

# Exercícios Complementares

## Capítulo 2 Exercícios

### Exercícios para a Seção 2.1

Forneça uma descrição razoável do espaço amostral para cada um dos experimentos aleatórios nos Exercícios 2.1.1 a 2.1.11. Poderá haver mais de uma interpretação aceitável de cada experimento. Descreva qualquer suposição que você faça.

**2.1.1** Cada um dos quatro *bits* transmitidos é classificado como com erro e sem erro.

**2.1.2** O número de visitas a *sites* da internet com alto volume de visitas é gravado em um dia.

**2.1.3** Na inspeção final de suprimentos eletrônicos de potência, as unidades podem ser aprovadas ou três tipos de não conformidades podem ocorrer: funcional, menor ou cosmética. Três unidades são inspecionadas.

**2.1.4** Um amperímetro, que mostra três dígitos, é usado para medir corrente em miliampères.

**2.1.5** As duas questões seguintes aparecem em questionários na inspeção de empregados. Cada resposta é escolhida a partir de uma escala de cinco pontos: 1 (nunca), 2, 3, 4 e 5 (sempre).

A corporação está disposta a ouvir e avaliar de maneira justa novas ideias?

Quão frequentes meus colaboradores são importantes no desempenho global do trabalho?

**2.1.6** O tempo até que uma transação de serviço seja solicitada de um computador, com a precisão de milissegundo.

**2.1.7** A leitura de pH de uma amostra de água com a precisão de décimo de uma unidade.

**2.1.8** Os vazios em uma placa de óxido de ferro são classificados como pequenos, médios ou grandes. O número de vazios em cada categoria é medido por uma inspeção óptica de uma amostra.

**2.1.9** Uma peça amostrada, moldada por injeção, poderia ter sido produzida em uma das duas prensas e em qualquer uma das oito cavidades em cada prensa.

**2.1.10** Um pedido de compra de um automóvel pode especificar uma transmissão automática ou padrão, com ou sem ar-condicionado, e qualquer uma das quatro cores: vermelha, azul, preta ou branca. Descreva o conjunto de pedidos possíveis para esse experimento.

**2.1.11** Ligações são repetidamente feitas em uma linha telefônica ocupada até que uma conexão seja alcançada.

**2.1.12** Em um aparelho magnético de armazenagem, três tentativas são feitas para ler dados antes que seja usado um procedimento de recuperação de erro que reposiciona a cabeça magnética. O procedimento de recuperação de erros tenta três reposições antes que uma mensagem de “abortar” seja mandada para o operador. Seja

$s$  o sucesso de uma operação de leitura

$f$  a falha de uma operação de leitura

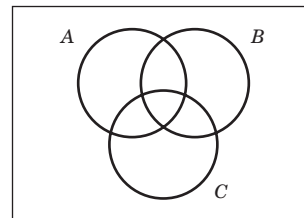
$S$  o sucesso de um procedimento de recuperação de erro

$F$  a falha de um procedimento de recuperação de erro

$A$  uma mensagem de abortar, enviada para o operador

Utilizando o diagrama em forma de árvore, descreva o espaço amostral desse experimento.

**2.1.13** Três eventos são mostrados no diagrama de Venn na seguinte figura:



Reproduza a figura e sombreie a região que corresponde a cada um dos seguintes eventos:

- a.  $A'$       b.  $A \cap B$       c.  $(A \cap B) \cup C$   
d.  $(B \cup C)'$       e.  $(A \cap B)' \cup C$

**2.1.14** Em uma operação de moldagem por injeção, o comprimento e a largura, denotados por  $X$  e  $Y$ , respectivamente, de cada peça moldada são avaliados. Seja

$A$  o evento em que  $48 < X < 52$  centímetros

$B$  o evento em que  $9 < Y < 11$  centímetros

Construa um diagrama de Venn que inclua esses eventos. Sombreie as áreas que representam o seguinte:

- a.  $A$       b.  $A \cap B$   
c.  $A' \cup B$       d.  $A \cap B$

e. Se esses eventos forem mutuamente excludentes, quão bem-sucedida seria essa operação de produção? O processo produziria peças com  $X = 50$  centímetros e  $Y = 10$  centímetros?

**2.1.15** Uma balança digital é usada para fornecer pesos em gramas.

a. Qual é o espaço amostral para esse experimento?

Seja  $A$  o evento em que um peso excede 11 gramas; seja  $B$  o evento em que um peso é menor que ou igual a 15 gramas; e seja  $C$  o evento em que um peso é maior que ou igual a 8 gramas e menor que 12 gramas.

Descreva os seguintes eventos:

- b.  $A \cup B$       c.  $A \cap B$   
d.  $A'$       e.  $A \cup B \cup C$   
f.  $(A \cup C)'$       g.  $A \cap B \cap C$   
h.  $B' \cap C$       i.  $A \cup (B \cap C)$

**2.1.16** Em fotossíntese dependente de luz, a qualidade da luz se refere aos comprimentos de onda de luz que são importantes. O comprimento de onda de uma amostra de radiações fotossinteticamente ativas (RFA) é medido em nanômetro. A faixa do vermelho é 675-700 nm e a faixa

## 2 Exercícios Complementares

do azul é 450-500 nm. Seja  $A$  o evento em que RFA ocorre na faixa do vermelho e seja  $B$  o evento em que RFA ocorre na faixa do azul. Descreva o espaço amostral e indique cada um dos seguintes eventos:

- a.  $A$    b.  $B$    c.  $A \cap B$    d.  $A \cup B$

2.1.17 Quatro *bits* são transmitidos em um canal digital de comunicações. Cada *bit* é distorcido ou recebido sem distorção. Seja  $A_i$  o evento em que o  $i$ -ésimo *bit* é distorcido,  $i = 1, \dots, 4$ .

- a. Descreva o espaço amostral para esse experimento.  
b. São os eventos  $A_i$ 's mutuamente excludentes?

Descreva os resultados em cada um dos seguintes eventos:

- c.  $A_1$    d.  $A_1'$   
e.  $A_1 \cap A_2 \cap A_3 \cap A_4$    f.  $(A_1 \cap A_2) \cup (A_3 \cap A_4)$

2.1.18 Discos de plástico de policarbonato, provenientes de um fornecedor, são analisados com relação à resistência a arranhões e a choque. Os resultados de 100 discos estão resumidos a seguir:

		Resistência a Choque	
		Alta	Baixa
Resistência a arranhões	Alta	70	9
	Baixa	16	5

Seja  $A$  o evento em que um disco tem alta resistência a choque e seja  $B$  o evento em que um disco tem alta resistência a arranhões. Indique o número de discos em  $A \cap B$ ,  $A'$  e  $A \cup B$ .

2.1.19 Em replicação controlada, células são replicadas em um período de dois dias. Dois mecanismos de controle foram identificados — um positivo e um negativo. Suponha que uma replicação seja observada em três células. Seja  $A$  o evento em que todas as células são identificadas como positivas e seja  $B$  o evento em que todas as células são negativas. Descreva o espaço amostral graficamente e mostre cada um dos seguintes eventos:

- a.  $A$    b.  $B$    c.  $A \cap B$    d.  $A \cup B$

2.1.20 Amostras de emissões de três fornecedores são classificadas com relação a satisfazer as especificações de qualidade do ar. Os resultados de 100 amostras são resumidos a seguir:

		Satisfaz	
		Sim	Não
Fornecedor	1	22	8
	2	25	5
	3	30	10

Seja  $A$  o evento em que uma amostra seja proveniente do fornecedor 1 e seja  $B$  o evento em que uma amostra atenda às especificações. Determine o número de amostras em  $A' \cap B$ ,  $B'$  e  $A \cup B$ .

2.1.21 O tempo de enchimento de um reator é medido em minutos (e frações de minutos). Seja o espaço amostral formado por números reais e positivos. Defina os eventos  $A$  e  $B$  como segue:  $A = \{x \mid x < 72,5\}$  e  $B = \{x \mid x > 52,5\}$ .

Descreva cada um dos seguintes eventos:

- a.  $A'$    b.  $B'$   
c.  $A \cap B$    d.  $A \cup B$

2.1.22 A tabela seguinte resume 204 reações endotérmicas envolvendo bicarbonato de sódio.

Condições Finais de Temperatura	Calor Absorvido (cal)	
	Abaixo do Valor Desejado	Acima do Valor Desejado
266 K	12	40
271 K	44	16
274 K	56	36

Seja  $A$  o evento em que a temperatura final de uma reação é 271 K ou menor. Seja  $B$  o evento em que o calor absorvido está abaixo do valor desejado. Determine o número de reações em cada um dos seguintes eventos:

- a.  $A \cap B$    b.  $A'$    c.  $A \cup B$   
d.  $A \cup B'$    e.  $A' \cap B'$

2.1.23 Um anúncio na internet pode ser projetado a partir de quatro cores diferentes, de três tipos de fonte, de cinco tamanhos de fonte, de três imagens e de cinco frases. Quantos projetos diferentes são possíveis?

2.1.24 Considere os dados da sala de emergência do hospital do Exemplo 2.6. Seja  $A$  o evento em que uma visita seja para o Hospital 1 e seja  $B$  o evento em que uma visita resulte na internação do paciente em qualquer hospital. Determine o número de pessoas em cada um dos seguintes eventos.

- a.  $A \cap B$    b.  $A'$    c.  $A \cup B$   
d.  $A \cup B'$    e.  $A' \cap B'$

2.1.25 O artigo "Term Efficacy of Ribavirin Plus Interferon Alfa in the Treatment of Chronic Hepatitis C" [*Gastroenterology* (1996, Vol. 111(5), pp. 1307-1312)] considerou o efeito de dois tratamentos e um controle para tratamento de hepatite C. A seguinte tabela fornece os pacientes totais em cada grupo e o número que mostrou uma resposta completa (positiva) depois de 24 semanas de tratamento.

	Resposta	
	Completa	Total
Ribavirina mais alfa interferon	16	21
Alfa interferon	6	19
Controles não tratados	0	20

Seja  $A$  o evento em que o paciente foi tratado com ribavirina mais alfa interferon e seja  $B$  o evento em que a resposta foi completa. Determine o número de pacientes em cada um dos seguintes eventos.

- a.  $A$    b.  $A \cap B$    c.  $A \cup B$    d.  $A' \cap B'$

2.1.26 Um sistema de computadores usa senhas que contêm exatamente oito caracteres, e cada caractere é uma das 26 letras minúsculas ( $a-z$ ) ou 26 letras maiúsculas ( $A-Z$ ) ou 10 inteiros (0-9). Seja  $\Omega$  o conjunto de todas as senhas possíveis e sejam  $A$  e  $B$  os eventos que consistem em senhas com somente letras ou com somente inteiros, respectivamente. Determine o número de senhas em cada uma das seguintes opções:

- a.  $\Omega$    b.  $A$    c.  $A' \cap B'$   
d. Senhas que conttenham no mínimo 1 inteiro.  
e. Senhas que conttenham exatamente 1 inteiro.

## Exercícios para a Seção 2.2

2.2.1 Uma amostra de duas placas de circuito impresso é selecionada sem reposição a partir de uma batelada. Descreva o espaço amostral (ordenado) para cada uma das seguintes bateladas:

- a. A batelada contém 90 placas que são não defeituosas, oito placas com pequenos defeitos e duas placas com grandes defeitos.  
b. A batelada contém 90 placas que são não defeituosas, oito placas com pequenos defeitos e uma placa com grandes defeitos.

2.2.2. Uma amostra de dois itens é selecionada sem reposição a partir de uma batelada. Descreva o espaço amostral (ordenado) para cada uma das seguintes bateladas:

- A batelada contém os itens  $\{a, b, c, d\}$ .
- A batelada contém os itens  $\{a, b, c, d, e, f, g\}$ .
- A batelada contém quatro itens defeituosos e 20 itens bons.
- A batelada contém um item defeituoso e 20 itens bons.

2.2.3 O controle remoto de abertura de um portão de garagem tem um código determinado pelo movimento, para cima ou para baixo, de 12 interruptores. Quantos resultados estão no espaço amostral de códigos possíveis?

2.2.4 Em uma operação de fabricação, um item é produzido por usinagem, polimento e pintura. Se há três ferramentas de usinagem, quatro ferramentas de polimento e três ferramentas de pintura, quantas rotinas diferentes (consistindo em usinagem, seguida de polimento e depois pintura) são possíveis para o item?

2.2.5 Novos projetos para um tanque de tratamento de águas residuais têm proposto três formas possíveis, quatro tamanhos diferentes, três localizações diferentes para válvulas de alimentação de material e quatro localizações diferentes para as válvulas de retirada de material. Quantos projetos diferentes do produto são possíveis?

2.2.6 Um processo de fabricação consiste em dez operações que podem ser completadas em qualquer ordem. Quantas sequências diferentes de produção são possíveis?

