2, referidos ao período-base de 1982. Tomando o ano de 1991 como base, determine o salário real dessa classe operária em 1994.

3. Dada a tabela:

ANOS	FATURAMENTO (R$)	IP (1986 = 100)
1989	385.000	234
1990	474.200	280
1991	612.500	329
1992	983.200	380
1993	1.230.000	490
1994	1.984.000	625
1995	3.038.000	894

determine o valor do faturamento relativamente ao período de 1991.

4. Se os preços dos cigarros aumentam 70% e, como consequência, o ICV sobe 1,8%, que ponderação tem esse bem econômico dentro do custo de vida?

5. O IP, em dado período, aumenta de 15%. Qual deve ser o aumento dos salários dos empregados de uma empresa para que tenham um aumento real de 5%?

APÊNDICE: INSTRUMENTAL MATEMÁTICO

1. Números aproximados e arredondamento de dados

1.1 Números aproximados

Como sabemos, os números resultam de uma mensuração (no seu sentido mais amplo), a qual só pode ser **exata** quando assume a forma de contagem ou enumeração, em números naturais, de coisas ou unidades mínimas indivisíveis. Em tais casos, a variável pode assumir somente **valores discretos** ou **descontínuos**.

Outras mensurações se dão numa **escala contínua**, que pode, teoricamente, ser indefinidamente subdividida. Na prática, porém, há sempre um limite para a precisão com a qual a mensuração pode ser feita, o que nos leva a concluir que o valor verdadeiro nunca é conhecido. Na verdade, os valores observados são **discretos** e **aproximados**.

Assim é que, se o comprimento de um parafuso, medido em centímetros, foi dado por 4,6 cm, devemos considerar que o valor exato desse comprimento será algum valor entre 4,55 cm e 4,65 cm, que foi aproximado para 4,6 cm devido ao fato de a precisão adotada na medida ser apenas de décimos de centímetro.

Em nossos estudos, faremos uso da seguinte convenção: **a precisão da medida será automaticamente indicada pelo número de decimais com que se escrevem os valores da variável**.

Assim, um valor 4,60 indica que a variável em questão foi medida com a precisão de centésimos, não sendo exatamente o mesmo que 4,6, valor correspondente a uma precisão de décimos.

1.2 Arredondamento de dados

Muitas vezes, é necessário ou conveniente suprimir unidades inferiores às de determinada ordem. Esta técnica é denominada **arredondamento de dados.**

De acordo com as normas da Fundação IBGE, o arredondamento é feito da seguinte maneira:

- Quando o primeiro algarismo a ser abandonado é **0**, **1**, **2**, **3**, ou **4**, fica inalterado o último algarismo a permanecer.

  Exemplo: 53,24 passa a 53,2.

- Quando o primeiro algarismo a ser abandonado é **5**, **6**, **7**, **8** ou **9**, aumenta-se de uma unidade o algarismo a permanecer.

  Exemplos: 42,87 passa a 42,9
  25,08 passa a 25,1
  53,99 passa a 54,0

NOTA:

- Não devemos nunca fazer arredondamentos sucessivos.

  Exemplo: 17,3452 passa a 17,3 e não a 17,35, a 17,4.

  Se tivermos necessidade de um novo arredondamento, fica recomendada a volta aos dados originais.

Resolva

1. Arredonde cada um dos dados abaixo, deixando-os com apenas uma casa decimal:
   a. 2,38 = 2,4
   b. 24,65 = 24,7
   c. 0,351 = ...
   d. 4,24 = ...
   e. 328,35 = ...
   f. 2,97 = ...
   g. 6,829 = ...
   h. 5,550 = ...
   i. 89,99 = ...

1.3 Compensação

Suponhamos os dados abaixo, aos quais aplicamos as regras do arredondamento:

25,32		25,3
17,84		17,8
10,44		10,4
+ 31,15		+ 31,2
84,75		(84,7)
84,8	(arredondando o resultado)	

Verificamos que houve uma pequena discordância: a soma é exatamente 84,7 quando, pelo arredondamento, deveria ser 84,8. Entretanto, para a apresentação dos resultados, é necessário que desapareça tal diferença, o que é possível pela prática do que denominamos **compensação**, conservando o mesmo número de casas decimais.

Praticamente, usamos "descarregar" a diferença na(s) maior(es) parcela(s). Assim, passaríamos a ter:

$$\begin{array}{r} 25{,}3 \\ 17{,}8 \\ 10{,}4 \\ +\ 31{,}3 \\ \hline 84{,}8 \end{array}$$

NOTA:

- Convém, ainda, observar que, se a maior parcela é igual ao dobro de qualquer outra parcela (ou maior que esse dobro), "descarregamos" a diferença (maior que uma unidade) apenas na maior parcela.

Exercícios

1. Arredonde cada um dos numerais abaixo, conforme a precisão pedida:

   a. Para o décimo mais próximo:

23,40	48,85002	120,4500
234,7832	78,85	129,98
45,09	12,35	199,97

   b. Para o centésimo mais próximo:

46,727	253,65	28,255
123,842	299,951	37,485

   c. Para a unidade mais próxima:

26,6	67,5	128,5
49,98	68,2	39,49

   d. Para a dezena mais próxima:

42,3	265,31	295
59	265,0	302,7
446,4	265	2.995,000

2. Arredonde para o centésmo mais próximo e compense, se necessário:
   $0{,}060 + 0{,}119 + 0{,}223 + 0{,}313 + 0{,}164 + 0{,}091 + 0{,}030 = 1{,}000$

3. Arredonde para a unidade mais próxima e compense, se necessário:
   $4{,}0 + 7{,}6 + 12{,}4 + 27{,}4 + 11{,}4 + 8{,}0 = 70{,}8$

2. Frações

2.1 Conceito

Fração é um par ordenado de números naturais, com o segundo elemento diferente de zero.

$$\frac{a}{b}, \text{ com } a \in \mathbb{N} \text{ e } b \in \mathbb{N}^*$$

NOTA:
- $\mathbb{N}$ é o conjunto dos números naturais e $\mathbb{N}^*$ é o conjunto dos números naturais com exclusão do zero.

2.2 Frações própria, imprópria e aparente

Fração própria é aquela cujo numerador (diferente de zero) é menor que o denominador.

Exemplos: $\frac{2}{3}, \frac{4}{5}, \frac{12}{17}$ etc.

Fração imprópria é aquela cujo numerador é igual ao denominador ou maior que ele.

Exemplos: $\frac{5}{2}, \frac{3}{3}, \frac{8}{4}$ etc.

Fração aparente é a fração imprópria cujo numerador é múltiplo do denominador.

Exemplos: $\frac{3}{3}, \frac{8}{4}$ etc.

NOTAS:
- A fração aparente representa o **número natural**, que é o **quociente** do numerador pelo denominador.
  Assim, $\frac{8}{4}$ representa o número natural 2, pois $8 : 4 = 2$.
- Se o numerador é **zero**, a fração representa o número zero.
  Assim, $\frac{0}{3} = 0$.
- Todo número natural pode ser representado por uma fração com denominador **1** e numerador igual ao número considerado.
  Assim, 5 pode ser representado por $\frac{5}{1}$.

2.3 Frações equivalentes

Duas frações são **equivalentes** quando os produtos do numerador de uma pelo denominador da outra são iguais.

Exemplo:

Para $\frac{2}{3}$ e $\frac{4}{6}$, temos: $2 \times 6 = 3 \times 4$. Logo: $\frac{2}{3} = \frac{4}{6}$

2.4 Simplificação de frações

Simplificar uma fração é obter uma fração equivalente à primeira com termos menores.

Para obter uma fração simplificada, basta dividir ambos os termos por um divisor comum.

Exemplo:

$$\frac{18}{30} = \frac{18 : 6}{30 : 6} = \frac{3}{5}$$

2.5 Fração irredutível

Fração irredutível é aquela cujos termos são números primos entre si (isto é, não possuem outro divisor comum a não ser o número 1).

Exemplo:

$\frac{7}{12}$ é uma fração irredutível, pois 7 e 12 são números primos entre si.

2.6 Redução de frações ao mesmo denominador

a. Calcula-se o menor múltiplo comum (**m.m.c.**) dos denominadores.

b. Escreve-se como denominador comum das frações o m.m.c. calculado; em seguida, divide-se o m.m.c. por cada um dos denominadores das frações dadas e multiplica-se o resultado pelo respectivo numerador.

Exemplo:

Reduzir ao mesmo denominador as frações:

$$\frac{7}{8}, \frac{3}{4}, \frac{1}{6}$$

a. cálculo do m.m.c.:

$$\begin{array}{ccc|c} 8, & 4, & 6 & 2 \\ 4, & 2, & 3 & 2 \\ 2, & 1, & 3 & 2 \\ 1, & 1, & 3 & 3 \\ 1, & 1, & 1 & \end{array}$$

$$\text{m.m.c.} = 2^3 \times 3 = 8 \times 3 = 24$$

b. $\frac{7\times3}{24}, \frac{3\times6}{24}, \frac{1\times4}{24}$

Logo:

$$\frac{21}{24}, \frac{18}{24}, \frac{4}{24}$$

NOTA:

- As frações que têm denominadores iguais são chamadas **frações homogêneas** e as que têm denominadores diferentes, **frações heterogêneas.**

2.7 Comparação de frações

Se queremos comparar duas ou mais frações, devemos reduzi-las ao mesmo denominador e lembrar que, **de duas frações com o mesmo denominador, é maior a que tem maior numerador**.

2.8 Operações com frações

Adição e subtração

a. **Frações homogêneas:** conserva-se o denominador e adicionam-se (ou subtraem-se) os numeradores.

Exemplos:

$$\frac{4}{5}+\frac{2}{5}=\frac{4+2}{5}=\frac{6}{5}$$

$$\frac{4}{7}-\frac{3}{7}=\frac{4-3}{7}=\frac{1}{7}$$

b. **Frações heterogêneas:** reduzem-se as frações ao mesmo denominador, obtendo-se, assim, frações homogêneas.

Exemplos:

$$\frac{4}{5}+\frac{2}{3}=\frac{12+10}{15}=\frac{22}{15}$$

$$\frac{6}{7}-\frac{1}{2}=\frac{12-7}{14}=\frac{5}{14}$$

NOTA:

- Sempre que possível, o resultado deve ser simplificado.

Exemplo:

$$\frac{5}{6}+\frac{2}{3}=\frac{5+4}{6}=\frac{\overset{3}{\cancel{9}}}{\underset{2}{\cancel{6}}}=\frac{3}{2}$$

Multiplicação

O **produto** de duas frações é uma fração cujo numerador é o produto dos numeradores e cujo denominador é o produto dos denominadores.

Exemplo:

$$\frac{2}{3}\times\frac{3}{5}=\frac{2\times3}{3\times5}=\frac{\overset{2}{\cancel{6}}}{\underset{5}{\cancel{15}}}=\frac{2}{5}$$

NOTAS:

- A operação multiplicação pode ser facilitada, realizando-se a simplificação pelo cancelamento dos fatores comuns dos numeradores e dos denominadores.

  Exemplos:

$$\frac{2}{\cancel{3}_{1}} \times \frac{\cancel{3}^{1}}{5} = \frac{2 \times 1}{1 \times 5} = \frac{2}{5}$$

$$\frac{\cancel{2}^{1}}{\cancel{3}_{1}} \times \frac{\cancel{3}^{1}}{\cancel{4}_{1}} \times \frac{\cancel{2}^{1}}{7} = \frac{1 \times 1 \times 1}{1 \times 1 \times 7} = \frac{1}{7}$$

- O dobro de 4 é $2 \times 4 = 8$; o triplo de $\frac{4}{7}$ é $3 \times \frac{4}{7} = \frac{12}{7}$. Por analogia:

$$\frac{2}{3} \text{ de } 4 \text{ é } \frac{2}{3} \times 4 = \frac{8}{3} \text{ e } \frac{3}{5} \text{ de } \frac{1}{4} \text{ é } \frac{3}{5} \times \frac{1}{4} = \frac{3}{20}$$

Divisão

O **quociente** de duas frações é o produto da primeira pelo inverso da segunda.

Exemplo:

$$\frac{4}{5} : \frac{5}{6} = \frac{4}{5} \times \frac{6}{5} = \frac{24}{25}$$

Potenciação

Para **elevar** uma fração a um expoente dado, devemos elevar o numerador e o denominador a esse expoente.

Exemplo:

$$\left(\frac{3}{2}\right)^2 = \frac{3^2}{5^2} = \frac{9}{25}$$

Exercícios

1. Que fração da semana corresponde a um dia?

2. Que fração do ano corresponde a dois meses?

3. Que fração do mês de fevereiro de um ano não bissexto corresponde a uma semana?

4. Três inteiros quantos quintos são?

5. Reduza 8 a sétimos, 12 a décimos e 7 a treze avos.

6. Simplifique as seguintes frações, de modo a torná-las irredutíveis:

a. $\frac{8}{12}$ d. $\frac{34}{6}$

b. $\frac{15}{30}$ e. $\frac{96}{144}$

c. $\frac{121}{11}$ f. $\frac{360}{600}$

7. Reduza ao mesmo denominador cada grupo de frações:

a. $\frac{2}{5}, \frac{5}{8}, \frac{3}{2}$ b. $\frac{4}{9}, \frac{5}{18}, \frac{7}{36}$ c. $\frac{8}{9}, \frac{5}{3}, \frac{7}{6}$

8. Escreva em ordem crescente de seus valores cada um dos seguintes grupos de frações:

a. $\frac{3}{8}, \frac{1}{8}, \frac{5}{8}$ b. $\frac{4}{6}, \frac{3}{5}, \frac{1}{3}$

9. Efetue as operações, simplificando os resultados quando possível:

a. $\frac{3}{4}+\frac{4}{5}$

b. $\frac{7}{12}+\frac{5}{8}$

c. $\frac{16}{14}-\frac{5}{7}$

d. $\frac{7}{4}-\frac{3}{2}+\frac{1}{2}$

e. $\frac{4}{10}+2\frac{3}{5}-1$

f. $\frac{6}{12}\times\frac{2}{9}$

g. $\frac{9}{15}\times 7$

h. $2\frac{1}{2}\times\frac{3}{5}$

i. $\frac{3}{5}\times\frac{1}{4}\times\frac{5}{6}$

j. $8\times\frac{10}{3}\times\frac{3}{8}\times\frac{1}{6}$

l. $\frac{9}{5} : \frac{1}{5}$

m. $\frac{2}{5} : 5$

n. $3\frac{1}{2} : \frac{7}{4}$

o. $\left(\frac{3}{5}\right)^2$

p. $\left(\frac{4}{5}\right)^3$

q. $\left(1\frac{2}{3}\right)^3$

r. $\left(\frac{8}{15}\right)^0$

s. $\left(\frac{1}{2}+\frac{2}{3}\right)\frac{2}{3}$

2.9 Frações decimais

Frações decimais são as frações cujos denominadores são potências de **10**.