2.2.7 Um lote de 140 *chips* semicondutores é inspecionado, escolhendo-se uma amostra de cinco *chips*, sem reposição. Suponha que dez dos *chips* não obedecem aos requerimentos dos consumidores.

- Quantas amostras diferentes são possíveis?
- Quantas amostras contêm exatamente um *chip* não conforme?
- Quantas amostras contêm no mínimo um *chip* não conforme?

2.2.8 Em uma operação de chapa metálica, três entalhes e quatro dobramentos são requeridos. Se as operações podem ser feitas em qualquer ordem, quantas maneiras diferentes são possíveis para completar a fabricação?

2.2.9 No laboratório de análises de amostras de um processo químico, cinco amostras do processo são analisadas diariamente. Além disso, uma amostra controle é analisada duas vezes por dia para verificar a calibração dos instrumentos do laboratório.

- Quantas sequências diferentes de amostras de processo e de controle são possíveis por dia? Suponha que as cinco amostras de processo sejam consideradas idênticas e que as duas amostras de controle sejam consideradas idênticas.
- Quantas sequências diferentes de amostras de processo e de controle são possíveis se considerarmos que as cinco amostras de processo sejam diferentes e que as duas amostras de controle sejam consideradas idênticas?
- Para a mesma situação do item (b), quantas sequências são possíveis se o primeiro teste de cada dia tiver de ser uma amostra de controle?

2.2.10 Na disposição de uma placa de circuito impresso para um produto eletrônico, há 12 localizações diferentes que podem acomodar *chips*.

- Se cinco tipos diferentes de *chips* devem ser colocados na placa, quantas disposições diferentes são possíveis?
- Se os cinco *chips* colocados na placa são do mesmo tipo, quantas disposições diferentes são possíveis?

2.2.11 Considere o projeto de um sistema de comunicação.

- Quantos prefixos de telefone com três dígitos, que são usados para representar uma área geográfica particular (tal como um código de área), podem ser criados a partir dos dígitos de 0 a 9?
- Como no item (a), quantos prefixos de telefone com três dígitos são possíveis que não comecem com 0 ou 1, porém contêm 0 ou 1 como dígitos intermediários?
- Quantos prefixos de telefone com três dígitos são possíveis em que nenhum dígito apareça mais de uma vez em cada prefixo?

2.2.12 No projeto de um produto eletromecânico, 12 componentes diferentes devem ser empilhados em um revestimento cilíndrico que prende 12 componentes de uma maneira que minimiza o impacto a choques. Uma ponta do revestimento é projetada como o fundo e outra como o topo.

- Se todos os componentes forem diferentes, quantos projetos diferentes serão possíveis?
- Se sete componentes forem todos idênticos, quantos projetos diferentes serão possíveis?
- Se três componentes forem de um tipo, mas idênticos entre si, e quatro componentes forem de outro tipo, mas idênticos entre si, porém os outros sendo diferentes, quantos projetos diferentes são possíveis?

2.2.13 Um silo de 50 itens contém cinco que são defeituosos. Uma amostra de dez itens é selecionada ao acaso, sem reposição. Quantas amostras contêm no mínimo quatro itens defeituosos?

2.2.14 Itens plásticos produzidos por uma operação de moldagem por injeção são verificados em relação a conformidades relativas a certas especificações. Cada ferramenta contém 12 cavidades, em que os itens são produzidos. Esses itens caem em um transportador quando a prensa se abre. Um inspetor escolhe aleatoriamente três itens entre os 12. Duas cavidades são afetadas por mau funcionamento do controlador de temperatura, que resulta em itens que não obedecem às especificações.

- Quantas amostras contêm exatamente um item não conforme?
- Quantas amostras contêm no mínimo um item não conforme?

2.2.15 Similar à programação de um hospital do Exemplo 2.9, suponha que um centro cirúrgico necessita lidar com cirurgias em três joelhos, em quatro quadris e em cinco ombros.

- Quantas sequências diferentes são possíveis?
- Quantas sequências diferentes de todas as cirurgias nos quadris, no joelho e no ombro foram programadas consecutivamente?
- Quantas programações diferentes começam e terminam com uma cirurgia de joelho?

## Exercícios para a Seção 2.3

2.3.1 O espaço amostral de um experimento aleatório é  $\{a, b, c, d, e\}$ , com probabilidades 0,1; 0,1; 0,2; 0,4 e 0,2, respectivamente. Seja  $A$  o evento  $\{a, b, c\}$  e seja  $B$  o evento  $\{c, d, e\}$ . Determine o seguinte:

- $P(A)$
- $P(B)$
- $P(A')$
- $P(A \cup B)$
- $P(A \cap B)$

2.3.2 Uma peça selecionada para teste é igualmente provável de ter sido produzida por qualquer uma das seis ferramentas de corte.

- Qual é o espaço amostral?
- Qual é a probabilidade de a peça ser proveniente da ferramenta 1?
- Qual é a probabilidade de a peça ser proveniente da ferramenta 3 ou 5?
- Qual é a probabilidade de a peça não ser proveniente da ferramenta 4?

#### 4 Exercícios Complementares

2.3.3 Uma peça moldada por injeção é igualmente provável de ser obtida a partir de qualquer uma das oito cavidades de um molde.

- Qual é o espaço amostral?
- Qual é a probabilidade de a peça ser proveniente da cavidade 1 ou 2?
- Qual é a probabilidade de a peça não ser proveniente nem da cavidade 3 nem da 4?

2.3.4 Um cartão de crédito contém 16 dígitos entre 0 e 9. No entanto, somente 100 milhões de números são válidos. Se um número for selecionado aleatoriamente, qual é a probabilidade de que ele seja um número válido?

2.3.5 Em uma bateria de NiCd, uma célula completamente carregada é composta de hidróxido de níquel. Níquel é um elemento que tem múltiplos estados de oxidação, sendo geralmente encontrado nos seguintes estados:

Carga de Níquel	Proporções Encontradas
0	0,17
+2	0,35
+3	0,33
+4	0,15

- Qual é a probabilidade de uma célula ter no mínimo uma das opções de níquel carregada positivamente?
  - Qual é a probabilidade de uma célula *não* ser composta de uma carga positiva de níquel maior do que +3?
- 2.3.6 Uma mensagem pode seguir diferentes caminhos por meio de servidores em uma rede. A mensagem do remetente pode ir para um de cinco servidores na primeira etapa; cada um deles pode enviar para cinco servidores na segunda etapa; cada um dos quais pode enviar para quatro servidores na terceira etapa; e então a mensagem vai para o servidor do remetente.
- Quantos caminhos são possíveis?
  - Se todos os caminhos forem igualmente prováveis, qual é a probabilidade de uma mensagem passar pelo primeiro de quatro servidores na terceira etapa?

2.3.7 Suponha que seu veículo seja licenciado em um estado que emita placas que consistam em três dígitos (entre 0 e 9), seguidos de três letras (entre A e Z). Se um número de placa for selecionado ao acaso, qual é a probabilidade de seu veículo ser selecionado?

2.3.8 Discos de plástico de policarbonato, provenientes de um fornecedor, são analisados com relação às resistências a arranhões e a choque. Os resultados de 100 discos são resumidos abaixo.

		Resistência a Choque	
		Alta	Baixa
Resistência a arranhões	Alta	70	9
	Baixa	16	5

Seja  $A$  o evento em que um disco tem alta resistência a choque e seja  $B$  o evento em que um disco tem alta resistência a arranhões. Se um disco for selecionado aleatoriamente, determine as seguintes probabilidades:

- $P(A)$
- $P(B)$
- $P(A')$
- $P(A \cap B)$
- $P(A \cup B)$
- $P(A' \cup B)$

2.3.9 Álcalis de magnésio são usados como catalisadores homogêneos na produção de polietileno linear de baixa densidade (PEBD), que requer um pó mais fino de magnésio para sustentar uma reação. Experimentos de reação redox, usando quatro quantidades diferentes de pó de magnésio, são realizados. Cada resultado pode ou não ser mais reduzido em uma segunda etapa usando três quantidades diferentes de pó de magnésio. Cada um dos resultados pode ou não ser mais reduzido em uma terceira etapa usando três quantidades diferentes de pó de magnésio.

- Quantos experimentos são possíveis?

b. Se todos os resultados forem igualmente prováveis, qual é a probabilidade de que o melhor resultado seja obtido a partir de um experimento que usa todas as três etapas?

c. O resultado no item (b) mudará, se cinco, seis ou sete quantidades diferentes forem usadas na primeira etapa? Explique.

2.3.10 Um artigo publicado no *Journal of Database Management* ["Experimental Study of a Self-Tuning Algorithm for DBMS Buffer Pools" (2005, Vol. 16, pp. 1-20)] forneceu a carga de trabalho usada na marca de referência TPC-C OLTP (Transaction Processing Performance Council's Version C On-Line Transaction Processing), que simula uma aplicação típica de ordem de entrada.

#### Frequências Médias e Operações em TPC-C

Transação	Frequência	Seleciona	Atualiza
Nova ordem	43	23	11
Pagamento	44	4,2	3
Status do pedido	4	11,4	0
Entrega	5	130	120
Nível do estoque	4	0	0

A frequência de cada tipo de transação (na segunda coluna) pode ser usada como a percentagem de cada tipo de transação. Mostre-se o número médio requerido de operações *seleciona* para cada tipo de transação. Seja  $A$  o evento de transações com o número médio de operações *seleciona* igual a 12 ou menos. Seja  $B$  o evento de transações com o número médio de operações *atualiza* igual a 12 ou menos. Calcule as seguintes probabilidades:

- $P(A)$
- $P(B)$
- $P(A \cap B)$
- $P(A \cap B')$
- $P(A \cup B)$

2.3.11 Amostras de emissões de três fornecedores são classificadas com relação a satisfazer as especificações de qualidade do ar. Os resultados de 100 amostras são resumidos a seguir.

		Conforme	
		Sim	Não
Fornecedor	1	22	8
	2	25	5
	3	30	10

Seja  $A$  o evento em que uma amostra seja proveniente do fornecedor 1 e seja  $B$  o evento em que uma amostra atenda às especificações. Se uma amostra for selecionada ao acaso, determine as seguintes probabilidades:

- $P(A)$
- $P(B)$
- $P(A')$
- $P(A \cap B)$
- $P(A \cup B)$
- $P(A' \cap B)$

2.3.12 Considere os dados da sala de emergência do hospital no Exemplo 2.6. Seja  $A$  o evento em que uma visita seja para o Hospital 4 e seja  $B$  o evento em que uma visita resulte em PNEM (em qualquer hospital). Determine as seguintes probabilidades:

- $P(A \cap B)$
- $P(A')$
- $P(A \cup B)$
- $P(A \cup B')$
- $P(A' \cap B')$

2.3.13 Considere os axiomas da probabilidade para mostrar o seguinte:

a. Para qualquer evento  $E$ ,  $P(E') = 1 - P(E)$ .

b.  $P(\emptyset) = 0$ .

c. Se  $A$  estiver contido em  $B$ , então  $P(A) \leq P(B)$ .

2.3.14 Suponha que um paciente seja selecionado aleatoriamente a partir daqueles descritos no Exercício 2.1.25. Seja  $A$  o evento em que o paciente está no grupo tratado com alfa interferon e seja  $B$  o evento em que o paciente tem uma resposta completa. Determine as seguintes probabilidades:

- $P(A)$
- $P(B)$
- $P(A \cap B)$
- $P(A \cup B)$
- $P(A' \cup B)$

2.3.15 Um anúncio na internet pode ser projetado a partir de quatro cores diferentes, de três tipos de fonte, de cinco tamanhos de fonte, de três imagens e de cinco frases. Um diagrama específico é gerado aleatoriamente pelo servidor da internet quando você visita o *site*. Se você visitar o *site* cinco vezes, qual será a probabilidade de que você não veja o mesmo projeto?

2.3.16 Similar à programação de um hospital do Exemplo 2.9, suponha que um centro cirúrgico necessite programar três cirurgias de joelho, quatro de quadris e cinco de ombros. Considere que todas as programações sejam igualmente prováveis. Determine a probabilidade para cada uma das seguintes situações:

- a. Todas as cirurgias de quadris são concluídas antes de outro tipo de cirurgia.

b. A programação começa com uma cirurgia de quadris.

c. A primeira e última cirurgias são de quadris.

d. As duas primeiras cirurgias são de quadris.

2.3.17 Um sistema de computadores usa senhas, que são seis caracteres, sendo cada caractere uma das 26 letras minúsculas ( $a-z$ ) ou 26 letras maiúsculas ( $A-Z$ ) ou dez inteiros (0-9). Seja  $\Omega$  o conjunto de todas as senhas possíveis e sejam  $A$  e  $B$  os eventos que consistem em senhas com somente letras ou com somente inteiros, respectivamente. Suponha que todas as senhas em  $\Omega$  sejam igualmente prováveis. Determine a probabilidade de cada uma das seguintes opções:

a.  $A$     b.  $B$

c. Uma senha contém no mínimo um inteiro.

d. Uma senha contém exatamente dois inteiros.

## Exercícios para a Seção 2.4

2.4.1 Se  $A$ ,  $B$  e  $C$  forem eventos mutuamente excludentes, com  $P(A) = 0,2$ ,  $P(B) = 0,3$  e  $P(C) = 0,4$ , determine as seguintes probabilidades:

- a.  $P(A \cup B \cup C)$     b.  $P(A \cap B \cap C)$   
 c.  $P(A \cap B)$     d.  $P[(A \cup B) \cap C]$   
 e.  $P(A' \cap B' \cap C')$

2.4.2 Se  $P(A) = 0,3$ ,  $P(B) = 0,2$  e  $P(A \cap B) = 0,1$ , determine as seguintes probabilidades:

- a.  $P(A')$     b.  $P(A \cup B)$     c.  $P(A' \cap B)$   
 d.  $P(A \cap B')$     e.  $P[(A \cup B)']$     f.  $P(A' \cup B)$

2.4.3 Um fabricante de faróis para automóveis testa lâmpadas sob ambientes com alta umidade e com alta temperatura, usando intensidade e vida útil como as respostas de interesse. A seguinte tabela mostra o desempenho de 130 lâmpadas.

		Vida Útil	
		Satisfatória	Não satisfatória
Intensidade	Satisfatória	117	3
	Não satisfatória	8	2

a. Encontre a probabilidade de uma lâmpada selecionada aleatoriamente fornecer resultados não satisfatórios sob qualquer critério.

b. Os consumidores dessas lâmpadas demandam 95 % de resultados satisfatórios. O fabricante de lâmpadas pode encontrar essa demanda?

2.4.4 No artigo "ACL Reconstruction Using Bone-Patellar Tendon-Bone Press-Fit Fixation: 10-Year Clinical Results", em *Knee Surgery, Sports Traumatology, Arthroscopy* (2005, Vol. 13, pp. 248-255), foram consideradas as seguintes causas para lesões no joelho:

Atividade	Porcentagem de Lesões no Joelho
Esporte de contato	46 %
Esporte sem contato	44 %
Atividade da vida diária	9 %
Dirigir motocicleta	1 %

a. Qual a probabilidade de uma lesão no joelho ter resultado de um esporte (contato ou sem contato)?

b. Qual a probabilidade de uma lesão no joelho ter resultado de uma atividade diferente de esporte?

2.4.5 Considere os dados da sala de emergência do hospital no Exemplo 2.6. Seja  $A$  o evento em que uma visita seja para o Hospital 4 e seja  $B$  o evento em que uma visita resulte em PNEM (em qualquer hospital). Aplique as regras de adição para calcular as seguintes probabilidades:

- a.  $P(A \cup B)$     b.  $P(A \cup B')$     c.  $P(A' \cup B')$

2.4.6 Cabos de fio de cobre, provenientes de um fabricante, são analisados em relação à resistência e à condutividade. Os resultados de 100 cabos são dados a seguir.

	Resistência	
	Alta	Baixa
Alta condutividade	74	8
Baixa condutividade	15	3

a. Se um cabo for selecionado aleatoriamente, qual será a probabilidade de sua condutividade ser alta e sua resistência ser alta?

b. Se um cabo for selecionado aleatoriamente, qual será a probabilidade de sua condutividade ser baixa ou sua resistência ser baixa?

c. Considere o evento em que um cabo tem baixa condutividade e o evento em que o cabo tem baixa resistência. Esses dois eventos são mutuamente excludentes?

2.4.7 Um sistema de computadores usa senhas que contêm exatamente seis caracteres e cada caractere é uma das 26 letras ( $a-z$ ) ou dez inteiros (0-9). Letras maiúsculas não são usadas. Seja  $A$  o evento em que uma senha comece com uma vogal ( $a, e, i, o$  ou  $u$ ) e seja  $B$  o evento em que a senha termine com um número par (0, 2, 4, 6 ou 8). Suponha que um invasor selecione uma senha ao acaso. Determine as seguintes probabilidades:

- a.  $P(A)$     b.  $P(B)$   
 c.  $P(A \cap B)$     d.  $P(A \cup B)$

2.4.8 Um anúncio na internet pode ser projetado a partir de quatro cores diferentes, de três tipos de fonte, de cinco tamanhos de fonte, de três imagens e de cinco frases. Um diagrama específico é gerado aleatoriamente pelo servidor da internet quando você visita o *site*. Seja  $A$  o evento em que a cor do diagrama seja vermelha e seja  $B$  o evento em que o tamanho da fonte não seja o menor de todos. Use as regras de adição para calcular as seguintes probabilidades:

- a.  $P(A \cup B)$     b.  $P(A \cup B')$     c.  $P(A' \cup B')$

6 Exercícios Complementares

2.4.9 Um sistema de computadores usa senhas que contêm exatamente oito caracteres, sendo cada caractere uma das 26 letras minúsculas ( $a-z$ ) ou 26 letras maiúsculas ( $A-Z$ ) ou dez inteiros (0-9). Suponha que todas as senhas sejam igualmente prováveis. Sejam  $A$  e  $B$  os eventos que consistem em senhas com somente letras ou com somente inteiros, respectivamente. Determine as seguintes probabilidades:

- a.  $P(A \cup B)$     b.  $P(A' \cup B)$   
 c.  $P(\text{Senha contém exatamente um ou dois inteiros})$

2.4.10 Considere os três grupos de pacientes no Exercício 2.1.25. Seja  $A$  o evento em que o paciente foi tratado com ribavirina mais alfa interferon e seja  $B$  o evento em que a resposta tenha sido completa. Determine as seguintes probabilidades:

- a.  $P(A \cup B)$     b.  $P(A' \cup B)$     c.  $P(A \cup B')$

2.4.11 O artigo “Clinical and Radiographic Outcomes of Four Different Treatment Strategies in Patients with Early Rheumatoid

Arthritis” [*Arthritis & Rheumatism* (2005, Vol. 52, pp. 3381-3390)] considerou quatro grupos de tratamento. Os grupos consistiram em pacientes com diferentes terapias com drogas (tais como prednisona e infliximab): monoterapia sequencial (grupo 1), terapia intensiva de combinação (grupo 2), terapia inicial de combinação (grupo 3) ou terapia inicial de combinação com infliximab (grupo 4). Radiografias de mãos e pés foram usadas para avaliar a progressão da doença. O número de pacientes sem progresso das injúrias nas articulações foi igual a 76 de 114 pacientes (67%), 82 de 112 pacientes (73%), 104 de 120 pacientes (87%) e 113 de 121 pacientes (93%) nos grupos 1-4, respectivamente. Suponha que um paciente seja selecionado aleatoriamente. Seja  $A$  o evento em que o paciente seja do grupo 1 e seja  $B$  o evento em que não haja progresso. Determine as seguintes probabilidades:

- a.  $P(A \cup B)$     b.  $P(A' \cup B')$     c.  $P(A \cup B')$

Exercícios para a Seção 2.5

2.5.1 A análise de resultados de um experimento de transmutação de uma folha (tornando uma folha em uma pétala) é resumida pelo tipo de transformação completada:

		Transformação Total da Textura	
		Sim	Não
Transformação total da cor	Sim	243	26
	Não	13	18

- a. Se a folha completar a transformação da cor, qual é a probabilidade de que ela completará a transformação na textura?  
 b. Se a folha não completar a transformação na textura, qual é a probabilidade de que ela completará a transformação da cor?

2.5.2 Amostras de uma pele com descamação são analisadas em relação ao teor de umidade e de melanina. Os resultados de 100 amostras de pele são dados a seguir.

		Conteúdo de Melanina	
		Alto	Baixo
Conteúdo de umidade	Alto	13	7
	Baixo	48	32

Seja  $A$  o evento em que uma amostra tem baixo teor de melanina e seja  $B$  o evento em que a amostra tem alto teor de umidade. Determine as seguintes probabilidades:

- a.  $P(A)$     b.  $P(B)$     c.  $P(A | B)$     d.  $P(B | A)$

2.5.3 A tabela a seguir resume o número de besouros mortos por autólise (a destruição de uma célula depois de sua morte pela ação de suas próprias enzimas) e por putrefação (decomposição de matéria orgânica, especialmente proteína, por microrganismos, resultando na produção de matéria malcheirosa).

		Autólise	
		Alta	Baixa
Putrefação	Alta	14	59
	Baixa	18	9

- a. Se a autólise de uma amostra for alta, qual será a probabilidade de que a putrefação seja baixa?  
 b. Se a putrefação de uma amostra for alta, qual será a probabilidade de que a autólise seja alta?

- c. Se a putrefação de uma amostra for baixa, qual será a probabilidade de que a autólise seja baixa?

2.5.4 Uma firma de manutenção tem coletado a seguinte informação relativa aos mecanismos de falha para os sistemas de ar-condicionado:

		Evidência de Vazamentos de Gás	
		Sim	Não
Evidência de falha elétrica	Sim	55	17
	Não	32	3

As unidades sem evidência de vazamentos de gás ou de falha elétrica mostraram outros tipos de falha. Se essa for uma amostra representativa de falha em ar-condicionado, encontre a probabilidade:

- a. De que a falha envolva um vazamento de gás.  
 b. De que haja evidência de uma falha elétrica, dado que tenha havido um vazamento de gás.  
 c. De que haja evidência de um vazamento de gás, dado que tenha havido uma falha elétrica.

2.5.5 Considere as reações endotérmicas no Exercício 2.1.22. Seja  $A$  o evento em que uma temperatura final de reação seja 271 K ou menos. Seja  $B$  o evento em que o calor absorvido esteja acima do valor desejado. Determine as seguintes probabilidades:

- a.  $P(A | B)$     b.  $P(A' | B)$   
 c.  $P(A | B')$     d.  $P(B | A)$

2.5.6 Uma batelada de 500 reservatórios para suco congelado de laranja contém cinco que são defeituosos. Três são selecionados, ao acaso e sem reposição, a partir da batelada.

- a. Qual é a probabilidade de que o segundo reservatório selecionado seja defeituoso, dado que o primeiro foi defeituoso?  
 b. Qual é a probabilidade de que ambos sejam defeituosos?  
 c. Qual é a probabilidade de que ambos sejam aceitáveis?  
 d. Qual é a probabilidade de que o terceiro reservatório selecionado seja defeituoso, dado que o primeiro e o segundo selecionados foram defeituosos?  
 e. Qual é a probabilidade de que o terceiro reservatório selecionado seja defeituoso, dado que o primeiro selecionado foi defeituoso e o segundo selecionado foi aceitável?  
 f. Qual é a probabilidade de que todos os três reservatórios sejam defeituosos?

2.5.7 Suponha que  $A$  e  $B$  sejam eventos mutuamente excludentes. Construa um diagrama de Venn que contenha os três eventos  $A$ ,  $B$  e  $C$ , de modo que  $P(A | C) = 1$  e  $P(B | C) = 0$ .

2.5.8 Um artigo em *The Canadian Entomologist* (Harcourt *et al.*, 1977, Vol. 109, pp. 1521-1534) estudou a vida da praga da alfafa a partir dos ovos até a vida adulta. A tabela seguinte mostra o número de larvas que sobreviveram em cada estágio do desenvolvimento, desde ovos até a vida adulta.

Ovos	Fase	Fase	Pré-pupa	Pupa	Adultos
	Precoce da Larva	Madura da Larva			
421	412	306	45	35	31

- Qual é a probabilidade de um ovo sobreviver até a vida adulta?
- Qual é a probabilidade de sobrevivência até a vida adulta, dada a sobrevivência para a fase madura da larva?
- Que estágio tem a menor probabilidade de sobrevivência para o próximo estágio?

2.5.9 Considere os dados da sala de emergência do hospital no Exemplo 2.6. Seja  $A$  o evento em que uma visita seja para o Hospital 4 e seja  $B$  o evento em que uma visita resulte em PNEM (em qualquer hospital). Determine as seguintes probabilidades:

- $P(A | B)$
- $P(A' | B)$
- $P(A | B')$
- $P(B | A)$

2.5.10 Um sistema computacional usa senhas que contêm exatamente oito caracteres e cada caractere é uma das 26 letras minúsculas ( $a-z$ ) ou 26 letras maiúsculas ( $A-Z$ ) ou dez inteiros (0-9). Seja  $\Omega$  o conjunto de todas as senhas possíveis. Suponha que todas as senhas em  $\Omega$  sejam igualmente prováveis. Determine a probabilidade de cada uma das seguintes opções:

- A senha é constituída só por letras minúsculas, dado que ela contém somente letras.
- A senha contém no mínimo uma letra maiúscula, dado que ela contém somente letras.
- A senha contém somente números pares, dado que ela contém todos os números.