Exemplos:

$$\frac{1}{10}, \frac{1}{100}, \frac{42}{1.000} \text{ etc.}$$

As frações decimais podem ser representadas por outro numeral, denominado **número decimal**, o qual é obtido pela seguinte **convenção: são dadas ao numerador tantas ordens decimais (casas) quantos são os zeros do denominador**.

Exemplos:

$\frac{1}{10} = 0,1$ (um décimo)

$\frac{1}{100} = 0,01$ (um centésimo)

$\frac{1}{1.000} = 0,001$ (um milésimo)

$\frac{1}{10.000} = 0,0001$ (um décimo milésimo)

$\frac{452}{100} = 4,52$ (quatro inteiros e cinquenta e dois centésimos)

NOTA:

- Deixamos de apresentar as técnicas de operações com números decimais, na suposição de que os alunos farão uso de calculadoras.

Exercícios

1. Represente, na forma decimal, os números:

   a. $\frac{3}{10}$

   b. $\frac{128}{100}$

   c. $\frac{50}{1.000}$

   d. $\frac{45}{10}$

2. Represente na forma de fração:

   a. 0,7

   b. 0,12

   c. 12,75

   d. 0,018

3. Calcule:

a. $0{,}532 + 1{,}2403 + 62{,}7 + 0{,}007$
b. $15{,}208 + 7{,}06 + 100{,}4 + 2$
c. $12{,}703 - 3{,}8$
d. $3 - 0{,}04$
e. $0{,}05 - 0{,}005$
f. $5{,}13 \times 0{,}3$
g. $27{,}5 \times 3$
h. $0{,}62 \times 10$
i. $3{,}8 \times 100$
j. $0{,}002 \times 6$
l. $6{,}36 \times 0{,}53$
m. $0{,}1575 \times 0{,}63$
n. $14{,}18 : 0{,}2$
o. $50 : 0{,}05$
p. $0{,}072 : 8$
q. $15 : 0{,}003$
r. $\sqrt{10{,}24}$
s. $\sqrt{127{,}69}$
t. $\sqrt{0{,}36}$
u. $\sqrt{0{,}0081}$

3. Razões

3.1 Razão de dois números

Razão do número **a** para o número **b** (diferente de zero) é o quociente exato de **a** por **b**.

Indicamos:

$$\frac{a}{b} \text{ (e lemos: a para b)}$$

Os números **a** e **b** são os **termos** da razão; **a** é chamado **antecedente** e **b**, **consequente** da razão.

Exemplos:

A razão de 3 para 12 é $\frac{3}{12} = \frac{1}{4}$.

A razão de 20 para 5 é $\frac{20}{5} = 4$.

3.2 Razão de duas grandezas

Razão de duas grandezas é o quociente dos números que expressam essas grandezas.

Exemplo:

Um automóvel percorre 36 km com 4 l de álcool. A razão entre distância percorrida e álcool gasto é:

$$\frac{36 \text{ km}}{4\ell} = 9 \text{ km}/\ell$$

Podemos dizer, então, que esse automóvel faz 9 km por litro de álcool ou 9 km/ℓ.

4. Percentagem

4.1 Conceito

Para evidenciar a participação de uma parte no todo e para facilitar comparações, costumamos usar razões com consequentes iguais a 100.

Denominamos **razões percentuais** as razões cujos consequentes sejam iguais a 100.

Exemplos:

$$\frac{25}{100}, \frac{4}{100}, \frac{212}{100}$$

A razão percentual $\frac{20}{100}$ pode também ser indicada pelo símbolo 20% (lemos: vinte por cento).

Assim, quando dizemos que 90% dos alunos de uma classe foram aprovados, isto significa que, se a classe tivesse 100 alunos, 90 desses alunos teriam sido aprovados.

Temos, então:

$$90\% = \frac{90}{100}$$

90 é a **percentagem** e 90% é a **taxa percentual**.

Os problemas de percentagem podem ser resolvidos com o emprego da regra de três simples.

Exercícios resolvidos

1. Em uma classe de 40 alunos, 32 foram aprovados. Qual a taxa percentual de aprovação?

Temos:

$$\frac{32}{40} = \frac{x}{100}$$

Logo:

$$40x = 32 \times 100 \Rightarrow x = \frac{32 \times 100}{40} = 80$$

80% é a resposta.

2. Ao comprar um livro, obtive um desconto de R$ 3. Qual o preço do livro, sabendo que a taxa de desconto foi de 5%?

Temos:

$$\frac{3}{x} = \frac{5}{100}$$

Logo:

$$5x = 3 \times 100 \Rightarrow x = \frac{3 \times 100}{5} = 60$$

R$ 60 é a resposta.

3. Uma pessoa comprou uma calça por R$ 20. Obteve um desconto de 15%. De quanto foi o desconto?

Temos:

$$\frac{x}{20} = \frac{15}{100}$$

Logo:

$$100x = 15 \times 20 \Rightarrow x = \frac{15 \times 20}{100} = 3$$

R$ 3 é a resposta.

NOTA:

- Neste caso, podemos resolver mais rapidamente, lembrando o conceito de fração:

$$15\% \text{ de } 20 = \frac{15}{100} \text{ de } 20 = \frac{15}{100} \times 20 = 3$$

Exercícios

1. Escreva sob a forma de percentagem as frações:
   - a. $\frac{2}{5}$
   - b. $\frac{3}{4}$
   - c. $\frac{3}{50}$
   - d. $\frac{1}{20}$
   - e. $\frac{5}{2}$

2. Escreva as taxas percentuais abaixo como razões, sob a forma mais simples possível:
   - a. 30%
   - b. 40%
   - c. 60%
   - d. 200%
   - e. 2,5%

3. Calcule:
   - a. 20% de 300;
   - b. 15% de R$ 150;
   - c. 70% de 80 animais;
   - d. 9% de 50.

4. Em uma classe de 60 alunos, faltaram 15. Qual a taxa de percentagem dos alunos presentes?

5. Em São Paulo colheram-se 1.300.000 sacas de café. Se 25% desta produção destinam-se ao consumo interno, qual a quantidade de sacas para este consumo?

6. Uma nota promissória, cujo valor era R$ 50.000, foi paga com um desconto de R$ 2.500. Qual a taxa de desconto?

7. Quarenta por cento dos alunos de uma escola são meninos. O total de alunos é 2.500. Quantas são as meninas e quantos são os meninos?

8. Doze por cento dos alunos de um colégio são internos. Os alunos externos são 924. Qual é o total de alunos do colégio? Quantos são os internos?

9. Vendi um objeto por R$ 60 e tive um lucro de 30% sobre o custo. Qual foi o lucro?

10. Vendi uma mercadoria recebendo 25% de entrada e o restante em três prestações de R$ 160 e uma de R$ 180. Qual o preço da mercadoria?

11. Por quanto devo vender um objeto que me custou R$ 150, para ter um lucro de 20% sobre o custo?

12. Um objeto foi vendido com 15% de lucro e outro semelhante com 35%. Por quanto foi vendido cada um, se os dois foram vendidos por R$ 180?

5. Sequência
Somatório

5.1 Sequência ou sucessão

Sequência ou **sucessão** é uma função cujo domínio é o conjunto dos números inteiros positivos (N*) ou um subconjunto finito do mesmo ({1, 2, 3, ..., n}).

No primeiro caso, dizemos que a sequência é **infinita** e no segundo, **finita**.

O conjunto imagem de uma sequência pode ser um conjunto qualquer.

Em nossos estudos, ficaremos restritos às **sequências reais finitas**, isto é, aquelas que têm para domínio um subconjunto finito dos números inteiros positivos e para conjunto imagem um subconjunto dos números reais.

Para indicarmos os elementos de uma sequência, lançamos mão de um recurso, o **índice**, que nada mais é que um numeral escrito à direita e um pouco abaixo da letra e que **indica a ordem** que esse elemento ocupa na sequência.

Assim, representando por:

$\mathbf{a_1}$: o primeiro termo (lemos: **a** índice **1**);

$\mathbf{a_2}$: o segundo termo (lemos: **a** índice **2**);

....

$\mathbf{a_n}$: o n-ésimo termo (lemos: **a** índice **n**),

indicamos uma sequência por:

$$(a_1, a_2, ..., a_n)$$

ou:

$$a_i \ (i = 1, 2, ..., n)$$

(lemos: **a** índice **i** sendo **i** igual a 1, 2, ..., n), onde a_i é o **termo geral**, a_n é o **último termo** e **n** é o **número de termos**.

5.2 Somatório

Para indicarmos a soma dos x_i valores de uma variável **x**, isto é, $x_1 + x_2 + \dots + x_n$, lançamos mão do símbolo Σ (letra grega, maiúscula: **sigma**), denominado, em Matemática, **somatório**.

Assim, a soma $x_1 + x_2 + \dots + x_n$ pode ser representada por:

$\sum_{i=1}^{5} x_i$ (lemos: somatório de **x** índice **i, i** variando de 1 até 5), isto é:

$$x_1 + x_2 + \dots + x_n = \sum_{i=1}^{5} x_i$$

Não havendo possibilidade de dúvidas, podemos indicar, mais simplesmente, por:

$$\sum x_i$$

Assim:

$$\sum x_i = x_1 + x_2 + \dots + x_n$$

Exemplo:

Sendo x ∈ (2, 4, 6), temos:

$$\left.\begin{array}{l} x_1 = 2 \\ x_2 = 4 \\ x_3 = 6 \end{array}\right\} \Rightarrow \sum x_i = x_1 + x_2 + x_3 = 2 + 4 + 6 = 12$$

Propriedades

1ª) Sendo c uma constante:

$$\sum_{i=1}^{n} c = n \times c$$

2ª) Sendo c uma constante e x uma variável:

$$\sum_{i=1}^{n} (c \times x_i) = c \sum_{i=1}^{n} x_i$$

3ª) Sendo x e y duas variáveis:

$$\sum_{i=1}^{n} (x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$$

NOTAS:

- $\sum_{i=1}^{n} (x_i y_i) \neq \sum_{i=1}^{n} x_i \times \sum_{i=1}^{n} y_i$
- $\left(\sum_{i=1}^{n} x_i\right)^2 \neq \sum_{i=1}^{n} x_i^2$

Exercícios resolvidos

1. Sendo $x \in (2, 5, 8, 9)$, dê os valores de x_1, x_2, x_3 e x_4:

$$x_1 = 2, x_2 = 5, x_3 = 8 \text{ e } x_4 = 9$$

2. Desenvolva o somatório $\sum_{i=1}^{4} x_i$:

$$\sum_{i=1}^{4} x_i = x_1 + x_2 + x_3 + x_4$$

3. Escreva $x_3 + x_4 + x_5$ sob a forma de somatório:

$$x_3 + x_4 + x_5 = \sum_{i=3}^{5} x_i$$

Exercícios

1. Desenvolva os somatórios:
   a. $\sum_{i=1}^{8} x_i$ b. $\sum_{i=3}^{6} x_i$ c. $\sum_{i=1}^{5} x_i$

2. Escreva sob a forma de somatório:
   a. $x_1 + x_2 + x_3 + x_4$
   b. $x_1 + x_2 + ... + x_7$
   c. $x_4 + x_5 + x_6 + x_7$
   d. $x_5 + x_6 + ... + x_{10}$

3. Dada a sequência (2, 5, 7, 10, 12, 13, 15) e sendo x_i o termo geral, determine os valores de x_1, x_2, x_3, ..., x_7.

4. Calcule, considerando a sequência do exercício anterior:
   a. $\sum_{i=1}^{7} x_i$ b. $\sum_{i=1}^{4} x_i$ c. $\sum_{i=3}^{7} x_i$ d. $\sum_{i=4}^{6} x_i$

6. Média aritmética

6.1 Média aritmética simples

Chamamos de **média aritmética** de um conjunto de valores o quociente da divisão da soma desses valores pelo número deles.

Indicando por $x_1, x_2, ..., x_n$ os **n** valores que a variável **x** pode assumir, e por $\bar{\mathbf{x}}$ a média aritmética, temos:

$$\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n} \quad \text{ou} \quad \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Exercício resolvido

1. Calcular a média aritmética do seguinte conjunto de números: 2, 3, 4, 5, 6.

Temos:

$$\bar{x} = \frac{\sum_{i=1}^{5} x_i}{5} = \frac{2 + 3 + 4 + 5 + 6}{5} = \frac{20}{5} = 4$$

6.2 Média aritmética ponderada

No caso de os valores estarem afetados por **pesos**, que são números indicadores da intensidade do valor no conjunto, a **média aritmética** se diz **ponderada**.

A **média aritmética ponderada** é igual ao quociente da divisão cujo dividendo é formado pela soma dos produtos dos valores pelos respectivos pesos e cujo divisor é a soma dos pesos.

Assim, se os valores $x_1, x_2, ..., x_n$ ocorrem $p_1, p_2, ..., p_n$ vezes, respectivamente, a média aritmética ponderada é dada por:

$$\bar{x} = \frac{x_1p_1 + x_2p_2 + ... + x_np_n}{p_1 + p_2 + ... + p_n} \quad \text{ou} \quad \bar{x} = \frac{\sum_{i=1}^{n}(x_ip_i)}{\sum_{i=1}^{n} p_i}$$

Exercício resolvido

1. Sabendo que um aluno obteve as notas 7, 6, 5 e 8 e que essas notas têm, respectivamente, os pesos 2, 2, 3 e 3, calcule a sua média.