2.5.11 Suponha que um paciente seja selecionado aleatoriamente a partir daqueles descritos no Exercício 2.4.11. Seja  $A$  o evento em que o paciente está no grupo 1 e seja  $B$  o evento para o qual não há progresso. Determine as seguintes probabilidades:

- $P(B | A)$
- $P(A | B)$
- $P(A | B')$
- $P(A' | B)$

## Exercícios para a Seção 2.6

2.6.1 Suponha que  $P(A | B) = 0,4$  e  $P(B) = 0,5$ . Determine o seguinte:

- $P(A \cap B)$
- $P(A' \cap B)$

2.6.2 Suponha que  $P(A | B) = 0,2$ ,  $P(A | B') = 0,3$  e  $P(B) = 0,8$ . Qual é  $P(A)$ ?

2.6.3 A probabilidade é 1 % de que um conector elétrico, que seja mantido seco, falhe durante o período de garantia. Se o conector for molhado, a probabilidade de falha durante o período de garantia será de 5 %. Se 90 % dos conectores forem mantidos secos e 10 % forem mantidos molhados, qual será a proporção de conectores que falharão durante o período de garantia?

2.6.4 Falhas no coração são por causa tanto de ocorrências naturais (87 %) como por fatores externos (13 %). Fatores externos estão relacionados com substâncias induzidas (73 %) ou com objetos estranhos (27 %). Ocorrências naturais são causadas por bloqueio arterial (56 %), doenças (27 %) e infecção (por exemplo, infecção por estafilococos) (17 %).

- Determine a probabilidade de uma falha ser causada por substância induzida.
- Determine a probabilidade de uma falha ser causada por doença ou infecção.

2.6.5 A aspereza nas bordas de produtos de papel cortado aumenta, à medida que as lâminas de uma faca vão sendo gastas. Somente 1 % dos produtos cortados com novas lâminas tem bordas ásperas, 3 % dos produtos cortados com lâminas mediamente afiadas exibem rugosidade e 5 % dos produtos cortados com lâminas gastas exibem rugosidade. Se 25 % das lâminas na fabricação de papel forem novas, 60 % forem mediamente afiadas e 15 % forem gastas, qual será a proporção dos produtos que exibem aspereza nas bordas?

2.6.6 Um lote de 100 *chips* semicondutores contém 20 que são defeituosos.

- Dois são selecionados, ao acaso e sem reposição, do lote. Determine a probabilidade de o segundo *chip* selecionado ser defeituoso.

b. Três são selecionados, ao acaso e sem reposição, do lote. Determine a probabilidade de todos os *chips* selecionados serem defeituosos.

2.6.7 Falhas em teclados de computadores ocorrem por causa de conexões elétricas imperfeitas (12 %) ou de defeitos mecânicos (88 %). Defeitos mecânicos estão relacionados com teclas soltas (27 %) ou com montagens impróprias (73 %). Defeitos de conexão elétrica são causados por fios defeituosos (35 %), por conexões impróprias (13 %) ou por fios deficientemente soldados (52 %).

- Encontre a probabilidade de uma falha ocorrer por causa de teclas soltas.
- Encontre a probabilidade de uma falha ocorrer por causa de fios imprópriamente conectados ou deficientemente soldados.

2.6.8 Um artigo publicado no *British Medical Journal* ["Comparison of Treatment of Renal Calculi by Operative Surgery, Percutaneous Nephrolithotomy, and Extracorporeal Shock Wave Lithotripsy" (1986, Vol. 82, pp. 879-892)] forneceu a seguinte discussão acerca de taxas de sucesso na remoção de pedras nos rins. Cirurgia aberta tem uma taxa de sucesso de 78 % (273/350), enquanto um método mais novo, nefrolitotomia percutânea (NP), tem uma taxa de sucesso de 83 % (289/350). Esse novo método pareceu melhor, mas os resultados mudaram quando o diâmetro da pedra foi considerado. Para pedras com diâmetros menores do que 2 cm, 93 % (81/87) de casos de cirurgia aberta obtiveram sucesso, comparados com somente 83 % (234/270) de casos de NP. Para pedras maiores do que ou iguais a 2 cm, as taxas de sucesso foram 73 % (192/263) e 69 % (55/80) para cirurgia aberta e NP, respectivamente. Cirurgia aberta é melhor para ambos os tamanhos de pedra, porém tem menos sucesso no total. Em 1951, E. H. Simpson alertou para essa aparente contradição (conhecida como **paradoxo de Simpson**), porém o risco ainda hoje persiste. Explique como a cirurgia aberta pode ser melhor para ambos os tamanhos de pedra, porém pior para o total.

## 8 Exercícios Complementares

2.6.9 Considere os dados da sala de emergência do hospital no Exemplo 2.6. Seja  $A$  o evento em que uma visita seja para o Hospital 4 e seja  $B$  o evento em que uma visita resulte em PNEM (em qualquer hospital). Determine as seguintes probabilidades:

- a.  $P(A \cap B)$     b.  $P(A \cup B)$     c.  $P(A' \cup B')$

d. Use a regra da probabilidade total para determinar  $P(A)$ .

2.6.10 Suponha que um paciente seja selecionado aleatoriamente a partir daqueles pacientes descritos no Exercício 2.4.11. Seja  $A$  o evento em que o paciente está no grupo 1 e seja  $B$  o evento para o qual não há progresso. Determine as seguintes probabilidades:

- a.  $P(A \cap B)$     b.  $P(B)$     c.  $P(A' \cap B)$

- d.  $P(A \cup B)$     e.  $P(A' \cup B)$

2.6.11 Um anúncio na internet pode ser projetado a partir de quatro cores diferentes, de três tipos de fonte, de cinco tamanhos de fonte, de três imagens e de cinco frases. Um diagrama específico é gerado

aleatoriamente pelo servidor da internet quando você visita o *site*. Determine a probabilidade de que o anúncio tenha cor vermelha e o tamanho da fonte não seja o menor.

2.6.12 Similar ao cronograma de um hospital no Exemplo 2.9, suponha que um centro cirúrgico necessita programar três cirurgias de joelho, quatro de quadris e cinco de ombros. Considere que todos os cronogramas sejam igualmente prováveis. Determine as seguintes probabilidades:

- a. Todas as cirurgias de quadris são concluídas em primeiro lugar, dado que todas as cirurgias de joelho são as últimas.  
 b. A programação começa com uma cirurgia de quadris, dado que todas as cirurgias de joelho são as últimas.  
 c. A primeira e a última cirurgias são de quadris, dado que as cirurgias de joelho são programadas em períodos de tempo 2 a 4.  
 d. As duas primeiras cirurgias são de quadris, dado que todas as cirurgias de joelho são as últimas.

## Exercícios para a Seção 2.7

2.7.1 Se  $P(A|B) = 0,3$ ,  $P(B) = 0,8$  e  $P(A) = 0,3$ , o evento  $B$  e o evento complementar de  $A$  são independentes?

2.7.2 Se  $P(A) = 0,2$ ,  $P(B) = 0,2$  e  $A$  e  $B$  são mutuamente excludentes, eles são independentes?

2.7.3 Uma batelada de 500 reservatórios para suco congelado de laranja contém cinco que são defeituosos. Dois são selecionados, ao acaso e sem reposição, a partir da batelada. Sejam  $A$  e  $B$  os eventos em que o primeiro e o segundo reservatórios selecionados sejam defeituosos, respectivamente.

- a. Os eventos  $A$  e  $B$  são independentes?  
 b. Se a amostragem fosse feita com reposição,  $A$  e  $B$  seriam independentes?

2.7.4 Discos de plástico de policarbonato, provenientes de um fornecedor, são analisados com relação à resistência a arranhões e a choque. Os resultados de 100 discos são resumidos a seguir.

		Resistência a Choque	
		Alta	Baixa
Resistência a arranhões	Alta	70	9
	Baixa	16	5

Seja  $A$  o evento em que um disco tenha alta resistência a choque e seja  $B$  o evento em que um disco tenha alta resistência a arranhões. Os eventos  $A$  e  $B$  são independentes?

2.7.5 Matriz redundante de discos baratos (RAID — *Redundant Array of Inexpensive Disks*) é uma tecnologia que usa discos rígidos múltiplos para aumentar a velocidade de transferência de dados e fornecer cópia de segurança instantânea de dados. Suponha que a probabilidade de qualquer disco rígido falhar em um dia seja de 0,001 e que as falhas do disco sejam independentes.

- a. Um esquema de RAID 0 usa dois discos rígidos, cada um contendo uma imagem do outro no espelho. Qual é a probabilidade de perda de dados? Considere que a perda de dados ocorrerá se ambos os discos falharem dentro do mesmo dia.  
 b. Um esquema de RAID 1 divide os dados em dois discos rígidos. Qual é a probabilidade de perda de dados? Considere que a perda de dados ocorrerá se no mínimo um disco falhar dentro do mesmo dia.

2.7.6 Em um teste de placa de circuito impresso, usando um teste-padrão aleatório, um arranjo de dez *bits* é igualmente provável de ser 0 ou 1. Considere que os *bits* sejam independentes.

- a. Qual é a probabilidade de todos os *bits* serem iguais a 1?  
 b. Qual é a probabilidade de todos os *bits* serem iguais a 0?  
 c. Qual é a probabilidade de exatamente cinco *bits* serem iguais a 1 e cinco *bits* serem iguais a 0?

2.7.7 A probabilidade de um espécime de laboratório conter altos níveis de contaminação é 0,10. Cinco amostras são verificadas, sendo elas independentes.

- a. Qual é a probabilidade de nenhum espécime conter altos níveis de contaminação?  
 b. Qual é a probabilidade de exatamente um espécime conter altos níveis de contaminação?  
 c. Qual é a probabilidade de pelo menos um espécime conter altos níveis de contaminação?

2.7.8 Um jogador de *videogame* é confrontado com uma série de quatro oponentes e uma probabilidade de 80 % de derrotar cada oponente. Considere que os resultados dos oponentes são independentes (e que quando o jogador é derrotado por um oponente, o jogo acaba).

- a. Qual é a probabilidade de um jogador derrotar todos os quatro oponentes em um jogo?  
 b. Qual é a probabilidade de um jogador derrotar no mínimo dois oponentes em um jogo?  
 c. Se o jogo for jogado três vezes, qual será a probabilidade de um jogador derrotar todos os quatro oponentes no mínimo uma vez?

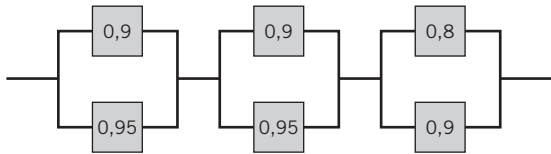
2.7.9 Oito cavidades de uma ferramenta de moldagem por injeção produzem conectores de plástico que caem em uma corrente comum. Uma amostra é escolhida a cada intervalo de vários minutos. Considere que as amostras sejam independentes.

- a. Qual é a probabilidade de cinco amostras sucessivas terem sido produzidas na cavidade 1 do molde?  
 b. Qual é a probabilidade de cinco amostras sucessivas terem sido produzidas na mesma cavidade do molde?  
 c. Qual é a probabilidade de quatro de cinco amostras sucessivas terem sido produzidas na cavidade 1 do molde?

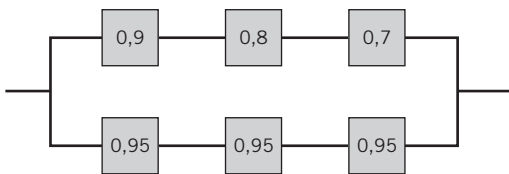
2.7.10 Um cartão de crédito contém 16 dígitos. Ele contém também mês e ano de expiração. Suponha que haja um milhão de usuários de cartão de crédito com números únicos de cartão. Um invasor seleciona aleatoriamente um número de cartão de crédito de 16 dígitos.

- a. Qual é a probabilidade de que ele pertença a um usuário?  
 b. Suponha que um invasor tenha uma chance de 25 % de adivinhar corretamente o ano de expiração de seu cartão e que selecione aleatoriamente um dos 12 meses. Qual é a probabilidade de o invasor selecionar corretamente o mês e o ano de expiração?

2.7.11 O circuito a seguir opera se, e somente se, houver uma rota de dispositivos funcionais, da esquerda para a direita. A probabilidade de cada aparelho funcionar é mostrada. Suponha que a probabilidade de um dispositivo ser funcional não depende de outros dispositivos serem ou não funcionais. Qual será a probabilidade de o circuito operar?