Temos:

$$\bar{x} = \frac{\sum_{i=1}^{4} x_i p_i}{\sum_{i=1}^{4} p_i} = \frac{7 \times 2 + 6 \times 2 + 5 \times 3 + 8 \times 3}{2 + 2 + 3 + 3} = \frac{14 + 12 + 15 + 24}{10} = \frac{65}{10} = 6{,}5$$

Logo:

$$\bar{x} = 6{,}5$$

Exercícios

1. Os tempos de reação de um indivíduo a certos estímulos foram medidos por um psicologista como sendo (em segundos) 0,53; 0,46; 0,50; 0,49; 0,52; 0,53; 0,44; e 0,55, respectivamente. Determine o tempo médio de reação do indivíduo a esses estímulos.

2. Os graus de um estudante nas disciplinas de laboratório, leitura e declamação foram 7,1; 7,8 e 8,9, respectivamente. Se os pesos atribuídos a esses graus são 2, 4 e 5, respectivamente, qual é o grau médio do estudante?

3. Três professores de Economia atribuíram os graus médios de exame 7,5; 8,2 e 8,4 a suas respectivas classes, que se compunham de 32, 25 e 17 estudantes, respectivamente. Determine o grau médio para todas as classes.

4. Um conjunto de números é composto de seis **6**, sete **7**, oito **8**, nove **9** e dez **10**. Qual é a média aritmética dos números?

7. Fatorial

Sendo **n** um número natural diferente de zero, temos:

$$n! = n(n-1)\ (n-2) \times \ldots \times 3 \times 2 \times 1$$

Assim:

> **n!** (lemos: **ene fatorial**) é o produto de todos os números naturais de **n** até 1.

Exemplos:

$2! = 2 \times 1 = 2$

$3! = 3 \times 2 \times 1 = 6$

$4! = 4 \times 3 \times 2 \times 1 = 24$

> **NOTA:**
> - Por definição, tomamos:
>
> $0! = 1$
>
> $1! = 1$

Exercício resolvido

1. Calcule:

a. $\dfrac{5!}{5}$ b. $\dfrac{7!}{5!}$ c. $(n-2)!$ d. $\dfrac{n!}{(n-1)!}$

Temos:

a. $\dfrac{5!}{5} = \dfrac{\cancel{5} \times 4 \times 3 \times 2 \times 1}{\cancel{5}} = 4 \times 3 \times 2 \times 1 = 24$

b. $\dfrac{7!}{5!} = \dfrac{7 \times 6 \times \cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2} \times 1}{\cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2} \times 1} = 7 \times 6 = 42$

c. $(n-2)! = (n-2)\ (n-3)\ (n-4) \times \ldots \times 3 \times 2 \times 1$

d. $\dfrac{n!}{(n-1)!} = \dfrac{n\ \cancel{(n-1)}\ \cancel{(n-2)}\ \cancel{(n-3)} \times \ldots \times \cancel{3} \times \cancel{2} \times 1}{\cancel{(n-1)}\ \cancel{(n-2)}\ \cancel{(n-3)} \times \ldots \times \cancel{3} \times \cancel{2} \times 1} = n$

Exercício

1. Calcule:

a. $\dfrac{8!}{7!}$ b. $\dfrac{5(4!)}{5!}$ c. $\dfrac{3(4!)}{4(3!)}$ d. $\dfrac{(n-2)!}{(n-4)!}$

8. Coeficientes binomiais

Sendo **n** e **k** números naturais diferentes de zero, indicamos por:

$$\binom{n}{k}$$

o **coeficiente binomial** de **n** sobre **k** ou, simplesmente, **binomial** de **n** sobre **k**.

Temos:

$$\binom{n}{k} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!}$$

NOTA:

- Por definição, tomamos:

$$\binom{n}{0} = 1$$

Exercício resolvido

1. Calcule:

a. $\binom{5}{3}$ b. $\binom{4}{2}$

Temos:

a. $\binom{5}{3} = \dfrac{5 \times 4 \times 3}{3!} = \dfrac{5 \times \overset{2}{\cancel{4}} \times 3}{3 \times \underset{1}{\cancel{2}} \times 1} = 5 \times 2 = 10$

b. $\binom{4}{2} = \dfrac{4 \times 3}{2!} = \dfrac{\overset{2}{\cancel{4}} \times 3}{\underset{1}{\cancel{2}} \times 1} = 2 \times 3 = 6$

NOTA:

- Observe que os números de fatores do numerador e do denominador são sempre iguais.

8.1 Coeficientes binomiais complementares

Os coeficientes binomiais $\binom{n}{k}$ e $\binom{n}{n-k}$ são chamados **complementares**. Demonstra-se que:

$$\binom{n}{k} = \binom{n}{n-k}$$

Exemplo:

$\binom{7}{5}$ e $\binom{7}{2}$ são complementares; logo, $\binom{7}{5} = \binom{7}{2}$.

NOTA:

- Os coeficientes binomiais complementares são usados para simplificar cálculos.

Assim, para calcularmos $\binom{50}{48}$, empregamos o complementar:

$$\binom{50}{48} = \binom{50}{2} = \frac{50 \times 49}{2!} = \frac{\overset{25}{\cancel{50}} \times 49}{\underset{1}{\cancel{2}} \times 1} = 25 \times 49 = 1.225$$

Exercício

1. Calcule:

a. $\binom{8}{3}$

b. $\binom{100}{98}$

c. $\dfrac{\binom{42}{25}}{\binom{42}{17}}$

9. Binômio de Newton

Denominamos **binômio de Newton** toda expressão da forma:

$$(x + a)^n, \text{onde } n \in \mathbb{N}$$

O desenvolvimento de $(x + a)^n$ é dado por:

$$(x + a)^n = \binom{n}{0} x^n + \binom{n}{1} a x^{n-1} + \binom{n}{2} a^2 x^{n-2} + \ldots + \binom{n\,n\,1}{n-1} a^{n-1} x + \binom{n}{n} a^n$$

onde o termo que ocupa o lugar de ordem $\mathbf{k + 1}$ é:

$$T_{k+1} = \binom{n}{k} a^k x^{n-k}$$

Exercícios resolvidos

1. Desenvolva o binômio $(x + y)^6$.

Temos:

$$(x+y)^6 = \binom{6}{0} x^6 + \binom{6}{1} x^5y + \binom{6}{2} x^4y^2 + \binom{6}{3} x^3y^3 + \binom{6}{4} x^2y^4 + \binom{6}{5} xy^5 + \binom{6}{6} y^6$$

Como:

$$\binom{6}{0} = 1 = \binom{6}{6} \qquad \binom{6}{2} = \frac{6 \times 5}{2 \times 1} = 15 = \binom{6}{4}$$

$$\binom{6}{1} = 6 = \binom{6}{5} \qquad \binom{6}{3} = \frac{6 \times 5 \times 4}{3 \times 1 \times 1} = 20$$

vem:

$$(x + y)^6 = x^6 + 6x^5y + 15x^4y^2 + 20x^3y^3 + 15x^2y^4 + 6xy^5 + y^6$$

2. Determine o 5º termo do desenvolvimento de $(x + 2)^{10}$.

Lembrando que:

$$T_{k+1} = \binom{n}{k} a^k x^{n-k},$$

temos:

$$k + 1 = 5 \Rightarrow k = 4.$$

Logo:

$$T_5 = \binom{10}{4} 2^4 x^{10-4} = \frac{10 \times 9 \times 8 \times 7}{4 \times 3 \times 2 \times 1} \times 16x^6 = 3.360x^6$$

Daí:

$$T_5 = 3.360x^6$$

Exercícios

1. Desenvolva:

   a. $(3y + 1)^4$

   b. $\left(\dfrac{y}{2} + 2\right)^6$

   c. $(2x + 1)^5$

2. Determine:

   a. o 5º termo em $(p + q)^{10}$;

   b. o 4º termo em $\left(2 + \dfrac{b}{2}\right)^{10}$;

   c. o 6º termo em $(x + 2)^{13}$;

   d. o 5º termo em $(x + 3)^8$.

10. Função

10.1 Definição

Seja a equação:

$$y = 2x$$

É fácil constatar que para cada valor dado a **x** obtemos um e um só valor para **2x**.

Assim, dando a **x** os valores $\{-2, -1, 0, 1, 2, 3\}$, obtemos para **y** os valores $\{-4, -2, 0, 2, 4, 6\}$, isto é:

	−2	−1	0	1	2	3
	−4	−2	0	2	4	6

Podemos, então, dizer que **para cada valor de x existe um único valor para y.**

Neste caso, dizemos que **y é função de x** e escrevemos:

$$f: x \rightarrow y = 2x;$$

x e **y** são as **variáveis** da função; **x** é a **variável independente** e **y**, a **dependente.**

A tabela acima dá origem aos pares ordenados $(-2, -4)$, $(-1, -2)$, $(0, 0)$, $(1, 2)$, $(2, 4)$ e $(3, 6)$, que dizemos **pertencerem** à função definida por $y = 2x$.

NOTA:

- Como **x** pode tomar os valores −2, −1, 0, 1, 2 e 3, dizemos que $x \in \{-2, -1, 0, 1, 2, 3\}$ (lemos: **x** pertence ao conjunto formado pelos elementos −2, −1, 0, 1, 2 e 3).

Resolva

1. Faça uma tabela para cada uma das funções abaixo, com $x \in \{-2, -1, 0, 1, 2, 3\}$:

a. $f: x \to y = 3x - 5$ b. $f: x \to y = x^2 - 3$ c. $f: x \to y = 2x^2 - x$

10.2 Gráfico de uma função

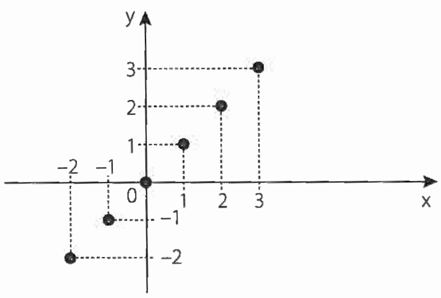
São dadas a função:

$$f: x \to y = x$$

e a tabela correspondente:

	−2	−1	0	1	2	3
	−2	−1	0	1	2	3

Representando, em um sistema coordenado cartesiano ortogonal, os pares ordenados resultantes da tabela: (−2, −2), (−1, −1), (0, 0), (1, 1), (2, 2), (3, 3), vem:



O conjunto de pontos de intersecção das perpendiculares forma o gráfico da função.

Seja, agora, a função:

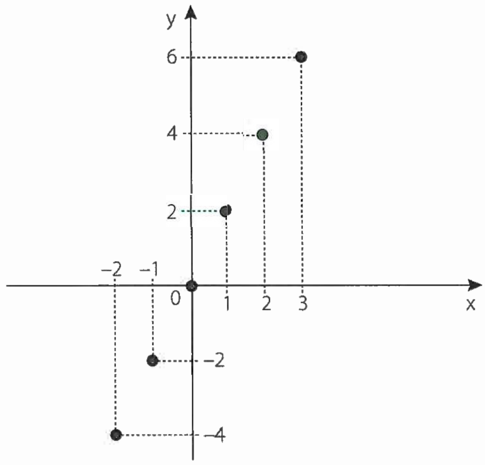
$$f: x \to y = 2x, \text{ com } x \in \mathbb{R}$$

Por ser o conjunto dos números reais um **conjunto denso**, os pontos do gráfico ficarão intimamente ligados entre si, dando origem a uma linha contínua.

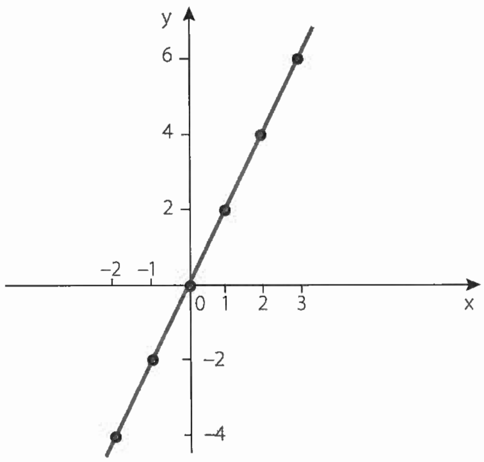
Na impossibilidade de representarmos todos os valores de **x** e de **y**, construímos uma tabela a partir de alguns valores de **x**:

	−2	−1	0	1	2	3
	−4	−2	0	2	4	6

Representando esses pontos no sistema de eixos coordenados, obtemos:



Podemos comprovar, com uma régua, que os seis pontos correspondentes estão em linha reta, o que nos leva a **prever** que o gráfico completo dessa função é **uma reta passando por esses pontos**:



NOTA:

- Podemos afirmar **apenas** que esse gráfico é **provavelmente** uma reta.

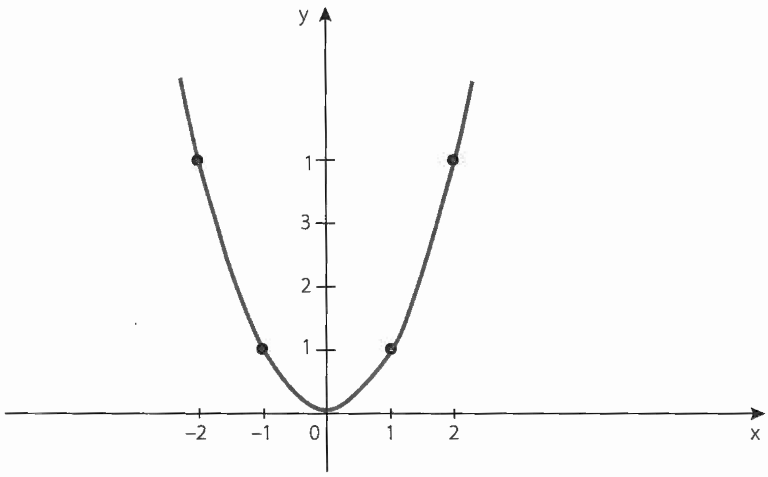
Consideremos, ainda, a função:

$$f: x \rightarrow y = x^2, \text{ com } x \in \mathbb{R}$$

Determinando os valores de **y** a partir de valores arbitrários de **x**, obtemos a tabela:

	−2	−1	0	1	2
	−4	1	0	1	4

que nos dá pontos do plano. Como x ∈ $\mathbb{R}$, podemos ligar esses pontos por meio de uma linha contínua:



> **NOTA:**
> - Como anteriormente, **presumimos** que a curva correspondente seja uma **parábola**.

10.3 Função do 1º grau

Denominamos **função do 1º grau** toda função definida por:

$$y = ax + b, \text{ com } a, b \in \mathbb{R} \text{ e } a \neq 0$$

Exemplos:

$y = 2x$, onde $a = 2$ e $b = 0$

$y = x - 2$, onde $a = 1$ e $b = -2$

$y = 4 - 3x$, onde $a = -3$ e $b = 4$

> **NOTA:**
> - Os números reais **a** e **b** são denominados **coeficientes** ou **parâmetros**.

10.4 Gráfico da função do 1º grau

Em um sistema coordenado cartesiano ortogonal, demonstra-se que:

> O gráfico de uma função do 1º grau é uma reta oblíqua.

Assim, como dois pontos determinam uma e uma só reta, para traçarmos o gráfico de uma função do 1º grau é o bastante determinarmos dois de seus pontos.

Exemplo:

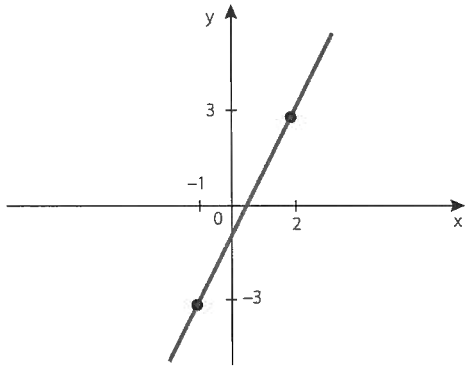
Seja a função do 1º grau:

$$f: x \to y = 2x - 1$$

Temos:

$$x = -1 \Rightarrow y = 2(-1) - 1 = -3 \Rightarrow (-1, -3) \in f$$
$$x = 2 \Rightarrow y = 2 \times 2 - 1 = 3 \Rightarrow (2, 3) \in f$$

Logo:



10.5 Equação da reta que passa por dois pontos dados

Consideremos o seguinte problema:

Qual a equação da reta que passa pelos pontos (5, 10) e (2, 1)?

Como toda função do 1º grau é definida por uma equação da forma:

$$y = ax + b \qquad (1)$$

e como a reta em questão passa pelos pontos (5, 10) e (2, 1), isto quer dizer que esses pares ordenados pertencem à equação (1). Logo:

$$10 = a \times 5 + b \text{ e } 1 = a \times 2 + b,$$

o que nos dá o sistema de equações simultâneas:

$$\begin{cases} 5a + b = 10 \\ 2a + b = 1 \end{cases}$$

Resolvendo pelo processo de adição, obtemos:

$$\begin{array}{l} 5a \cancel{+b} = 10 \\ \underline{-2a \cancel{-b} = -1} \\ 3a \qquad = 9 \quad \Rightarrow a = \dfrac{9}{3} \Rightarrow a = 3 \end{array}$$

Daí:

$$2 \times 3 + b = 1 \Rightarrow 6 + b = 1 \Rightarrow b = 1 - 6 \Rightarrow b = -5.$$

Substituindo esses valores de **a** e **b** em (1), temos:

$$y = 3x - 5,$$

que é a equação pedida.

Exercícios

1. Faça uma tabela de valores para cada uma das equações abaixo:
   a. $y = 3x + 1$
   b. $y = x - 3$
   c. $y = x^2 + 1$
   d. $y = \sqrt{x}$ (**Sugestão:** $x \in \{0, 1, 4, 9, 16\}$.)

2. Fazendo um exame das tabelas obtidas no exercício anterior, diga qual das equações não define uma função.

3. Represente graficamente as funções definidas por:
   a. $y = 2x - 3$
   b. $y = 4 - x$
   c. $y = 3x + 2$
   d. $y = x$

4. Determine a função do 1º grau que passa pelos pontos:
   a. (0, 0) e (2, 2)
   b. (5, 0) e (0, –3)
   c. (1, –2) e (0, 0)
   d. (1, 1) e (–2, –5)

10.6 Pontos notáveis

Ponto em que a reta corta o eixo dos x

O ponto em que a reta corta o eixo dos **x** é aquele de **ordenada nula**; por isso é denominado **abscissa na origem**.

$$\text{Se } y = 0 \Rightarrow 0 = ax + b \Rightarrow x = \frac{-b}{a}.$$

Logo, o ponto:

$$\left(\frac{-b}{a}, 0\right) \in f$$

é aquele em que a **reta corta o eixo dos x**.

Ponto em que a reta corta o eixo dos y

O ponto em que a reta corta o eixo dos **y** é aquele de **abscissa nula**; por isso é denominado **ordenada na origem**.

$$\text{Se } x = 0 \Rightarrow y = a \times 0 + b \Rightarrow y = b.$$

Logo, o ponto:

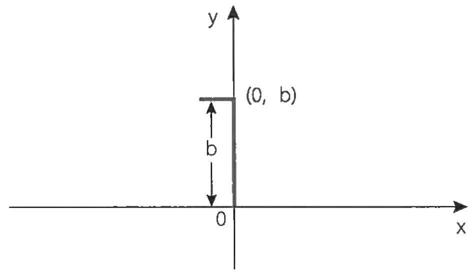
$$(0, b) \in f$$

é o ponto procurado.

10.7 Significado dos coeficientes

Coeficiente b

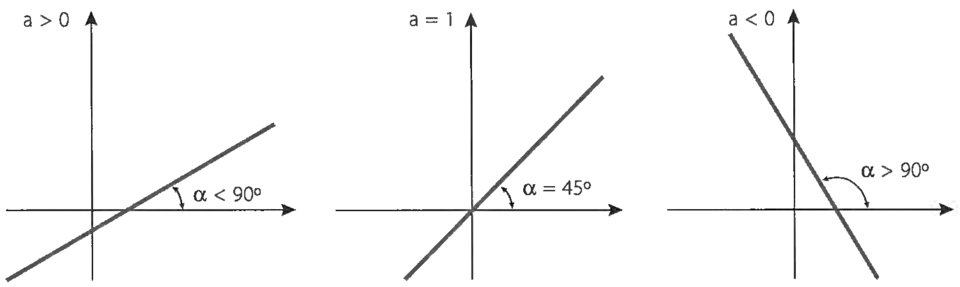
Como vimos, o coeficiente **b** determina o ponto em que a reta corta o eixo dos **y**, isto é, **b** é o valor algébrico do segmento determinado pela origem e pelo ponto de intersecção da reta com o eixo dos **y**:



Por essa razão o coeficiente **b** é denominado **coeficiente linear**.

Coeficiente a

Analisando os gráficos da função do 1º grau traçados até agora, vemos que:



- se $a > 0 \Rightarrow 0° < \alpha < 90°$;
- se $a = 1 \Rightarrow \alpha = 45°$;
- se $a < 0 \Rightarrow 90° < \alpha < 180°$.

Assim, podemos concluir que a medida do ângulo α, formado pela reta com o sentido positivo do eixo dos **x**, depende do valor do coeficiente **a**, razão pela qual o denominamos **coeficiente angular**.

COLETÂNEA DE QUESTÕES OBJETIVAS

1. Ao nascer, os bebês são pesados e medidos, para se saber se estão dentro das tabelas de peso e altura esperados. Estas duas variáveis são:
   a. qualitativas.
   b. ambas discretas.
   c. ambas contínuas.
   d. contínua e discreta, respectivamente.
   e. discreta e contínua, respectivamente.

2. A parcela da população convenientemente escolhida para representá-la é chamada de:
   a. variável.
   b. rol.
   c. amostra.
   d. dados brutos.
   e. nada podemos afirmar, porque a informação é incompleta.

3. Na administração de um sistema escolar de certo município, 70% das despesas vão para o ensino, 12% para a administração e manutenção e 18% para órgãos auxiliares, encargos fixos e despesas ocasionais. O gráfico que melhor representa essa situação é:
   a. o linear simples.
   b. o de barras.
   c. o de setores.
   d. o histograma.

4. Um conjunto de 100 notas de Matemática, de alunos do sexo masculino, tiradas dos arquivos da secretaria da escola, constitui:
   a. um rol.
   b. uma relação de dados brutos.
   c. uma tabela.
   d. uma distribuição de frequência.

5. Por definição, **rol é qualquer série ordenada de valores referentes a uma mesma variável**. Então, dadas as séries da mesma variável **x**:

I. –2, 4, 5, 6, 7
II. 1, 3, 3, 6, 7
III. 8, 7, 5, 2,1
IV. 5, 4, 4, –1

podemos afirmar que:

a. todas elas constituem róis.
b. só a série I constitui um rol.
c. a série II não é um rol, mas as outras sim.
d. apenas as séries I e IV não são róis.
e. somente a série III é um rol, as demais não.

Com base na distribuição abaixo, resultante de pesos de moças, responda às questões de **6** a **9**:

	42 ⊢ 44	44 ⊢ 46	46 ⊢ 48	48 ⊢ 50	50 ⊢ 52
	22	24	56	59	25

6. Nessa distribuição, o intervalo usado é:

a. aberto à esquerda.
b. fechado à esquerda.
c. aberto.
d. fechado.
e. aberto à esquerda e à direita.

7. Nessa distribuição, os pontos médios são:

a. 42, 44, 46, 48, 50.
b. 44, 46, 48, 50, 52.
c. 86, 90, 94, 98, 102.
d. 43, 45, 47, 49, 51.

8. Nessa distribuição, a amplitude total do fenômeno estudado é:

a. 42.
b. 10.
c. 52.
d. 2.
e. 94.

9. Nessa distribuição, a amplitude dos intervalos de classe é:

a. 10.
b. 2.
c. 52.
d. 94.
e. 50

10. As regras básicas para se construir uma distribuição de frequência são:

I. Nenhum dado deve ser excluído.
II. Nenhum dado deve ser contado mais de uma vez.
III. As classes têm de ser mutuamente exclusivas.
IV. O campo de variação da variável tem de ser esgotado.

Destas regras:

a. todas estão corretas.
b. todas estão erradas.
c. só a segunda está errada.
d. só a terceira está errada.
e. só a quarta está correta.

11. Os gráficos próprios de uma distribuição de frequência são:

a. colunas, curva de frequência e histograma.
b. polígono de frequência e histograma.
c. colunas, curva de frequência e polígono de frequência.
d. gráfico em setor, gráfico em barra, curva de frequência e curva normal.
e. colunas, barra, setor e curva de frequência.

12. Um teste de inteligência, aplicado aos alunos das 4ªs séries do 1º grau da Escola **A**, apresentou os seguintes resultados:

PONTOS DO QI	90 ⊢ 95	95 ⊢ 100	100 ⊢ 105	105 ⊢ 110	110 ⊢ 115	115 ⊢ 120	120 ⊢ 125	125 ⊢ 130	130 ⊢ 135	135 ⊢ 140
Nº DE ALUNOS	40	60	140	160	180	120	40	30	20	10

A frequência relativa da classe modal é:

a. 0,200.
b. 0,225.
c. 0,250.
d. 0,500.

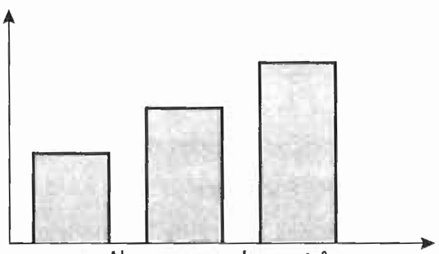
13. Na construção de qual dos gráficos citados — **histograma** e **polígono de frequência** — usamos, obrigatoriamente, as frequências acumuladas?

a. Só no primeiro.
b. Só no segundo.
c. Em ambos.
d. Em nenhum.
e. No primeiro, às vezes, dependendo do tipo de variável.

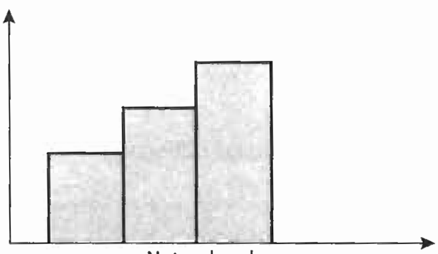
14. As classes de uma distribuição de frequência devem ser mutuamente exclusivas para que:

a. nenhum dado seja excluído.
b. nenhum dado seja contado mais de uma vez.
c. todos os dados sejam computados.
d. possam exaurir totalmente o campo de variação.
e. os limites inferiores e superiores sejam levados em consideração.

15.



Alunos aprovados em três classes de 5ª série



Notas dos alunos de 5ª série

Estes dois gráficos são, respectivamente:

a. gráficos em colunas.
b. histogramas.
c. gráfico em colunas e polígono de frequência.
d. histograma e polígono de frequência.
e. gráfico em colunas e histograma.

16. Das afirmações:

I. Tanto o histograma como o polígono de frequência são gráficos próprios da distribuição de frequência, são gráficos de análise, os quais devem ser feitos só quando a variável for contínua.

II. Tanto o polígono de frequência como o histograma são gráficos próprios da distribuição de frequência, são gráficos de análise, e devem ser feitos só quando a variável for discreta.

III. Tanto o histograma como o polígono de frequência são gráficos de análise, próprios da distribuição de frequência, e podem ser feitos para qualquer tipo de variável, desde que ela seja quantitativa.