2.7.12 O circuito a seguir opera se, e somente se, houver um caminho de dispositivos funcionais, da esquerda para a direita. A probabilidade de cada aparelho funcionar é mostrada. Suponha que a probabilidade de um dispositivo ser funcional não depende de outros dispositivos serem ou não funcionais. Qual é a probabilidade de o circuito operar?



2.7.13 Considere os dados da sala de emergência do hospital no Exemplo 2.6. Seja  $A$  o evento em que uma visita é para o Hospital 4 e seja  $B$  o evento em que uma visita resulte em PNEM (em qualquer hospital). Esses eventos são independentes?

2.7.14 Um anúncio na internet pode ser projetado a partir de quatro cores diferentes, de três tipos de fonte, de cinco tamanhos de fonte, de três imagens e de cinco frases. Um diagrama específico é gerado aleatoriamente pelo servidor da internet quando você visita o *site*. Seja  $A$  o evento em que a cor do diagrama é vermelha e seja  $B$  o evento em que o tamanho da fonte não é o menor de todos. Os eventos  $A$  e  $B$  são independentes? Explique por que sim ou por que não.

2.7.15 Um circuito integrado contém dez milhões de portas lógicas (cada um pode ser um circuito lógico E ou OU). Admita que a probabilidade de uma falha na porta seja  $p$  e que as falhas sejam independentes. O circuito integrado para de funcionar, se qualquer porta falhar. Determine o valor para  $p$ , de modo que a probabilidade de o circuito integrado funcionar seja 0,95.

2.7.16 A Tabela 2.1 fornece dados sobre pastilhas, categorizados por localização e níveis de contaminação. Considere  $A$  o evento em que a contaminação seja *baixa* e  $B$  o evento em que a localização esteja no *centro*. Os eventos  $A$  e  $B$  são independentes? Por que sim ou por que não?

## Exercícios para a Seção 2.8

2.8.1 Consumidores são usados para avaliar projetos iniciais de produtos. No passado, 95 % dos produtos altamente aprovados recebiam bons conceitos, 60 % dos produtos moderadamente aprovados recebiam bons conceitos e 10 % dos produtos ruins recebiam bons conceitos. Além disso, 40 % dos produtos tinham sido altamente aprovados, 35 % moderadamente aprovados e 25 % tinham sido produtos ruins.

- a. Qual é a probabilidade de um produto atingir um bom conceito?  
 b. Se um novo projeto atingir um bom conceito, qual a probabilidade de que ele será um produto altamente aprovado?  
 c. Se um produto não atingir uma boa revisão, qual a probabilidade de que ele será um produto altamente aprovado?

2.8.2 Suponha que  $P(A | B) = 0,7$ ,  $P(A) = 0,5$  e  $P(B) = 0,2$ . Determine  $P(B | A)$ .

2.8.3 Está sendo testado um novo método analítico de detectar poluentes em água. Esse novo método de análise química é importante porque, se adotado, poderia ser usado para detectar três diferentes contaminantes — poluentes orgânicos, solventes voláteis e compostos clorados — em vez de ter de usar um único teste para cada poluente. As pessoas que elaboraram o teste afirmam que ele pode detectar altos níveis de poluentes orgânicos com 99,7 % de acurácia, solventes voláteis com 99,95 % de acurácia e compostos clorados com 89,7 % de acurácia. Se um poluente não estiver presente, o teste não acusa. Amostras são preparadas para a calibração do teste e 60 % delas são

contaminados com poluentes orgânicos, 27 % com solventes voláteis e 13 % com traços de compostos clorados. Uma amostra teste é selecionada aleatoriamente.

- a. Qual é a probabilidade de o teste acusar?  
 b. Se o teste acusar, qual a probabilidade de os compostos clorados estarem presentes?

2.8.4 Um programa computacional para detectar fraudes em cartões telefônicos dos consumidores rastreia, todo dia, o número de áreas metropolitanas de onde as ligações se originam. Sabe-se que 1 % dos usuários legítimos faz suas ligações de duas ou mais áreas metropolitanas em um único dia. Entretanto, 30 % dos usuários fraudulentos fazem suas ligações de duas ou mais áreas metropolitanas em um único dia. A proporção de usuários fraudulentos é 0,01 %. Se o mesmo usuário fizer suas ligações de duas ou mais áreas metropolitanas em um único dia, qual será a probabilidade de o usuário ser fraudulento?

2.8.5 Considere os dados da sala de emergência do hospital no Exemplo 2.6. Empregue o Teorema de Bayes para calcular a probabilidade de uma pessoa visitar o Hospital 4, dado que eles são PNEM.

2.8.6 Suponha que um paciente seja selecionado aleatoriamente daqueles descritos no Exercício 2.4.11. Seja  $A$  o evento em que o paciente esteja no grupo 1 e seja  $B$  o evento em que não haja progresso. Determine as seguintes probabilidades:

- a.  $P(B)$    b.  $P(B | A)$    c.  $P(A | B)$

## 10 Exercícios Complementares

2.8.7 Duas cores são usadas em um anúncio na internet. Se um visitante do *site* for de uma afiliada, as probabilidades das cores azul ou verde são 0,8 e 0,2, respectivamente. Se um visitante do *site* for de um *site* de busca, as cores azul e verde são 0,4 e 0,6, respectivamente. As proporções de visitantes a partir das afiliadas e de *sites* de busca são 0,3 e 0,7, respectivamente. Qual é a probabilidade de que um visitante seja proveniente do *site* de busca, dado que o anúncio azul foi visto?

2.8.8 Um fornecedor de equipamentos para recreação encontra que entre ordens de compra incluindo tendas, 40 % também incluem sacos de dormir. Somente 5 % das ordens de compra que não incluem tendas incluem sacos de dormir. Além disso, 20 %

das ordens de compra incluem tendas. Determine as seguintes probabilidades:

- A ordem de compra inclui colchões de dormir.
- A ordem de compra inclui tenda, dado que ela inclui colchão de dormir.

2.8.9 Um filtro de *e-mail* é planejado para separar *e-mails* válidos de *spam*. A palavra *livre* ocorre em 60 % das mensagens em *spam* e em somente 4 % das mensagens válidas. Além disso, 20 % das mensagens são *spam*. Determine as seguintes probabilidades:

- A mensagem contém a palavra *livre*.
- A mensagem é *spam*, dado que ela contém a palavra *livre*.
- A mensagem é válida, dado que ela não contém a palavra *livre*.

## Exercícios para a Seção 2.9

2.9.1 Decida se uma variável aleatória discreta ou contínua é o melhor modelo para cada uma das variáveis a seguir.

- O número de rachaduras que excedem meia polegada em dez milhas de uma autoestrada interestadual.
- O peso de uma peça plástica moldada por injeção.
- O número de moléculas em uma amostra de gás.
- A concentração de saída de um reator.
- A corrente em um circuito elétrico.

2.9.2 Decida se uma variável aleatória discreta ou contínua é o melhor modelo para cada uma das variáveis a seguir.

- O tempo que um projétil gasta para retornar à Terra.
- O número de vezes que um transistor em uma memória de computador muda de estado em uma operação.

- O volume de gasolina que é perdido por evaporação, durante o enchimento de um tanque de gasolina.
- O diâmetro externo de um eixo usinado.

2.9.3 Decida se uma variável aleatória discreta ou contínua é o melhor modelo para cada uma das variáveis a seguir.

- O tempo para um algoritmo computacional atribuir uma imagem a uma categoria.
- O número de *bytes* usado para estocar um arquivo em um computador.
- A concentração de ozônio em microgramas por metro cúbico.
- A fração de ejeção (fração volumétrica de sangue bombeado em cada batimento, a partir de um ventrículo no coração).
- A taxa volumétrica em litros por minuto.

## Exercícios Suplementares para o Capítulo 2

2.S4 Amostras de vidrarias de laboratório são embaladas em pacotes pequenos e leves ou em pacotes grandes e pesados. Suponha que 2 % e 1 % da amostra despachada em pequenos e grandes pacotes, respectivamente, se quebrem durante o transporte. Se 60 % das amostras forem despachados em grandes pacotes e 40 % delas forem despachados em pequenos pacotes, qual será a proporção de amostras que se quebrarão durante o transporte?

2.S5 Amostras de uma peça de alumínio fundido são classificadas com base no acabamento (em micropolegadas) na superfície e nas bordas. Os resultados de 100 peças são resumidos a seguir.

		Acabamento das Bordas	
		Excelente	Bom
Acabamento da superfície	Excelente	80	2
	Bom	10	8

Seja  $A$  o evento em que uma amostra tem excelente acabamento na superfície e seja  $B$  o evento em que uma amostra tem excelente acabamento nas bordas. Se um item for selecionado aleatoriamente, determine as seguintes probabilidades:

- $P(A)$
- $P(B)$
- $P(A')$
- $P(A \cap B)$
- $P(A \cup B)$
- $P(A' \cup B)$

2.S6 Um lote contém 15 moldes provenientes de um fornecedor local e 25 moldes provenientes de um fornecedor de um estado vizinho.

Dois moldes são selecionados ao acaso, sem reposição, do lote de 40. Seja  $A$  o evento em que o primeiro molde selecionado seja proveniente do fornecedor local e seja  $B$  o evento em que o segundo molde selecionado seja proveniente do fornecedor local. Determine:

- $P(A)$
- $P(B | A)$
- $P(A \cap B)$
- $P(A \cup B)$

Suponha que três moldes sejam selecionados ao acaso, sem reposição, do lote de 40. Além da definição dos eventos  $A$  e  $B$ , seja  $C$  o evento em que o terceiro molde selecionado seja proveniente do fornecedor local. Determine:

- $P(A \cap B \cap C)$
- $P(A \cap B \cap C')$

2.S7 Se  $A$ ,  $B$  e  $C$  forem eventos mutuamente excludentes, é possível que  $P(A) = 0,3$ ,  $P(B) = 0,4$  e  $P(C) = 0,5$ ? Por que sim ou por que não?

2.S8 Ligações recebidas por um centro de serviço de atendimento ao consumidor são classificadas como reclamações (75 % delas) ou pedidos de informação (25 % delas). Das reclamações, 40 % se referem a computadores que não respondem e 57 % se referem à instalação incompleta de programas; nos 3 % restantes das reclamações, o usuário não seguiu apropriadamente as instruções de instalação. Os pedidos de informação são igualmente divididos entre perguntas técnicas (50 %) e solicitações de compra de mais produtos (50 %).

a. Qual é a probabilidade de uma ligação recebida pelo centro de serviço de atendimento ao consumidor ser proveniente de um consumidor que não tenha seguido apropriadamente as instruções de instalação?

b. Encontre a probabilidade de uma ligação recebida ser uma solicitação de compra de mais produtos.

2.S9 Na fabricação de um adesivo químico, 3 % de todas as bateladas têm matérias-primas provenientes de dois lotes diferentes. Isso ocorre quando tanques de retenção são reabastecidos e a porção restante de um lote é insuficiente para encher os tanques.

Somente 5 % das bateladas com material de um único lote requerem reprocessamento. Entretanto, a viscosidade das bateladas consistindo em dois ou mais lotes de material é mais difícil de controlar e 40 % de tais bateladas requerem processamento adicional para atingir a viscosidade requerida.

Seja  $A$  o evento em que uma batelada é formada de dois lotes diferentes e seja  $B$  o evento em que um lote requer processamento adicional. Determine as seguintes probabilidades:

- a.  $P(A)$       b.  $P(A')$       c.  $P(B | A)$
- d.  $P(B | A')$     e.  $P(A \cap B)$     f.  $P(A \cap B')$     g.  $P(B)$

2.S10 Lasers semicondutores usados em produtos ópticos de armazenagem requerem níveis mais altos de potência para operações de escrita do que para operações de leitura. Operações com níveis de alta potência baixam a vida útil do *laser*.

*Lasers* em produtos usados para realizar cópias de discos magnéticos de alta velocidade gravam, principalmente, e a probabilidade de que a vida útil exceda cinco anos é de 0,95. *Lasers* que estejam em produtos que são usados para armazenagem principal gastam aproximadamente uma quantidade igual de tempo na leitura e na gravação, e a probabilidade de que a vida útil exceda cinco anos é de 0,995. Agora, 25 % dos produtos de um fabricante são usados para cópias e 75 % dos produtos são usados para armazenagem principal.

Seja  $A$  o evento em que a vida útil de um *laser* excede cinco anos e seja  $B$  o evento em que um *laser* está em um produto que é usado para copiar.

Considere o diagrama em forma de árvore para determinar o seguinte.

- a.  $P(B)$       b.  $P(A | B)$       c.  $P(A | B')$
- d.  $P(A \cap B)$     e.  $P(A \cap B')$     f.  $P(A)$

g. Qual é a probabilidade de que a vida útil de um *laser* exceda cinco anos?

h. Qual é a probabilidade de um *laser*, que falhou antes de cinco anos, ser proveniente de um produto usado para cópias?