IV. O histograma é um gráfico em colunas, mas qualquer gráfico em colunas não é necessariamente um histograma.

a. II e III são falsas.
b. a IV é falsa.
c. apenas a I é verdadeira.
d. todas são verdadeiras.
e. todas são falsas.

17. Das afirmações:

I. A média aritmética ficará aumentada (ou diminuída) da quantidade que for adicionada (ou subtraída) a (de) todos os valores da série.

II. A média aritmética, por ser um valor representativo, depende de todos os valores da série ou distribuição de frequência.

III. A média aritmética pode não ser considerada um valor típico da distribuição de frequência ou rol.

IV. A moda pode ser considerada como um valor representativo que envolve todos os elementos do rol ou distribuição de frequência.

V. A média, a moda e a mediana são valores de posição.

a. somente a I é correta.
b. todas são corretas.
c. II e III são incorretas.
d. IV é incorreta.
e. todas são incorretas.

18. Na tabela primitiva abaixo:

6, 2, 7, 6, 5, 4,

a soma dos desvios em relação à média é igual a:

a. –4.
b. 8.
c. 0.
d. 25.
e. 4.

19. Dados os conjuntos de valores abaixo:

A = {3, 5, 6, 8, 9, 10, 10, 10, 11, 12, 17}
B = {4, 5, 7, 10, 11, 13, 15}
C = {2, 3, 4, 5, 5, 5, 5, 6, 7, 8, 8, 8, 8, 9, 10, 11}

em relação à moda, podemos dizer que:

I. **A** é unimodal e a moda é 10.
II. **B** é unimodal e a moda é 10.
III. **C** é bimodal e as modas são 5 e 8.

Então:

a. estas afirmações estão todas corretas.
b. estas afirmações estão todas erradas.
c. I e II estão corretas.
d. I e III estão corretas.
e. II e III estão corretas.

20. Um professor, após verificar que toda a classe obteve nota baixa, eliminou as

questões que não foram respondidas pelos alunos. Com isso, as notas de todos os alunos foram aumentadas de três pontos. Então:

a. a média aritmética ficou alterada, assim como a mediana.
b. apenas a média aritmética ficou alterada.
c. apenas a mediana ficou alterada.
d. não houve alteração nem na média nem na mediana.
e. nada podemos afirmar sem conhecer o número total de alunos.

21. No conjunto abaixo, correspondente a notas de Inglês de 15 alunos:

{1, 2, 3, 8, 5, 7, 6, 9, 4, 6, 2, 10, 3, 5, 3},

a mediana é:

a. 5,0 alunos.
b. nota 5,0.
c. 9,0 alunos.
d. nota 9,0.
e. nota 5,5.

22. Das afirmações abaixo:

A. Quando se ordenam valores não agrupados segundo sua grandeza, a mediana é o ponto médio desta série.
B. Quando os valores de uma série contínua estão agrupados em uma distribuição de frequência, a mediana é, por definição, o ponto que corresponde a 50% da distribuição.
C. Quando desejamos o ponto médio exato de uma distribuição de frequência, basta calcular a mediana.
D. Quando existem valores extremos que afetam muito o cálculo da média, para representá-la devemos dar preferência à mediana.

a. todas estão incorretas.
b. todas estão corretas.
c. apenas a **A** está incorreta.
d. apenas a **D** está incorreta.
e. apenas a **B** está correta.

Com base na tabela abaixo, que corresponde às notas de Estatística de uma classe, responda às questões **23** e **24**:

	1	2	3	4	5	6	7	8	9
	2	6	9	12	14	9	5	4	1

23. Para essa tabela, a mediana é:

a. 31. b. 5. c. 6. d. 7. e. 5,5.

24. Então, acima da mediana temos:

a. 15 alunos.
b. 18 notas.
c. 33 notas.
d. 19 alunos.

25. A média aritmética dos valores 2, 3, –5, 6, –7, 2, 0, 8, –3, 5, 10 é:

a. –1,9.
b. 1,9.
c. 3,2.
d. 4,7.

26. Na série abaixo, composta de notas de Matemática:

6, 2, 8, 6, 3, 0, 4, 2, 6, 7, 10, 3, 6,

a média aritmética, a mediana e a moda são, respectivamente:

a. 4,85; 6,5 e 6.
b. 4,85; 6 e 6.
c. 5,33; 6 e 6.
d. 5,33; 6,5 e 6.

27. A mediana da série 1 3 8 15 10 12 7 é:

a. 15.
b. 10.
c. 7.
d. 3,5.
e. Nenhuma das anteriores.

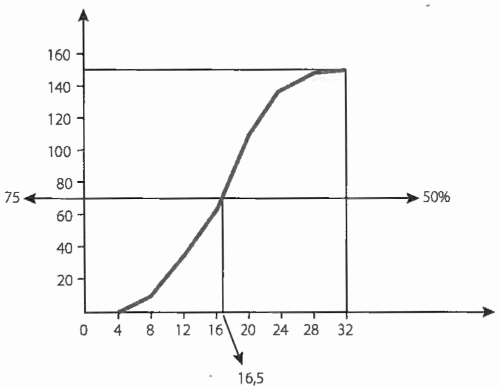
28. Numa pesquisa de opinião, 80 pessoas são favoráveis ao divórcio, 50 são desfavo-

ráveis, 30 são indiferentes e 20 ainda não têm opinião formada a respeito do assunto. Então, a média aritmética será:

a. 180, porque todos opinaram somente uma vez.
b. 40, porque é a média entre os valores centrais 50 e 30.
c. 45.
d. 1, porque todos opinaram somente uma vez.
e. Não há média aritmética.

29. O gráfico abaixo foi construído a partir da seguinte distribuição de frequência:

PONTOS DE UM TESTE	4 ⊢ 8	8 ⊢ 12	12 ⊢ 16	16 ⊢ 20	20 ⊢ 24	24 ⊢ 28	28 ⊢ 32
PESSOAS	10	25	35	40	25	10	5



Nesse caso, o valor 16,5 é:

a. a mediana.
b. a média aritmética.
c. a moda.
d. a média harmônica.

30. Qual a percentagem de valores que se localiza entre o último quartil e o P_{81}?

a. 6%
b. 19%
c. 56%
d. 77%
e. 81%

31. O sexagésimo percentil divide a área de uma distribuição em quantas partes?

a. 2
b. 6
c. 40
d. 60
e. 100

32. Se numa distribuição há 500 valores, então entre o segundo quartil e o quinquagésimo percentil quantos valores haverá?

a. 7
b. 13
c. 42
d. 48
e. Não haverá valores.

33. A nota média dos alunos de uma classe foi 7 e a das alunas, 9. O número de alunos era 20 e o das alunas, 30. Então, a nota média da classe toda foi:

a. 7.
b. 7,8.
c. 8.
d. 8,2.
e. 9.

34. Um relatório mostrou, entre outras coisas, que numa região polar a temperatura média é de –23 °C e o desvio padrão é –5 °C. Com base nestas informações, podemos afirmar que:

a. o relatório está impreciso e deve ser completado com o rol.
b. o relatório está correto e deve ser aceito.
c. o relatório está incompleto e deve ser completado com o rol.
d. o relatório está bom, desde que se tenha o rol das temperaturas.
e. o relatório está errado e deve ser rejeitado.

35. Um coeficiente de variação é uma razão, geralmente percentual, entre:

a. a média e a mediana.
b. o desvio padrão e a média aritmética.
c. o desvio padrão e a mediana.
d. a média aritmética e o número de casos.

36. Num teste de Conhecimentos Gerais, a média das questões certas foi 57,5 e o desvio padrão 5,98. A variabilidade relativa das classes foi de:

a. 5,75%. **c.** 10,4%.
b. 9,62%. **d.** 11,4%.

37. Para a série de valores 0, –1, –2, 5, 4, –3, –7, 2, –4 e 6:

a. a média é 3,4 e a variância 16.
b. a média é zero e a variância 4.
c. a média é zero e a variância 16.
d. a média é 3,4 e a variância 4.
e. a média é zero mas a variância é impossível calcular.

38. Os resultados de uma prova de Estudos Sociais estão normalmente distribuídos (curva de Gauss ou normal). Sabe-se que z = 0,5 corresponde, na curva normal, a uma área de 0,1915. Indique a percentagem dos resultados que diferem da média aritmética de mais da metade do desvio padrão.

a. 61,70% **c.** 38,30%
b. 57,45% **d.** 19,15%

39. Qual a percentagem de casos acima da mediana, numa distribuição normal?

a. 25% **c.** 68%
b. 50% **d.** 75%

40. O preço de determinado bem, em 1990, era R$ 10; considerando-se esse preço igual a 100, em 1993, o preço relativo para o mesmo bem, vendido a R$ 92, é:

a. R$ 950. **d.** R$ 920.
b. R$ 970. **e.** R$ 910.
c. R$ 930.

41. Em 1990, o preço de uma mercadoria era 60% menor do que o preço da mesma mercadoria em 1991 e, em 1992, era 80% superior ao de 1991. O aumento de preço em 1992, tendo por base o preço de 1990, foi de:

a. 120%. **d.** 300%.
b. 140%. **e.** 450%.
c. 148%.

42. Considere a seguinte série:

ANOS	1990	1991	1992	1993
EXPORTAÇÃO (toneladas)	48.000	54.000	40.500	57.500

Os índices relativos para 1991, 92 e 93, sendo 1990 = 100, são:

a. 112,5; 84,4 e 119,8.
b. 111,5; 83,2 e 112,8.
c. 112,5; 84,3 e 119,7.
d. 113,5; 82,3 e 111,4.
e. 114,5; 81,4 e 111,9.

43. Se os salários dos empregados de uma empresa aumentam em 20% em dado período, enquanto o Índice de Preços aumenta 10%, então, o aumento real de salário, durante o período, foi:

a. de 10%.
b. menor do que 10%.
c. maior do que 10%.
d. nulo.

44. Considerando a série abaixo:

MERCADORIAS	PREÇOS				
	1990	1991	1992	1993	1994
A	150	150	160	180	180
B	450	320	380	420	390
C	180	190	190	210	220

os índices médios relativos para 1990, 91, 92, 93 e 94, tomando como ano-base 1991, são:

a. 112, 100, 120, 110 e 121.
b. 119, 122, 115, 115 e 109.
c. 112, 100, 109, 121 e 119.
d. 113, 111, 112, 123 e 118.
e. 114, 109, 113, 116 e 101.

RESPOSTAS

CAPÍTULO 2 — POPULAÇÃO E AMOSTRA

RESOLVA (p. 9)
Qualitativa: a.
Quantitativas discretas: b, c, d.
Quantitativa contínua: e.

EXERCÍCIO (p. 10)
Quantitativas discretas: c, d, i, j.
Quantitativas contínuas: b, e, g, h, l.

EXERCÍCIOS (p. 15)
1. 002 - 014 - 016 - 034 - 039 - 053 - 054 - 056 - 062 - 066 - 076 - 082 - 094 - 096 - 099 - 105 - 110 - 118 - 123
3. 94 - 79 - 129 - 84 - 56 - 95 - 123 - 123 - 81 - 128 - 110 - 120 - 95 - 76 - 52 - 62 - 65 - 71 - 80 - 63 - 95 - 75 - 80 - 149 - 103 - 108
6. 30
7. 1.648°

CAPÍTULO 3 — SÉRIES ESTATÍSTICAS

EXERCÍCIOS (p. 21)
1. **a.** histórica **d.** histórica
   b. específica **e.** específica-histórica
   c. geográfica **f.** geográfica-histórica

4.

IMPORTAÇÃO DE MERCADORIAS
BRASIL — 1993

PAÍSES	QUANTIDADE (t)	VALOR (US$ 1.000)
Arábia Saudita	14.839.804	1.469.104
Estados Unidos	10.547.889	6.034.946
Japão	561.024	1.519.943

FONTE: Ministério da Fazenda.

EXERCÍCIOS (p. 28)
1. 42,9 + 25,7 + 22,0 + 9,4 = 100,0
2. **a.** 1,0%; 0,4%; 1,4%; 0%
   b. 0,7%
3. **b.** 18,0 + 29,3 + 24,1 + 28,6 = 100,0
   c. 162,5; 82,3; 118,9
   d. 100,0; 162,5; 133,6; 158,9
4. 129,6 hab/km²
5. **a.** 27,2 hab/km²
   b. 18,3‰
   c. 6,2‰
6. cidade **B**

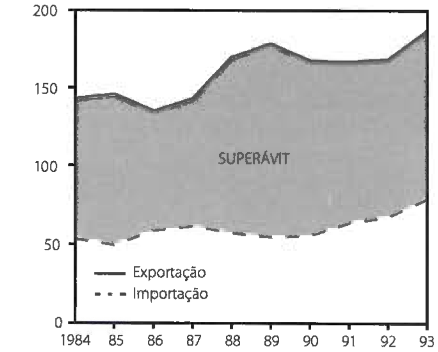
7. **a.** 10,2%; 14,3%; 25,5%; 14,9%
b. 16,1%
c. 90,9%; 100%; 85,7%; 82,5%
d. 90,1%
e. 83,3%; —; 37,5%; 0%
f. 53,3%
g. 46,7%
h. 85,1%
i. 5,6%

CAPÍTULO 4 — GRÁFICOS ESTATÍSTICOS

EXERCÍCIOS (p. 43)

1.

COMÉRCIO EXTERIOR
BRASIL — 1984-93

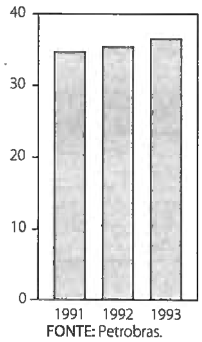


FONTE: Min. Indústria, Comércio e Turismo.

2.

a.

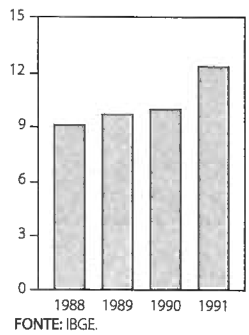
PRODUÇÃO BRASILEIRA
DE PETRÓLEO BRUTO
1991-93



FONTE: Petrobras.

b.