2.S11 Uma rede congestionada de computadores tem uma probabilidade de 0,002 de perder um pacote (*packet*) de dados, e as perdas desses pacotes são eventos independentes. Um pacote perdido tem de ser reenviado.

- a. Qual é a probabilidade de se precisar reenviar qualquer pacote de uma mensagem de *e-mail* com 100 blocos?
- b. Qual é a probabilidade de se precisar reenviar exatamente um pacote de uma mensagem de *e-mail* com três blocos?
- c. Se dez mensagens de *e-mail* são enviadas, cada uma com 100 pacotes, qual é a probabilidade de se precisar reenviar pelo menos um pacote?

2.S12 Um aparelho óptico de armazenagem usa um procedimento de recuperação de erro que requer uma imediata retroleitura satisfatória de qualquer dado gravado. Se a retroleitura não tiver sucesso depois de três operações de gravação, aquele setor do disco será eliminado como inaceitável para armazenagem de dados. Em uma porção aceitável do disco, a probabilidade de uma retroleitura é de 0,98. Suponha

que as retroleituras sejam independentes. Qual será a probabilidade de uma porção aceitável do disco ser eliminada como inaceitável para armazenagem de dados?

2.S13 Energia liberada de três células quebra a ligação molecular e converte ATP (adenosina trifostato) em ADP (adenosina difosfato). Armazenagem de ATP em células dos músculos (mesmo para um atleta) pode somente sustentar máxima potência do músculo por menos de cinco segundos (uma curta arrancada). Três sistemas são usados para prover novamente ATP: sistema fosfogênico, sistema glicogênico-ácido láctico (anaeróbico) e respiração aeróbica; porém, o primeiro é somente útil por menos de dez segundos e mesmo o segundo sistema fornece menos de dois minutos de ATP. Um atleta de enduro necessita de desempenho abaixo do limite anaeróbico para manter energia para períodos extras. Uma amostra de 100 indivíduos é descrita pelo sistema de energia usado em exercícios com diferentes níveis de intensidade.

Período	Principalmente Aeróbica	
	Sim	Não
1	50	7
2	13	30

Seja  $A$  o evento em que um indivíduo está no período 2 e seja  $B$  o evento em que a energia é principalmente aeróbica. Determine o número de indivíduos em

- a.  $A' \cap B$     b.  $B'$     c.  $A \cup B$

2.S14 A probabilidade de um pedido de compra de um consumidor não ser despachado em tempo é igual a 0,05. Determinado consumidor faz três pedidos, que são despachados em tempos espaçados o suficiente para considerá-los como eventos independentes.

- a. Qual é a probabilidade de todos os pedidos serem despachados no tempo exato?
- b. Qual é a probabilidade de exatamente um pedido não ser despachado no tempo exato?
- c. Qual é a probabilidade de dois ou mais pedidos não serem despachados no tempo exato?

2.S15 Em um teste de placas de circuito impresso, cada placa falha ou não durante o teste. Uma placa que falhe no teste é então verificada em detalhes, de modo a determinar qual dos cinco tipos de defeitos é o principal modo de falha. Represente o espaço amostral para esse experimento.

2.S16 A inclusão de novos itens ou a troca para itens prévios são transações realizadas em um banco computacional de dados. A adição de um item pode ser completada em menos de 100 milissegundos 90 % das vezes, mas somente 20 % das mudanças para um item prévio podem ser completadas em um tempo menor que esse. Se 30 % das transações são mudanças, qual é a probabilidade de uma transação poder ser completada em menos de 100 milissegundos?

2.S17 Sejam  $E_1, E_2$  e  $E_3$  as amostras que obedecem à especificação de percentual de sólidos, à especificação de massa molar e à especificação de cor, respectivamente. Um total de 240 amostras é classificado de acordo com as especificações de  $E_1, E_2$  e  $E_3$ , em que *sim* indica que a amostra obedece.

	$E_3$ sim	$E_2$		Total
		Sim	Não	
$E_1$	Sim	200	1	201
	Não	5	4	9
Total		205	5	210

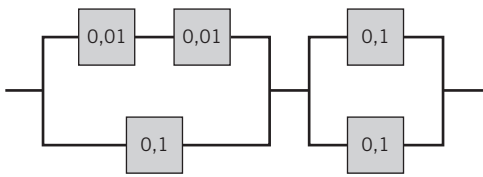
## 12 Exercícios Complementares

$E_3$  não

		$E_2$		Total
		Sim	Não	
$E_1$	Sim	20	4	24
	Não	6	0	6
Total		26	4	30

- Os eventos  $E_1$ ,  $E_2$  e  $E_3$  são mutuamente excludentes?
- Os eventos  $E'_1$ ,  $E'_2$  e  $E'_3$  são mutuamente excludentes?
- Qual é  $P(E'_1 \text{ ou } E'_2 \text{ ou } E'_3)$ ?
- Qual é a probabilidade de uma amostra obedecer às três especificações?
- Qual é a probabilidade de uma amostra obedecer à especificação  $E_1$  ou  $E_3$ ?
- Qual é a probabilidade de uma amostra obedecer às especificações  $E_1$  ou  $E_2$  ou  $E_3$ ?

2.S18 O circuito a seguir opera se, e somente se, houver um caminho de dispositivos funcionais da esquerda para a direita. Considere que dispositivos falhem independentemente, sendo a probabilidade de falha de cada um mostrada no diagrama. Qual é a probabilidade de o circuito operar?



2.S19 Uma placa de aço contém 20 parafusos. Considere que cinco parafusos não estejam apertados até o limite apropriado. Quatro parafusos são selecionados ao acaso, sem reposição, para verificação do torque.

- Qual é a probabilidade de todos os quatro parafusos selecionados estarem apertados até o limite apropriado?
- Qual é a probabilidade de no mínimo um dos parafusos selecionados não ter sido apertado até o limite apropriado?

2.S20 O governo britânico liberou, por envio de brochuras pelo correio para fazendeiros em todo o país, sua campanha de informação sobre doenças no pé e na boca. Estima-se que 99 % dos fazendeiros escoceses que recebem a brochura possuem informação suficiente para lidar com uma explosão da doença; porém, somente 90 % daqueles sem brochura podem lidar com uma explosão. Após os três primeiros meses de envio, 95 % dos fazendeiros na Escócia receberam a brochura informativa. Calcule a probabilidade de um fazendeiro selecionado aleatoriamente ter informação suficiente para lidar efetivamente com uma explosão da doença.

2.S21 Sabe-se que dois celulares defeituosos foram enviados erroneamente para um lote de remessa que tem agora um total de 75 telefones. Uma amostra de telefones será selecionada, sem reposição, do lote.

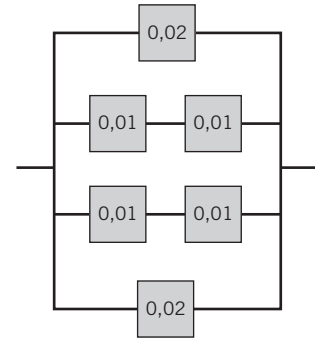
- Se três telefones forem inspecionados, determine a probabilidade de exatamente um dos telefones defeituosos ser encontrado.
- Se três telefones forem inspecionados, determine a probabilidade de ambos os telefones defeituosos serem encontrados.
- Se 73 telefones forem inspecionados, determine a probabilidade de ambos os telefones defeituosos serem encontrados. (Sugestão: Trabalhe com os telefones que continuam no lote.)

2.S22 Um sistema de codificação-decodificação consiste em três elementos: codifica, transmite e decodifica. Uma codificação falha ocorre em 0,5 % das mensagens processadas; erros de transmissão

ocorrem em 1 % das mensagens; e um erro de decodificação ocorre em 0,1 % das mensagens. Considere os erros como independentes.

- Qual é a probabilidade de se ter uma mensagem completamente livre de defeito?
- Qual é a probabilidade de uma mensagem ter tanto um defeito de codificação como de decodificação?

2.S23 O circuito a seguir opera se, e somente se, houver um caminho de dispositivos funcionais da esquerda para a direita. Considere que os dispositivos falhem independentemente e que a probabilidade de falha de cada equipamento esteja mostrada no diagrama. Qual é a probabilidade de o circuito não operar?



2.S24 Uma ferramenta de inserção robótica contém dez componentes principais. A probabilidade de qualquer componente falhar durante o período de garantia é de 0,01. Considere que os componentes falhem independentemente e que a ferramenta falhe se qualquer componente falhar. Qual é a probabilidade de a ferramenta falhar durante o período de garantia?

2.S25 Uma ferramenta de usinagem fica ociosa 15 % do tempo. Você requer uso imediato da ferramenta em cinco diferentes ocasiões durante o ano. Suponha que seus pedidos representem eventos independentes.

- Qual é a probabilidade de que a ferramenta esteja ociosa no período de todas as suas solicitações de uso?
- Qual é a probabilidade de que a ferramenta esteja ociosa no período de exatamente quatro de suas solicitações de uso?
- Qual é a probabilidade de que a ferramenta esteja ociosa no período de no mínimo três de suas solicitações de uso?

2.S26 Uma companhia que rastreia o uso de seu site na internet determinou que, quanto mais páginas um visitante visualizar, maior a probabilidade de o visitante solicitar mais informação (SMI). Use a seguinte tabela para responder às questões.

Número de páginas visualizadas:	1	2	3	4 ou mais
Porcentagem de visitantes:	40	30	20	10
Porcentagem de visitantes SMI:	10	10	20	40

- Qual é a probabilidade de um visitante ao site da internet solicitar mais informação?
- Se um visitante solicitar mais informação, qual será a probabilidade de o visitante ter visualizado quatro ou mais páginas?

2.S27 Um artigo em *Genome Research* ["An Assessment of Gene Prediction Accuracy in Large DNA Sequences" (2000, Vol. 10, pp. 1631-1642)] considerou a acurácia de um programa computacional comercial para prever nucleotídeos em sequências de genes. A tabela seguinte mostra o número de sequências para as quais os programas produziram previsões e o número de nucleotídeos corretamente previsto (calculado globalmente a partir do número total de sucessos e falhas de previsão em todas as sequências).

	Número de Sequências	Proporção
GenScan (Varredura de gene)	177	0,93
Blastx default (Algoritmo computacional)	175	0,91
Blastx topcomboN (Comando usado no Blastx)	174	0,97
Blastx em dois estágios	175	0,90
GeneWise (Algoritmo computacional)	177	0,98
Procrustes (Algoritmo computacional)	177	0,93

Considere que os sucessos e as falhas de previsão sejam independentes entre os programas.

- Qual é a probabilidade de que todos os programas prevejam corretamente um nucleotídeo?
  - Qual é a probabilidade de que todos os programas prevejam incorretamente um nucleotídeo?
  - Qual é a probabilidade de que no mínimo um programa Blastx preveja corretamente um nucleotídeo?
- 2.S28 Uma batelada contém 36 células de bactérias. Considere que 12 das células não sejam capazes de replicação celular. Seis células são selecionadas ao acaso, sem reposição, para verificar a replicação.

- Qual é a probabilidade de todas as seis células daquelas selecionadas serem capazes de replicação?
  - Qual é a probabilidade de no mínimo uma das células selecionadas não ser capaz de replicação?
- 2.S29 Um sistema de computadores usa senhas, que são exatamente sete caracteres, sendo cada caractere uma das 26 letras ( $a-z$ ) ou dez inteiros (0-9). Letras maiúsculas não são usadas.
- Quantas senhas são possíveis?
  - Se uma senha consistir em exatamente seis letras e um número, quantas senhas serão possíveis?
  - Se uma senha consistir em cinco letras seguidas por dois números, quantas senhas serão possíveis?
- 2.S30 Considere os tratamentos no Exercício 2.1.25. Suponha que um paciente seja selecionado aleatoriamente. Seja  $A$  o evento em que o paciente seja tratado com ribavirina mais alfa interferon ou alfa interferon, e seja  $B$  o evento em que o paciente teve uma resposta completa. Determine as seguintes probabilidades:
- $P(A | B)$
  - $P(B | A)$
  - $P(A \cap B)$
  - $P(A \cup B)$
- 2.S31 Considere os grupos de pacientes do Exercício 2.4.11. Suponha que um paciente seja selecionado aleatoriamente. Seja  $A$  o evento em que o paciente esteja no grupo 1 ou 2, e seja  $B$  o evento em que não haja progresso. Determine as seguintes probabilidades:
- $P(A | B)$
  - $P(B | A)$
  - $P(A \cap B)$
  - $P(A \cup B)$

## Capítulo 3 Exercícios

### Exercícios para a Seção 3.1

Para cada um dos exercícios seguintes, determine a faixa (valores possíveis) da variável aleatória.