ENTREGA DE GASOLINA
PARA CONSUMO
BRASIL — 1988-91

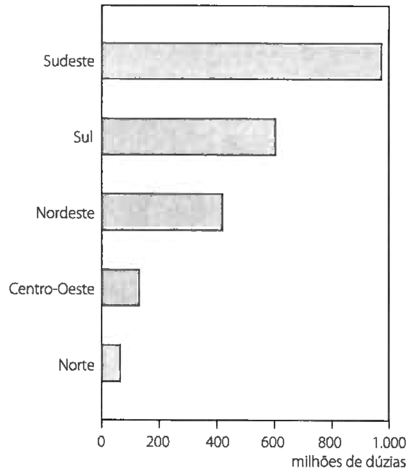


FONTE: IBGE.

3.

a.

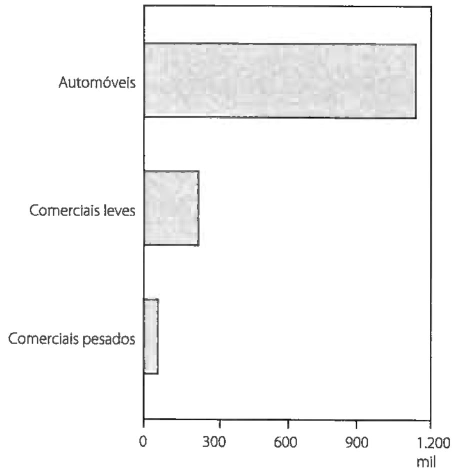
PRODUÇÃO DE OVOS
DE GALINHA
BRASIL — 1992



FONTE: IBGE.

b.

PRODUÇÃO DE VEÍCULOS DE
AUTOPROPULSÃO
BRASIL — 1993

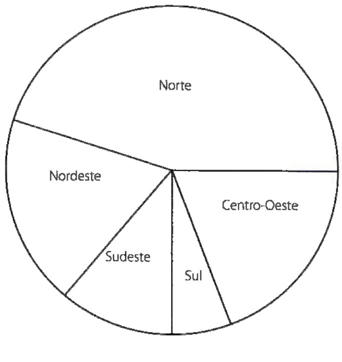


FONTE: IBGE.

4.

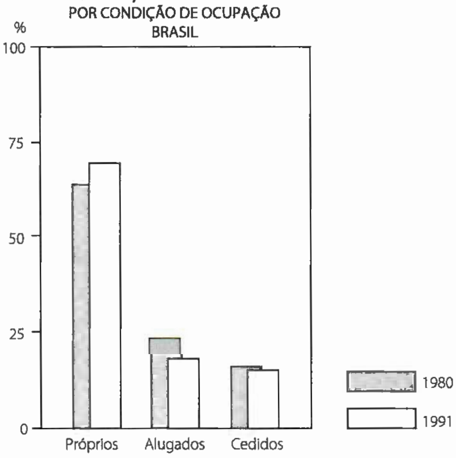
a.

ÁREA TERRESTRE BRASIL



b. PRODUÇÃO DE FERRO-GUSA
BRASIL — 1993

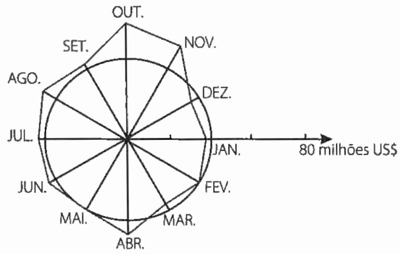
5. PROPORÇÃO DOS DOMICÍLIOS
POR CONDIÇÃO DE OCUPAÇÃO
BRASIL



FONTE: IBGE.

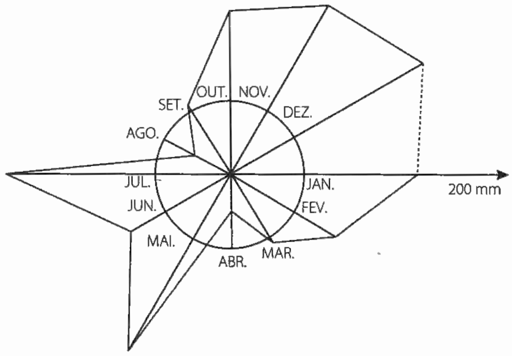
6.

a. VENDA DE VACINAS
CONTRA AFTOSA
BRASIL — 1992



FONTE: Sindan.

b. PRECIPITAÇÃO PLUVIOMÉTRICA
FLORIANÓPOLIS — 1993



FONTE: Ministério da Agricultura.

CAPÍTULO 5 — DISTRIBUIÇÃO DE FREQUÊNCIA

EXERCÍCIOS (p. 59)

1.

[redacted]	30 ⊢ 40	40 ⊢ 50	50 ⊢ 60	60 ⊢ 70	70 ⊢ 80	80 ⊢ 90	90 ⊢ 100
	4	6	9	11	9	7	4

2.

[redacted]	1	2	3	4	5	6
	6	8	9	7	10	10

3.

NOTAS	62 ⊢ 68	68 ⊢ 74	74 ⊢ 80	80 ⊢ 86	86 ⊢ 92	92 ⊢ 98	98 ⊢ 104	104 ⊢ 110
f_i	5	14	16	24	16	13	10	2

4.

[redacted]	10	11	12	13	14	15	16	17
	1	3	4	5	7	2	1	1

5. fr_i: 0,1; 0,25; 0,35; 0,225; 0,075
F_i: 4; 14; 28; 37; 40
Fr_i: 0,1; 0,35; 0,70; 0,925; 1,000

6. **a.** 40
b. 0,05; 0,125; 0,3; 0,25; 0,2; 0,075
c. 2; 7; 19; 29; 37; 40
d. 0,05; 0,175; 0,475; 0,725; 0,925; 1,000

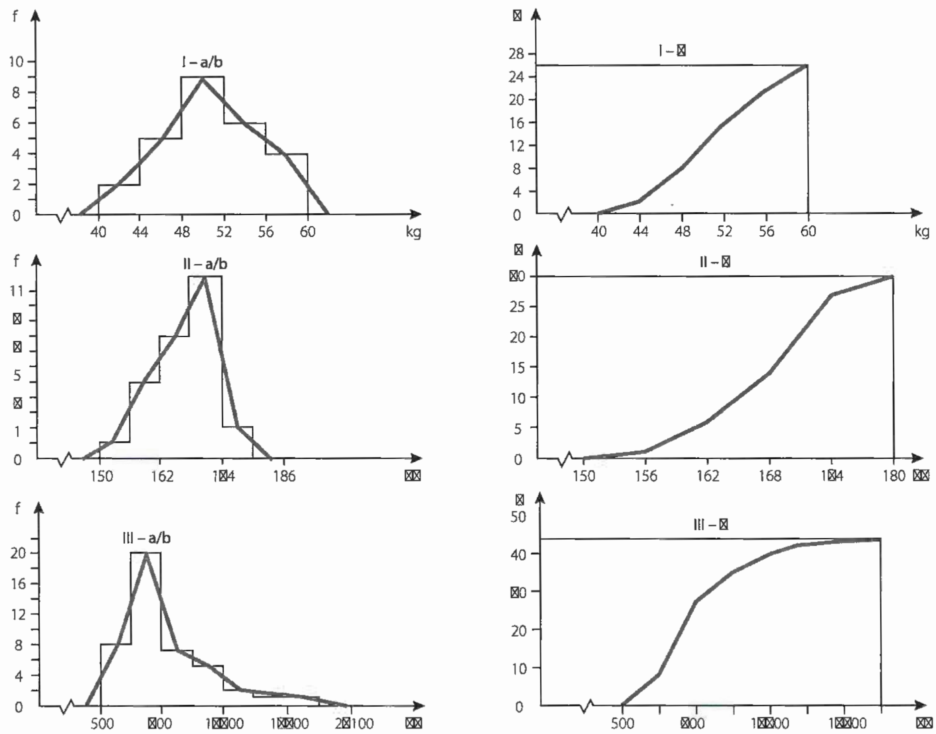
7. **a.** 900 **f.** 76
b. 800 **g.** 0,155
c. 1.000 **h.** 262
d. 950 **i.** 194
e. 100 **j.** 138
l. 29,5% **o.** i = 3
m. 19% **p.** i = 5
n. 78%

8. **a.** 20 **b.** 15 **c.** 46 **d.** 20 **e.** 65,7%

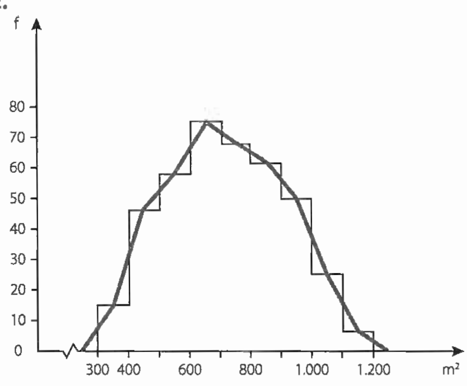
9. **a.** f_i: 1; 3; 4; 5; 3; 2; 1; 1
fr_i: 0,05; 0,15; 0,2; 0,25; 0,15; 0,1; 0,05; 0,05
F_i: 1; 4; 8; 13; 16; 18; 19; 20
b. classes: $6 \vdash 8$; $12 \vdash 14$
x_i: 3; 9; 11; 15
f_i: 18; 11; 7
F_i: 4; 12; 57; 100
fr_i: 0,08; 0,15; 0,11

EXERCÍCIOS (p. 69)

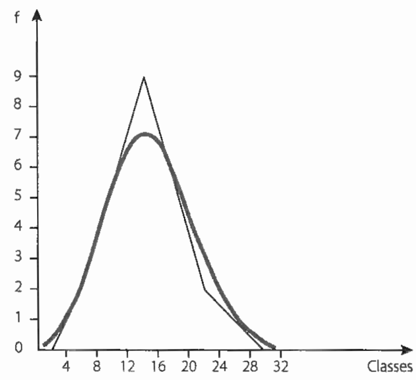
1.



2.



3.

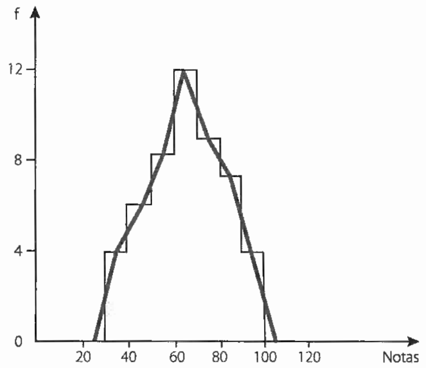


4. **a.** 100 ⊢ 110
b. 110
c. 139
d. 14
e. 80 ⊢ 90 e 90 ⊢ 100; 40 ⊢ 50 e 140 ⊢ 150
f. 50 ⊢ 60 e 120 ⊢ 130
g. 48
h. 54

5. **a.** J invertido **b.** J **c.** J **d.** J invertido **e.** U

6.

NOTAS	x_i	f_i	F_i	fr_i
30 ⊢ 40	35	4	4	0,08
40 ⊢ 50	45	6	10	0,12
50 ⊢ 60	55	8	18	0,16
60 ⊢ 70	65	12	30	0,24
70 ⊢ 80	75	9	39	0,18
80 ⊢ 90	85	7	46	0,14
90 ⊢ 100	95	4	50	0,08
		$\Sigma = 50$		$\Sigma = 1,00$

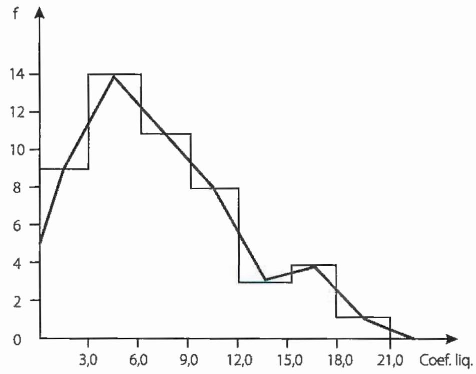


7.

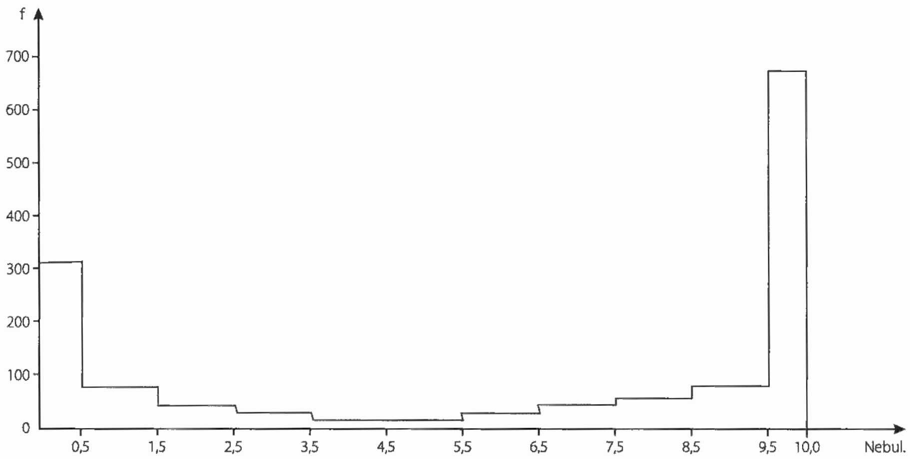
a.