- A variável aleatória é o número de conexões soldadas não conformes em uma placa de circuito impresso com 1000 conexões.
- Uma batelada de 500 peças usinadas contém dez que não atendem aos requisitos do consumidor. A variável aleatória é o número de peças em uma amostra de cinco peças que não atendem aos requisitos do consumidor.
- Uma batelada de 500 peças usinadas contém dez que não atendem aos requisitos do consumidor. Peças são selecionadas sucessivamente, sem reposição, até que uma peça não conforme seja obtida. A variável aleatória é o número de peças selecionadas.
- A variável aleatória é o número de ciclos do computador requerido para completar um cálculo aritmético selecionado.
- A variável aleatória é o número de falhas na superfície em uma grande serpentina de aço galvanizado.
- Um grupo de 10.000 pessoas é testado segundo um gene chamado *Ifi202*, que aumenta o risco de lúpus. A variável aleatória é o número de pessoas que carregam o gene.
- Mede-se o número de mutações em uma sequência de nucleotídeos de comprimento 40.000 em um filamento de DNA, depois da exposição à radiação. Cada nucleotídeo pode sofrer mutação.
- Uma clínica programa 30 minutos para cada visita de paciente, mas algumas visitas requerem tempo extra. A variável aleatória é o número de pacientes tratados em um dia de oito horas.
- O espaço amostral de um experimento aleatório é  $\{a, b, c, d, e, f\}$ , sendo cada resultado igualmente provável. Uma variável aleatória é definida como segue:

Resultado	$a$	$b$	$c$	$d$	$e$	$f$
$x$	0	0	1,5	1,5	2	3

Determine a função de probabilidade de  $X$ . Considere a função de probabilidade para determinar as seguintes probabilidades:

- $P(X = 1,5)$
- $P(0,5 < X < 2,7)$
- $P(X > 3)$
- $P(0 \leq X < 2)$
- $P(X = 0 \text{ ou } X = 2)$

Para os Exercícios 3.1.10 a 3.1.12, verifique se as seguintes funções são funções de probabilidade e determine as probabilidades requeridas.

3.1.10  $f(x) = (8/7)(1/2)^x$ ,  $x = 1, 2, 3$

- $P(X \leq 1)$
- $P(X > 1)$
- $P(2 < X < 6)$
- $P(X \leq 1 \text{ ou } X > 1)$

3.1.11  $f(x) = \frac{2x+1}{25}$ ,  $x = 0, 1, 2, 3, 4$

- $P(X = 4)$
- $P(X \leq 1)$
- $P(2 \leq X < 4)$
- $P(X > -10)$

3.1.12 Considere os pacientes do hospital no Exemplo 2.6. Dois pacientes são selecionados aleatoriamente, com reposição, a partir de todos os pacientes do Hospital 1. Qual é a função de probabilidade do número de pacientes na amostra que são admitidos?

$x$	1,25	1,5	1,75	2	2,25
$f(x)$	0,2	0,4	0,1	0,2	0,1

- $P(X \geq 2)$
- $P(X < 1,65)$
- $P(X = 1,5)$
- $P(X < 1,3 \text{ ou } X > 2,1)$

3.1.14 Um artigo em *Knee Surgery, Sports Traumatology, Arthroscopy* ["Arthroscopic Meniscal Repair with an Absorbable Screw: Results and Surgical Technique" (2005, Vol. 13, pp. 273-279)] menciona uma taxa de sucesso maior que 90 % para rompimentos de meniscos com rupturas com larguras menores que 3 mm, porém somente uma taxa de sucesso de 67 % para rompimentos de 3-6 mm. Se você for azarado o suficiente para sofrer um rompimento de menisco menor que 3 mm em seu joelho esquerdo e um de largura 3-6 mm em seu joelho direito, qual será a função de probabilidade do número de cirurgias com sucesso? Considere as cirurgias como independentes.

3.1.15 Em um processo de fabricação de semicondutores, três pastilhas de um lote são testadas. Cada pastilha é classificada como *passa* ou *falha*. Suponha que a probabilidade de uma pastilha passar no teste seja de 0,8 e que as pastilhas sejam independentes. Determine a função de probabilidade do número de pastilhas de um lote que passa no teste.

3.1.16 O setor de marketing estima que um novo instrumento para análise de amostras de solo terá grande sucesso, sucesso moderado, ou não terá sucesso, com probabilidades de 0,3, 0,6 e 0,1, respectivamente. A receita anual associada a um produto de grande sucesso, sucesso moderado, ou nenhum sucesso é de US\$ 10 milhões, US\$ 5 milhões e US\$ 1 milhão, respectivamente. Seja a variável aleatória  $X$  a renda anual do produto. Determine a função de probabilidade de  $X$ .

3.1.17 Um arranjo consiste em três componentes mecânicos. Suponha que as probabilidades de o primeiro, o segundo e o terceiro componentes satisfazerem as especificações sejam iguais a 0,95, 0,98 e 0,99, respectivamente. Considere que os componentes sejam independentes. Determine a função de probabilidade do número de componentes no arranjo que satisfazem as especificações.

3.1.18 A distribuição do tempo até uma página da internet mudar é importante para os indexadores automáticos (*Web crawlers*) que usam ferramentas de busca para manter a informação atual sobre *sites* da *Web*. A distribuição do tempo (em dias) até mudar uma página da internet é aproximada na seguinte tabela:

Dias até Mudanças	Probabilidade
1,5	0,05
3,0	0,25
4,5	0,35
5,0	0,20
7,0	0,15

Calcule a função de probabilidade dos dias até a mudança.

3.1.19 Os períodos de permanência efetiva em uma emergência de hospital em 2009 são mostrados na tabela seguinte (números arredondados para o inteiro mais próximo da hora). O período de permanência é o total do tempo de espera e de atendimento. Alguns períodos mais longos de permanência são também aproximados como 15 horas nessa tabela.

Horas	Frequência	Porcentagem
1	19	3,80
2	51	10,20
3	86	17,20
4	102	20,40
5	87	17,40
6	62	12,40
7	40	8,00
8	18	3,60
9	14	2,80
10	11	2,20
15	10	2,00

Calcule a função de probabilidade do tempo de espera para o atendimento.

## Exercícios para a Seção 3.2

3.2.1 Determine a função de distribuição cumulativa da variável aleatória do Exercício 3.1.9.

3.2.2 Determine a função de distribuição cumulativa da variável aleatória do Exercício 3.1.10.

3.2.3 Determine a função de distribuição cumulativa da variável aleatória do Exercício 3.1.11.

3.2.4 Determine a função de distribuição cumulativa da variável aleatória do Exercício 3.1.12. Determine também as seguintes probabilidades:

- a.  $P(X < 1,5)$     b.  $P(X \leq 3)$   
 c.  $P(X > 2)$     d.  $P(1 < X \leq 2)$

3.2.5 Determine a função de distribuição cumulativa da variável aleatória do Exercício 3.1.13.

3.2.6 Determine a função de distribuição cumulativa da variável aleatória do Exercício 3.1.16.

Verifique se as seguintes funções são funções de distribuição cumulativa e determine a função de probabilidade e as probabilidades requeridas.

$$3.2.7 F(x) = \begin{cases} 0 & x < 1 \\ 0,5 & 1 \leq x < 3 \\ 1 & 3 \leq x \end{cases}$$

- a.  $P(X \leq 3)$     b.  $P(X \leq 2)$   
 c.  $P(1 \leq X \leq 2)$     d.  $P(X > 2)$

$$3.2.8 F(x) = \begin{cases} 0 & x < -10 \\ 0,25 & -10 \leq x < 30 \\ 0,75 & 30 \leq x < 50 \\ 1 & 50 \leq x \end{cases}$$

- a.  $P(X \leq 50)$     b.  $P(X \leq 40)$   
 c.  $P(40 \leq X \leq 60)$     d.  $P(X < 0)$   
 e.  $P(0 \leq X < 10)$     f.  $P(-10 < X < 10)$

## Exercícios para a Seção 3.3

3.3.1 Se a faixa de  $X$  for o conjunto  $\{0, 1, 2, 3, 4\}$  e  $P(X = x) = 0,2$ , determine a média e a variância da variável aleatória.

3.3.2 Determine a média e a variância da variável aleatória no Exercício 3.1.10.

3.3.3 Determine a média e a variância da variável aleatória no Exercício 3.1.11.

3.3.4 Determine a média e a variância da variável aleatória no Exercício 3.1.12.

3.3.5 Determine a média e a variância da variável aleatória no Exercício 3.1.13.

3.3.6 Em uma bateria NiCd, uma célula completamente carregada é composta de hidróxido de níquel. Níquel é um elemento que tem múltiplos estados de oxidação. Considere as seguintes proporções dos estados:

Carga de Níquel	Proporções Encontradas
0	0,17
+2	0,35
+3	0,33
+4	0,15

a. Determine a função de distribuição cumulativa da carga de níquel.

b. Determine a média e a variância da carga de níquel.

3.3.7 A faixa da variável aleatória  $X$  é  $[0, 1, 2, 3, x]$ , em que  $x$  é uma incógnita. Se cada valor for igualmente provável e a média de  $X$  for igual a 6, determine  $x$ .

3.3.8 Árvores são sujeitas a diferentes níveis de atmosfera de dióxido de carbono com 6 % delas em uma condição mínima de crescimento a 350 partes por milhão (ppm) de  $\text{CO}_2$ , 10 % a 450 ppm (crescimento lento) de  $\text{CO}_2$ , 47 % a 550 ppm (crescimento moderado) de  $\text{CO}_2$  e 37 % a 650 ppm (crescimento rápido) de  $\text{CO}_2$ . Qual é a média e qual o desvio-padrão (em ppm) da atmosfera de dióxido de carbono para essas árvores?

## Exercícios para a Seção 3.4

3.4.1 Considere que os comprimentos de onda de radiações foto-sinteticamente ativas (RFA) estejam distribuídos uniformemente em nanômetros inteiros no espectro do vermelho de 675 a 700 nm.

a. Qual é a média e qual a variância da distribuição de comprimento de onda para essa radiação?

b. Se os comprimentos de onda estão uniformemente distribuídos em nanômetros inteiros de 75 a 100 nm, como podemos comparar a média e a variância da distribuição de comprimento de onda com o item anterior? Explique.

3.4.2 Considere que a variável aleatória  $X$  tem uma distribuição discreta uniforme nos inteiros  $0 \leq x \leq 99$ . Determine a média e a variância de  $X$ .

3.4.3 Suponha que  $X$  tenha uma distribuição discreta uniforme nos inteiros de 0 a 9. Determine a média, a variância e o desvio-padrão da variável aleatória  $Y = 5X$  e compare-os aos resultados correspondentes para  $X$ .

3.4.4 Medidas de espessura em um processo de recobrimento são feitas com a precisão de centésimo de milímetro. As medidas de espessura estão uniformemente distribuídas, com valores 0,15, 0,16, 0,17, 0,18 e 0,19. Determine a média e a variância da espessura de recobrimento para esse processo.

3.4.5 Suponha que 1000 números de telefones com sete dígitos dentro do seu código de área sejam discados aleatoriamente. Qual é a probabilidade de seu número ser chamado?

3.4.6 Cada questão de múltipla escolha em um exame tem quatro escolhas. Suponha que haja dez questões, e que a escolha seja selecionada aleatoriamente e independentemente para cada questão. Seja  $X$  o número de questões respondidas corretamente.  $X$  tem uma distribuição uniforme discreta? Por que sim ou por que não?

## Exercícios para a Seção 3.5

3.5.1 Para cada cenário descrito (a)-(j), estabeleça se a distribuição binomial é ou não um modelo razoável para a variável e por quê. Estabeleça qualquer suposição que você fizer.

a. Um processo de produção gera milhares de transdutores de temperatura. Seja  $X$  o número de transdutores não conformes em uma amostra de tamanho 30, selecionada ao acaso a partir do processo.

b. De uma batelada de 50 transdutores de temperatura, uma amostra de tamanho 30 é selecionada, sem reposição. Seja  $X$  o número de transdutores não conformes na amostra.

c. Quatro componentes eletrônicos idênticos são ligados a um controlador que pode permutar entre um componente defeituoso e um dos sobressalentes restantes. Seja  $X$  o número de componentes que falharam depois de um período especificado de operação.

d. Seja  $X$  o número de acidentes que ocorrem ao longo de uma autoestrada federal no Arizona, durante o período de um mês.

e. Seja  $X$  o número de respostas corretas dadas por um estudante ao fazer um teste de múltipla escolha, em que ele pode eliminar algumas das opções como incorretas em algumas questões e todas as opções incorretas nas outras questões.

f. Defeitos ocorrem aleatoriamente sobre a superfície de um *chip* semiconductor. Entretanto, somente 80 % dos defeitos podem ser encontrados por meio de testes. Uma amostra de 40 *chips*, cada um com um defeito, é testada. Seja  $X$  o número de *chips* em que o teste encontra um defeito.

g. Erros em um canal digital de comunicação ocorrem em cascatas que afetam vários *bits* consecutivos. Seja  $X$  o número de *bits* com erro em uma transmissão de 100.000 *bits*.

h. Seja  $X$  o número de falhas na superfície de uma grande serpentina de aço galvanizado.