COEFICIENTE LIQUIDEZ	f_i
0,0 ⊢ 3,0	9
3,0 ⊢ 6,0	14
6,0 ⊢ 9,0	11
9,0 ⊢ 12,0	8
12,0 ⊢ 15,0	3
15,0 ⊢ 18,0	4
18,0 ⊢ 21,0	1
	$\Sigma = 50$

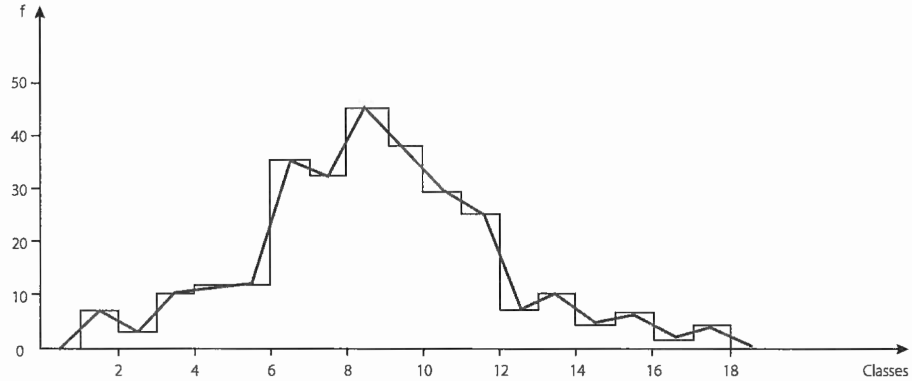
b.



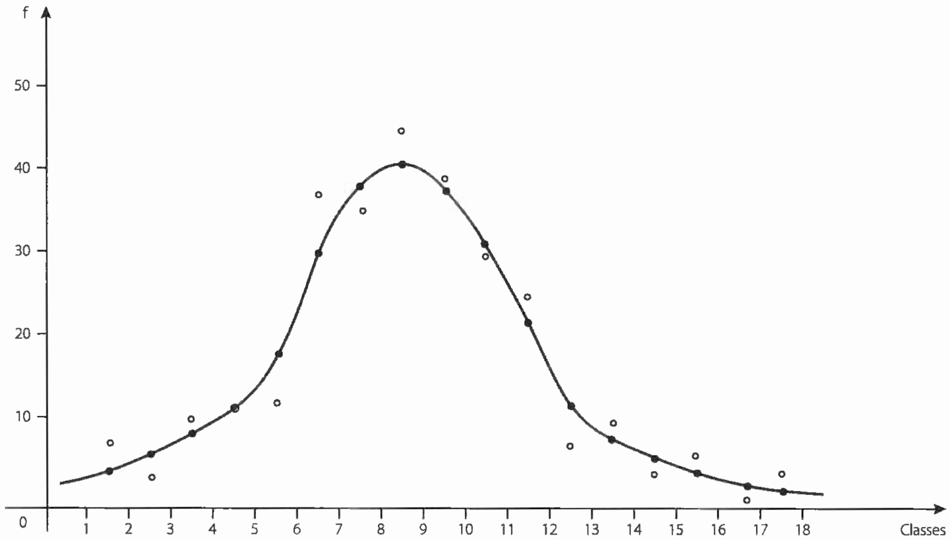
8.



9. a/b



c.



CAPÍTULO 6 — MEDIDAS DE POSIÇÃO

EXERCÍCIOS (p. 100)

1. **a.** $\bar{x} = 5{,}1$; Md = 5; Mo = 5
b. $\bar{x} = 11$; Md = 9; Mo = 7
c. $\bar{x} = 49{,}8$; Md = 49,5; $\nexists$ Mo
d. $\bar{x} = 15{,}1$; Md = 15; $\nexists$ Mo

2. **a.** R$ 96 **b.** R$ 88

3. **a.** 7,9 **b.** 7,8 **c.** 7,2

4. **a.** 5,4 **b.** 5 **c.** 5

5. **a.** 5,9 **b.** 6 **c.** 6

6. **a.** 64,5 **b.** 58,8

7. –2,5; –0,5; –3,5; 3,5; 2,5; –1,5; –4,5; 6,5

8. **a.** 5,3 **b.** 172,4 cm **c.** R$ 843 **d.** 159,4 kg

9. **a.** 5,3 **b.** 174 cm **c.** R$ 810 **d.** 157,8 kg

10. **a.** 5 **b.** 178 cm **c.** R$ 800 **d.** 148 kg

11. **a.** 3,5 e 7,2 **c.** R$ 694 e R$ 947
b. 166,2 cm e 179,2 cm **d.** 145 kg e 166 kg

12. $P_{10} = 159{,}3$ cm $P_1 = 151{,}1$ cm $P_{23} = 165{,}4$ cm
$P_{15} = 161{,}7$ cm $P_{90} = 183$ cm

13. c

14. c

CAPÍTULO 7 — MEDIDAS DE DISPERSÃO OU DE VARIABILIDADE

EXERCÍCIOS (p. 115)

1. **a.** 8 **b.** 8 **c.** 9,2 **d.** 20

2. **a.** 6 **b.** 0,7

3. **a.** 2,96 **b.** 2,81 **c.** 3,016 **d.** 7,04

4. **a.** 1,51 **b.** 0,159

5. 1,13

6. 4,45

7. **a.** 2,43 **b.** 8,8 cm **c.** R$ 229 **d.** 9,93 kg

8. 8,03%

9. Estatística

10. estatura

11. 3,72% e 3,71%, respectivamente; o segundo grupo
12. 5,41
13. 51,7

CAPÍTULO 8 — MEDIDAS DE ASSIMETRIA MEDIDAS DE CURTOSE

EXERCÍCIOS (p. 119)

1. simétrica; assimétrica negativa; assimétrica positiva
2. 0,283
3. a. assimétrica positiva b. 0,364
4. 0,021

EXERCÍCIOS (p. 121)

1. a. 0,252; 0,263; 0,287
   b. leptocúrtica; mesocúrtica; platicúrtica
2. 0,258 < 0,263 ⇒ leptocúrtica

CAPÍTULO 9 — PROBABILIDADE

EXERCÍCIOS (p. 130)

1. a. $\frac{1}{2}$ b. $\frac{3}{13}$ c. $\frac{1}{4}$ d. $\frac{3}{8}$
2. a. $\frac{1}{5}$ b. $\frac{1}{10}$ c. $\frac{1}{25}$ d. $\frac{2}{25}$
3. a. $\frac{1}{12}$ b. $\frac{1}{9}$ c. $\frac{5}{12}$ d. $\frac{5}{18}$
4. a. $\frac{1}{4}$ b. $\frac{1}{2}$
5. a. $\frac{3}{7}$ b. $\frac{1}{7}$
6. $\frac{4}{13}$
7. $\frac{1}{6}$
8. a. $\frac{1}{11}$ b. $\frac{14}{33}$ c. $\frac{19}{33}$
9. $\frac{2}{3}$
10. a. $\frac{1}{221}$ b. $\frac{4}{663}$
11. a. $\frac{1}{8}$ b. $\frac{3}{8}$
12. a. $\frac{1}{8}$ b. $\frac{3}{8}$ c. $\frac{3}{8}$
    d. $\frac{1}{8}$ e. $\frac{7}{8}$ f. $\frac{1}{2}$
13. a. $\frac{1}{6}$ b. $\frac{1}{6}$ c. $\frac{25}{36}$ d. $\frac{11}{36}$
14. a. $\frac{1}{50}$ b. $\frac{1}{2}$ c. $\frac{3}{5}$ d. $\frac{2}{5}$
15. a. $\frac{1}{3}$ b. $\frac{1}{11}$ c. $\frac{19}{33}$
16. a. $\frac{1}{18}$ b. $\frac{1}{12}$
17. a. $\frac{1}{9}$ b. $\frac{5}{18}$ c. $\frac{1}{9}$
18. a. $\frac{7}{8}$ b. $\frac{5}{8}$ c. $\frac{3}{4}$
19. a. $\frac{3}{8}$ b. $\frac{7}{8}$ c. $\frac{91}{120}$ d. $\frac{1}{8}$

CAPÍTULO 10 — DISTRIBUIÇÕES BINOMIAL E NORMAL

EXERCÍCIOS (p. 138)

1. $\frac{5}{16}$
2. $\frac{2}{9}$
3. a. $\frac{400}{729}$ b. $\frac{665}{729}$
4. $\frac{40}{243}$
5. 9,8415%

EXERCÍCIOS (p. 143)

1. a. 0,4251 e. 0,9788
   b. 0,3023 f. 0,1401
   c. 0,9104 g. 0,2546
   d. 0,2064 h. 0,7258
2. a. 0,0228 c. 0,8664
   b. 0,9772 d. 0,5
3. a. 0,6338 c. 0,6879
   b. 0,6480
4. a. 0,9998 c. 0,0062
   b. 0,8944

CAPÍTULO 11 — CORRELAÇÃO E REGRESSÃO

EXERCÍCIOS (p. 154)

1. 0,98
2. b. 0,89
3. a. 0,99 c. 47,5
   b. $\hat{Y} = 1{,}5X + 40$
4. a. 0,98 c. 1.007,5 mm
   b. $\hat{Y} = 0{,}56X - 2{,}6$ d. 1.017 mm
5. a. 0,94 c. R$ 12,66
   b. $\hat{Y} = 0{,}34X + 9{,}94$
6. a. –0,99 c. $\hat{Y} = 76{,}6$
   b. $\hat{Y} = -11{,}4X + 76{,}6$
7. a. –0,90 c. 274,6 e 162,4
   b. $\hat{Y} = -1{,}87X + 386{,}8$
8. a. 0,54 c. $\hat{X} = 0{,}16Y + 0{,}40$
   b. $\hat{Y} = 1{,}81X + 0{,}01$

CAPÍTULO 12 — NÚMEROS-ÍNDICES

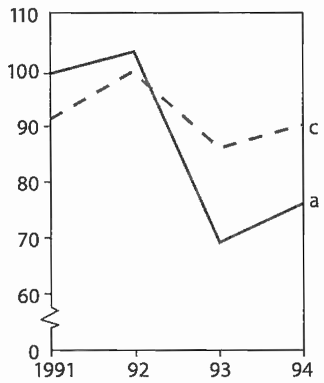
EXERCÍCIO (p. 162)

a. 100,0; 103,3; 69,2; 76,2
b. —; 109,2; 96,0; 97,3

c. 91,4; 100,0; 86,2; 84,9

d. $1{,}104 \times 1{,}170 \times 1{,}218 = 1{,}573$ e $q_{91,94} = 1{,}573$

e.



EXERCÍCIOS (p. 169)

1. 49,3; 74,9; 100,0; 158,1; 204,4; 285,7
2. R$ 629
3. R$ 541.491; R$ 557.227; R$ 612.500; R$ 851.255; R$ 826.058; R$ 1.044.210; R$ 1.118.145
4. 2,57
5. 20,75%

APÊNDICE — INSTRUMENTAL MATEMÁTICO

RESOLVA (p. 173)

c. 0,4	g. 6,8
d. 4,2	h. 5,6
e. 328,4	i. 90,0
f. 3,0	

EXERCÍCIOS (p. 174)

1. a. 23,4; 48,9; 120,4; 234,8; 78,9; 130,0; 45,1; 12,4; 200,0
   b. 46,73; 253,65; 28,26; 123,84; 299,95; 37,49
   c. 27,68; 129; 50; 68; 39
   d. 40; 270; 300; 60; 270; 300; 450; 270; 3.000
2. "descarregar" em 0,31
3. "descarregar" em 27

EXERCÍCIOS (p. 180)

1. $\frac{1}{7}$
2. $\frac{1}{6}$
3. $\frac{1}{4}$
4. 15
5. $\frac{56}{7}, \frac{120}{10}, \frac{91}{13}$
6. a. $\frac{2}{3}$ b. $\frac{1}{2}$ c. 11
   d. $\frac{17}{3}$ e. $\frac{2}{3}$ f. $\frac{3}{5}$
7. a. $\frac{16}{40}, \frac{25}{40}, \frac{60}{40}$
   b. $\frac{16}{36}, \frac{10}{36}, \frac{7}{36}$
   c. $\frac{16}{18}, \frac{30}{18}, \frac{21}{18}$
8. a. $\frac{1}{8} < \frac{3}{8} < \frac{5}{8}$ b. $\frac{1}{3} < \frac{3}{5} < \frac{4}{6}$
9. a. $\frac{31}{20}$ b. $\frac{29}{24}$ c. $\frac{3}{7}$ d. $\frac{3}{4}$
   e. 2 f. $\frac{1}{9}$ g. $\frac{21}{5}$ h. $\frac{3}{2}$
   i. $\frac{1}{8}$ j. $\frac{5}{3}$ l. 9 m. $\frac{2}{25}$
   n. 2 o. $\frac{9}{25}$ p. $\frac{64}{125}$ q. $\frac{125}{27}$
   r. 1 s. $\frac{7}{9}$

EXERCÍCIOS (p. 181)

1. a. 0,3 b. 1,28 c. 0,050 d. 4,5
2. a. $\frac{7}{10}$ b. $\frac{12}{100}$ c. $\frac{1.275}{100}$ d. $\frac{18}{1.000}$
3. a. 64,4793 b. 124,668 c. 8,903
   d. 2,96 e. 0,045 f. 1,539
   g. 82,5 h. 6,2 i. 380
   j. 0,012 l. 3,3708 m. 0,099225
   n. 70,9 o. 1.000 p. 0,009
   q. 5.000 r. 3,2 s. 11,3
   t. 0,6 u. 0,09

EXERCÍCIOS (p. 184)

1. a. 40% b. 75% c. 6% d. 5% e. 250%
2. a. $\frac{3}{10}$ b. $\frac{2}{5}$ c. $\frac{3}{5}$ d. 2 e. $\frac{1}{40}$
3. a. 60 b. 22,5 c. 56 d. 4,5
4. 75
5. 325.000 sacas
6. 5%
7. 1.500 meninas e 1.000 meninos
8. 1.050 e 126
9. R$ 18
10. R$ 880
11. R$ 180
12. R$ 83 e R$ 97

EXERCÍCIOS (p. 187)

1. a. $x_1 + x_2 + ... + x_8$
   b. $x_3 + x_4 + x_5 + x_6$
   c. $x_1 + x_2 + ... + x_5$
2. a. $\sum_{i=1}^{4} x_i$ b. $\sum_{i=1}^{7} x_i$ c. $\sum_{i=4}^{7} x_i$ d. $\sum_{i=5}^{10} x_i$
3. $x_1 = 2; x_2 = 5; x_3 = 7; x_4 = 10; x_5 = 12; x_6 = 13; x_7 = 15$
4. a. 64 b. 24 c. 57 d. 35

EXERCÍCIOS (p. 189)

1. 0,502
2. 8,17
3. 7,94
4. 8,25

EXERCÍCIO (p. 191)

a. 8 b. 1 c. 3 d. $n^2 - 5n + 6$

EXERCÍCIO (p. 193)

a. 56 b. 4.950 c. 1

EXERCÍCIOS (p. 195)

1. a. $81y^4 + 108y^3 + 54y^2 + 12y + 1$

   b. $\frac{1}{64}y^6 + \frac{3}{8}y^5 + \frac{15}{4}y^4 + 20y^3 + 60y^2 + 96y + 64$

   c. $32x^5 + 80x^4 + 80x^3 + 40x^2 + 10x + 1$

2. a. $210p^6q^4$ c. $41.184x^8$

   b. $1.920b^3$ d. $5.670x^4$

EXERCÍCIOS (p. 200)

2. $y = \sqrt{x}$
4. a. $y = x$ c. $2x + y = 0$

   b. $3x - 5y - 15 = 0$ d. $y = 2x - 1$

COLETÂNEA DE QUESTÕES OBJETIVAS

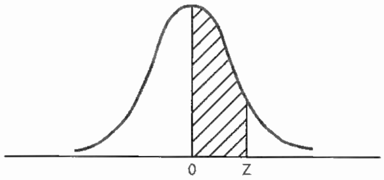
1. c	7. d	13. d	19. c	25. b	31. a	37. c	43. b
2. c	8. b	14. b	20. a	26. b	32. e	38. a	44. c
3. c	9. b	15. e	21. b	27. e	33. d	39. b	
4. b	10. a	16. a	22. b	28. e	34. e	40. d	
5. a	11. b	17. d	23. b	29. a	35. b	41. e	
6. b	12. b	18. c	24. d	30. a	36. c	42. a	

ANEXO I
TABELA DE NÚMEROS ALEATÓRIOS

5 7 7 2 0 0 3 9 8 4 8 4 4 1 7 9 6 7 7 1 4 0 2 1 1 3 9 7 5 6 4 9 8 6 5 4 0 8 9 3 2 9 6 8 7 4 5 4 8 3
2 8 8 0 5 3 5 1 5 9 0 9 9 3 9 8 8 7 5 8 7 0 2 7 7 1 7 7 1 7 0 6 3 2 0 2 7 8 6 2 1 6 7 4 6 9 6 5 1 7
9 2 5 9 1 8 5 2 8 7 3 0 4 8 8 6 9 7 4 8 3 5 2 5 1 8 8 8 7 4 0 3 6 2 9 8 3 8 5 8 6 5 8 6 4 2 4 1 0 3
9 0 3 8 1 2 9 1 7 4 3 0 1 9 7 5 8 9 0 7 5 0 6 4 1 5 5 9 7 1 8 8 1 3 7 4 9 5 3 0 5 2 7 8 3 0 1 1 7 5
8 0 9 1 1 6 9 4 6 7 5 8 6 0 8 2 0 6 6 6 9 0 4 7 5 6 1 8 4 6 4 5 1 1 1 2 3 5 3 2 4 5 5 0 4 1 1 3 4 3
2 2 0 1 7 0 3 1 3 2 9 6 9 1 9 2 7 5 4 0 1 6 5 4 2 9 7 2 7 4 9 9 0 0 9 5 9 7 6 1 0 0 9 8 2 4 3 0 0 7
5 6 2 4 1 0 0 4 3 0 2 0 4 6 2 9 9 0 5 3 5 3 1 1 0 5 8 4 4 1 2 1 6 4 7 9 1 9 7 6 2 9 5 1 6 2 6 0 6 6
7 9 4 4 9 2 6 2 0 2 9 6 8 6 6 4 3 0 0 0 9 4 5 6 6 9 3 0 2 0 5 9 8 7 8 7 3 5 4 4 2 2 5 0 9 7 7 8 1 9
5 3 9 9 6 6 4 5 0 8 8 9 7 8 5 0 7 7 5 3 3 7 2 5 7 7 4 1 2 7 6 2 3 8 0 2 2 3 5 7 6 2 0 1 4 1 6 0 3 5
1 8 9 2 8 7 3 5 8 8 5 5 0 5 2 1 3 6 5 1 3 9 2 8 5 0 1 4 6 6 8 5 7 9 3 0 1 9 7 9 7 2 6 6 6 4 3 1 4 5
5 3 0 8 5 8 9 6 6 3 0 5 6 1 2 5 7 0 2 2 5 0 4 1 2 8 9 6 6 2 6 6 4 3 6 3 0 6 6 3 0 1 3 2 7 9 8 5 2 2
0 3 5 8 8 0 2 9 2 8 7 6 8 9 5 1 1 8 2 4 8 8 8 9 4 6 4 7 4 8 5 9 1 9 2 9 8 7 0 3 1 0 3 3 9 9 6 7 1 2
2 7 0 7 8 1 8 8 6 5 6 9 4 9 9 8 0 0 2 8 0 4 7 0 5 1 3 0 0 1 4 7 1 8 9 7 3 3 2 1 8 5 8 2 4 5 4 3 2 4
0 5 2 1 0 8 5 9 0 1 0 6 2 2 2 4 9 8 9 1 8 1 1 7 5 5 4 4 6 6 1 6 0 7 7 3 0 7 6 6 1 0 1 2 3 1 7 8 5 8
4 0 3 6 1 3 2 7 8 4 3 0 8 2 3 3 3 6 3 9 6 9 4 2 0 5 5 8 6 4 6 1 1 2 3 3 8 9 2 7 8 9 5 2 6 6 7 1 9 3
5 4 6 0 2 5 2 8 8 5 8 8 2 0 0 0 1 0 5 9 6 1 0 5 3 6 6 1 3 3 7 2 0 1 0 1 1 9 0 1 6 1 1 0 5 1 2 0 9 1
7 1 5 1 6 3 4 0 7 6 7 1 1 1 7 3 7 3 5 2 3 7 3 1 6 0 4 5 8 8 9 2 7 3 4 3 7 1 2 8 0 4 9 8 0 9 0 2 4 8
6 1 0 2 0 1 8 1 7 3 9 2 6 0 6 6 7 3 5 8 5 3 3 4 4 2 6 8 2 6 3 8 3 4 0 3 2 7 4 4 9 6 0 4 4 6 6 5 9 3
8 2 5 5 9 3 1 3 4 6 3 0 9 5 2 6 5 5 0 6 9 6 1 7 6 5 9 1 7 2 3 9 7 9 9 6 1 2 4 9 5 2 8 0 6 3 2 6 9 9
8 9 9 8 5 4 1 4 2 1 7 4 1 3 5 7 6 8 1 9 8 6 2 8 6 0 8 9 4 7 3 3 1 5 2 6 2 8 7 7 4 5 3 8 4 8 0 8 0 8
0 0 9 9 8 4 8 4 1 4 6 7 9 5 1 3 7 7 5 8 9 0 1 4 5 0 7 9 4 2 7 3 6 3 3 1 0 6 6 0 4 3 4 0 1 2 5 5 0 4
6 2 4 1 5 0 7 8 2 0 4 8 0 5 8 8 4 3 5 2 9 8 0 3 1 9 9 3 9 2 0 3 0 4 9 7 2 5 8 4 9 5 9 5 0 3 6 3 3 1
9 4 2 7 9 0 6 9 2 4 6 8 0 9 9 2 1 1 8 6 0 7 6 3 8 3 1 9 3 2 9 9 5 1 1 5 5 5 7 1 0 9 2 7 0 2 6 7 0 0
4 4 8 9 2 9 2 8 8 4 3 6 2 8 2 5 1 5 8 2 8 7 7 4 1 8 9 7 2 5 7 6 1 0 6 3 2 6 7 6 0 2 2 6 7 4 5 3 2 8
9 7 3 0 7 6 9 5 3 3 2 1 1 0 5 4 2 6 9 5 6 6 6 5 5 2 0 4 9 9 3 6 5 8 4 8 0 3 0 8 9 3 6 3 5 8 1 7 9 6
3 9 1 6 5 8 0 4 4 4 8 0 1 5 5 9 5 9 8 3 9 0 9 5 5 4 6 6 8 1 8 4 3 9 6 0 8 5 3 8 8 8 6 6 3 3 3 5 6 9
6 0 7 8 1 1 0 3 2 6 6 7 5 0 3 4 0 9 6 1 3 1 3 0 2 0 7 6 9 3 6 6 3 0 8 3 5 1 0 9 3 3 8 3 6 4 7 6 0 5
0 3 1 9 2 3 4 7 6 2 8 9 5 7 7 7 9 1 3 3 8 8 4 7 6 0 5 9 3 7 5 4 3 9 4 8 7 7 6 7 4 9 8 5 3 8 4 3 9 1
4 1 2 8 5 2 6 7 5 6 2 5 3 9 5 9 9 6 6 5 5 1 3 6 9 0 3 2 2 2 3 9 3 3 0 5 2 2 9 9 0 3 3 9 9 7 9 6 9 9
7 7 5 4 9 8 5 0 3 9 2 5 3 7 4 2 5 2 9 7 1 0 0 3 5 6 0 4 9 2 8 1 6 6 8 6 7 0 0 1 4 8 8 9 5 5 8 2 1 0
2 8 6 3 4 1 6 1 9 1 6 4 2 4 8 3 8 1 3 7 3 4 4 8 8 3 2 7 9 6 3 8 7 1 6 9 7 3 0 6 7 7 5 0 2 5 6 4 6 0
7 4 2 4 4 8 8 5 4 0 1 2 3 3 5 9 6 7 5 0 1 4 9 8 1 4 2 6 4 2 7 9 7 9 1 3 5 2 8 9 6 9 7 8 8 0 4 4 7 1
0 0 2 4 0 3 3 7 9 6 4 6 6 8 7 5 0 5 3 2 4 2 1 6 6 3 3 3 2 8 9 7 2 6 3 6 4 7 2 7 7 3 6 5 3 8 3 4 4 6
0 5 4 1 4 7 6 9 6 9 4 5 3 6 1 6 7 1 1 8 9 5 5 1 9 7 2 2 0 4 1 3 2 3 9 6 5 8 6 0 0 3 6 9 4 8 7 9 8 3
6 2 6 9 8 4 9 7 9 7 4 7 2 3 6 6 5 1 5 6 1 3 0 8 6 9 1 1 5 2 7 5 5 9 2 6 8 6 8 1 8 0 4 3 0 0 9 8 9 2

NOTA: 0 — 10   00 — 100   000 — 1.000   etc.

ANEXO II
ÁREA SUBTENDIDA PELA CURVA NORMAL REDUZIDA DE 0 A Z



z	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0754
0,2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0,3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0,4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0,5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0,6	2258	2291	2324	2357	2389	2422	2454	2486	2518	2549
0,7	2580	2612	2642	2673	2704	2734	2764	2794	2823	2852
0,8	2881	2910	2939	2967	2996	3023	3051	3078	3106	3133
0,9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1,0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621
1,1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1,2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1,3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1,4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1,5	4332	4345	4357	4370	4382	4394	4406	4418	4429	4441
1,6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545
1,7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1,8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706
1,9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
2,0	4772	4778	4783	4788	4793	4798	4803	4808	4812	4817
2,1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857
2,2	4861	4864	4868	4871	4875	4878	4881	4884	4887	4890
2,3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916
2,4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936
2,5	4938	4940	4941	4943	4945	4946	4948	4949	4951	4952
2,6	4953	4955	4956	4957	4959	4960	4961	4962	4963	4964
2,7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974
2,8	4974	4975	4976	4977	4977	4978	4979	4979	4980	4981
2,9	4981	4982	4982	4983	4984	4984	4985	4985	4986	4986
3,0	4987	4987	4987	4988	4988	4989	4989	4989	4990	4990
3,1	4990	4991	4991	4991	4992	4992	4992	4992	4993	4993
3,2	4993	4993	4994	4994	4994	4994	4994	4995	4995	4995
3,3	4995	4995	4995	4996	4996	4996	4996	4996	4996	4997
3,4	4997	4997	4997	4997	4997	4997	4997	4997	4997	4998
3,5	4998	4998	4998	4998	4998	4998	4998	4998	4998	4998
3,6	4998	4998	4999	4999	4999	4999	4999	4999	4999	4999
3,7	4999	4999	4999	4999	4999	4999	4999	4999	4999	4999
3,8	4999	4999	4999	4999	4999	4999	4999	4999	4999	4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

LIVRO COM MATERIAL DE APOIO

O material de apoio consiste em conteúdo extra, que a **Editora Saraiva** disponibiliza no site www.editorasaraiva.com.br.

Voltado para professores e estudantes, ele é composto por itens que complementam o conteúdo do livro e servem de apoio didático para uso em sala de aula e para fixação do tema. Cada livro tem seu material e seu público específico, sendo necessária a visita ao site para ver qual material foi disponibilizado para você.

Além do material de apoio, o site traz os principais lançamentos e as novidades na área de livros Universitários e de Negócios, incluindo os livros digitais, contribuindo cada vez mais para a expansão do conhecimento.

Essa é mais uma forma de demonstrar a preocupação da **Editora Saraiva** em criar subsídios que facilitem a aprendizagem e que possam colaborar com a evolução da educação no Brasil.

Este livro é voltado para todos os estudantes de cursos técnicos de Contabilidade bem como para os alunos de cursos superiores e profissionais diversos que necessitam de uma abordagem introdutória do assunto.

Com uma linguagem extremamente objetiva, a principal preocupação de *Estatística fácil* foi a de apresentar de maneira clara todos os tópicos exigidos pelos programas dos cursos profissionalizantes, de modo a facilitar o aprendizado por parte do aluno.

Com características estritamente didáticas, evitando demonstrações e apresentando análises práticas e objetivas, o livro é instrumento essencial para todos aqueles que necessitam de uma abordagem introdutória — mas não superficial — da Estatística como um todo.

Aplicação

Este livro pode ser utilizado na disciplina de Estatística.

ISBN 978-85-02-08106-2