3.5.2 Seja  $X$  uma variável aleatória binomial com  $p = 0,1$  e  $n = 10$ . Calcule as seguintes probabilidades.

- a.  $P(X \leq 2)$     b.  $P(X > 8)$   
c.  $P(X = 4)$     d.  $P(5 \leq X \leq 7)$

3.5.3 A variável aleatória  $X$  tem uma distribuição binomial com  $n = 10$  e  $p = 0,01$ . Determine as seguintes probabilidades:

- a.  $P(X = 5)$     b.  $P(X \leq 2)$   
c.  $P(X \geq 9)$     d.  $P(3 \leq X < 5)$

3.5.4 A variável aleatória  $X$  tem uma distribuição binomial, com  $n = 10$  e  $p = 0,5$ . Esquematize a função de probabilidade de  $X$ .

- a. Qual é o valor mais provável de  $X$ ?  
b. Qual(is) é(são) o(s) valor(es) menos provável(is) de  $X$ ?  
c. Repita os itens prévios com  $p = 0,01$ .

3.5.5 Determine a função de distribuição cumulativa de uma variável aleatória binomial, com  $n = 3$  e  $p = 1/4$ .

3.5.6 Um produto eletrônico contém 40 circuitos integrados. A probabilidade de que qualquer circuito integrado seja defeituoso é de 0,01. Os circuitos integrados são independentes. O produto opera somente se não houver circuitos integrados defeituosos. Qual é a probabilidade de que o produto opere?

3.5.7 As linhas telefônicas em um sistema de reservas de uma companhia aérea estão ocupadas 40 % do tempo. Suponha que sejam independentes os eventos em que as linhas estejam ocupadas em sucessivas ligações. Considere que dez ligações aconteçam para a companhia aérea.

- a. Qual é a probabilidade de que, para exatamente três ligações, as linhas estejam ocupadas?  
b. Qual é a probabilidade de que, para no mínimo uma ligação, as linhas não estejam ocupadas?  
c. Qual é o número esperado de chamadas em que as linhas estejam todas ocupadas?

3.5.8 Um teste de múltipla escolha contém 25 questões, cada uma com quatro respostas. Suponha que um estudante apenas tente adivinhar (“chutar”) a resposta em cada questão.

- a. Qual é a probabilidade de que o estudante responda mais de 20 questões corretamente?  
b. Qual é a probabilidade de que o estudante responda menos de cinco questões corretamente?

3.5.9 As amostras de mitocôndrias rejuvenescidas são mutantes (defeituosas) em 1 % dos casos. Suponha que 15 amostras sejam estudadas e que elas possam ser consideradas independentes para mutação. Determine as seguintes probabilidades:

- a. Nenhuma amostra é mutante.  
b. No máximo uma amostra é mutante.  
c. Mais da metade das amostras é mutante.

3.5.10 Falhas no coração são provenientes de ocorrências naturais (87 %) ou de fatores externos (13 %). Fatores externos são relativos a substâncias induzidas ou a objetos alheios. Ocorrências naturais são causadas por bloqueio arterial, doença e infecção. Suponha que 20 pacientes vão a uma emergência por causa de falha no coração. Suponha que as causas de falha no coração entre os indivíduos sejam independentes.

a. Qual é a probabilidade de três indivíduos terem condições causadas por fatores externos?

b. Qual é a probabilidade de três ou mais indivíduos terem condições causadas por fatores externos?

c. Quais são a média e o desvio-padrão do número de indivíduos com condições causadas por fatores externos?

3.5.11 Um sistema de computadores usa senhas, que são exatamente seis caracteres, sendo cada caractere uma das 26 letras (*a-z*) ou dez inteiros (0-9). Suponha que haja 10.000 usuários do sistema com senhas únicas. Um invasor seleciona aleatoriamente (com reposição) 100.000 senhas do potencial conjunto, e a coincidência com a senha de um usuário é chamada de *êxito*.

- a. Qual é a distribuição do número de êxitos?  
b. Qual é a probabilidade de nenhum êxito?  
c. Quais são a média e o desvio-padrão do número de êxitos?

3.5.12 Amostras de 20 peças de um processo de um corte metálico são selecionadas a cada hora. Tipicamente, 1 % das peças requer retrabalho. Seja  $X$  o número de peças na amostra de 20 que requerem retrabalho. Suspeita-se de um problema no processo se  $X$  exceder sua média em mais de três desvios-padrão.

- a. Se a percentagem de peças que requerem retrabalho permanecer em 1 %, qual será a probabilidade de  $X$  exceder sua média em mais de três desvios-padrão?  
b. Se a percentagem de retrabalho aumentar para 4 %, qual será a probabilidade de  $X$  exceder 1?  
c. Se a percentagem de retrabalho aumentar para 4 %, qual será a probabilidade de  $X$  exceder 1 em, no mínimo, uma das próximas cinco horas de amostragem?

3.5.13 Uma vez que nem todos os passageiros de aviões aparecem na hora do embarque, uma companhia aérea vende 125 bilhetes para um voo que suporta somente 120 passageiros. A probabilidade de que um passageiro não apareça é 0,10 e os passageiros se comportam independentemente.

- a. Qual é a probabilidade de que cada passageiro que aparecer possa embarcar?  
b. Qual é a probabilidade de que o voo decole com assentos vazios?

3.5.14 Este exercício ilustra que a baixa qualidade pode afetar o agendamento e os custos. Um processo de fabricação tem 100 pedidos de consumidores para preencher. Cada pedido requer uma peça componente que é comprada de um fornecedor. No entanto, tipicamente, 2 % dos componentes são identificados como defeituosos, podendo os componentes ser considerados independentes.

- a. Se o fabricante estocar 100 componentes, qual será a probabilidade de que os 100 pedidos possam ser preenchidos sem refazer a solicitação dos componentes?  
b. Se o fabricante estocar 102 componentes, qual será a probabilidade de que os 100 pedidos possam ser preenchidos sem refazer a solicitação dos componentes?  
c. Se o fabricante estocar 105 componentes, qual será a probabilidade de que os 100 pedidos possam ser preenchidos sem refazer a solicitação dos componentes?

3.5.15 Considere os períodos de permanência em emergência de hospital no Exercício 3.1.19. Considere que cinco pessoas chegam independentemente para atendimento.

- a. Qual é a probabilidade de que a permanência de exatamente uma pessoa seja menor do que ou igual a quatro horas?  
b. Qual é a probabilidade de que exatamente duas pessoas esperem mais de quatro horas?  
c. Qual é a probabilidade de que no mínimo uma pessoa espere mais de quatro horas?

**3.5.16** A probabilidade de que um visitante de um *site* na internet forneça dados de contato para informações adicionais é de 0,01. Considere que 1000 visitantes da página se comportem independentemente. Determine as seguintes probabilidades:

- Nenhum visitante fornece dados de contato.
- Exatamente dez visitantes fornecem dados de contato.
- Mais de três visitantes fornecem dados de contato.

**3.5.17** Considere o tempo de recarga do *flash* em câmeras de celulares, como no Exemplo 3.1. Suponha que a probabilidade de a câmera passar no teste seja de 0,8 e as câmeras atuem de forma independente. Qual é o menor tamanho de amostra necessário, de modo que a probabilidade de pelo menos uma câmera falhar seja, no mínimo, de 95 %?

## Exercícios para a Seção 3.6

**3.6.1** Suponha que a variável aleatória  $X$  tenha uma distribuição geométrica, com  $p = 0,5$ . Determine as seguintes probabilidades:

- $P(X = 1)$
- $P(X = 4)$
- $P(X = 8)$
- $P(X \leq 2)$
- $P(X > 2)$

**3.6.2** Suponha que  $X$  seja uma variável aleatória binomial negativa, com  $p = 0,2$  e  $r = 4$ . Determine o seguinte:

- $E(X)$
- $P(X = 20)$
- $P(X = 19)$
- $P(X = 21)$

e. O valor mais provável de  $X$

**3.6.3** Considere a sequência de tentativas independentes de Bernoulli, com  $p = 0,2$ .

- Qual é o número esperado de tentativas, de modo a se obter o primeiro sucesso?
- Depois de oito sucessos ocorrerem, qual é o número esperado de tentativas, de modo a se obter o nono sucesso?

**3.6.4** Em um estudo clínico, voluntários são testados em relação a um gene que aumentou o risco de uma doença. A probabilidade de uma pessoa carregar o gene é igual a 0,1. As pessoas são supostamente independentes em relação ao gene.

- Qual é a probabilidade de exatamente quatro pessoas terem de ser testadas para detectar duas com o gene?
- Qual é a probabilidade de quatro ou mais pessoas terem de ser testadas para detectar duas com o gene?
- Qual o número esperado de pessoas que devem ser testadas para detectar duas com o gene?

**3.6.5** Suponha que cada uma das suas ligações para uma estação popular de rádio tenha probabilidade de 0,02 de se completar; ou seja, de não obter um sinal de ocupado. Considere que suas ligações sejam independentes.

- Qual é a probabilidade de que sua primeira ligação a se completar seja sua décima tentativa?
- Qual é a probabilidade de se necessitar mais de cinco ligações para que você se conecte?
- Qual é o número médio necessário de ligações para que você se conecte?

**3.6.6** Um jogador de videogame é confrontado com uma série de oponentes e tem uma probabilidade de 80 % de derrotar cada um deles. O sucesso com cada oponente é independente de confrontos prévios. O jogador continua a enfrentar oponentes até ser derrotado.

- Qual é a função de probabilidade do número de oponentes enfrentados em um jogo?
- Qual é a probabilidade de o jogador derrotar no mínimo dois oponentes em um jogo?
- Qual é o número esperado de oponentes enfrentados em um jogo?
- Qual é a probabilidade de o jogador enfrentar quatro ou mais oponentes em um jogo?
- Qual é o número esperado de jogos até que um jogador enfrente quatro ou mais oponentes?

**3.6.7** Falha no coração é causada por ocorrências naturais (87 %) ou por fatores externos (13 %). Fatores externos são relativos a substâncias induzidas ou a objetos alheios. Ocorrências naturais são causadas por bloqueio arterial, doença e infecção. Suponha que as causas de falha no coração entre os indivíduos sejam independentes.

- Qual é a probabilidade de o primeiro paciente com falha no coração que entre na emergência ter a condição por causa de fatores externos?
- Qual é a probabilidade de o terceiro paciente com falha no coração que entre na emergência ser o primeiro por causa de fatores externos?
- Qual é o número médio de pacientes, com falha no coração por causas naturais, que entram na emergência antes de o primeiro paciente com falha no coração causada por fatores externos entrar?

**3.6.8** Suponha que 20 peças sejam verificadas a cada hora e que  $X$  seja o número de peças na amostra de 20 que requeiram retrabalho. Peças são consideradas independentes com relação a retrabalho.

- Se a porcentagem de peças que requerem conserto permanecer 1 %, qual será a probabilidade de que na décima hora apareça a primeira amostra em que  $X$  excede 1?
- Se a porcentagem de conserto aumentar para 4 %, qual será a probabilidade de que na décima hora apareça a primeira amostra em que  $X$  excede 1?
- Se a porcentagem de retrabalho aumentar para 4 %, qual será o número esperado de horas até que  $X$  exceda 1?

**3.6.9** Uma companhia de comércio tem oito computadores que ela usa para negociar na Bolsa de Nova York. A probabilidade de um computador falhar em um dia é igual a 0,005, e os computadores falham independentemente. Computadores são reparados à noite e cada dia é uma tentativa independente.

- Qual é a probabilidade de todos os oito computadores falharem em um dia?
- Qual é o número médio de dias até que um computador específico falhe?
- Qual é o número médio de dias até que todos os oito computadores falhem no mesmo dia?

**3.6.10** No processo de meiose, uma célula diploide simples passa por oito fases diferentes. No entanto, somente 60 % dos processos passam pelas seis primeiras fases e somente 40 % passam por todas as oito fases. Considere que os resultados de cada fase sejam independentes.

- Se a probabilidade de uma passagem com sucesso de cada uma das seis primeiras fases for constante, qual será a probabilidade de uma passagem com sucesso de uma única dessas fases?
- Se a probabilidade de uma passagem com sucesso de cada uma das duas últimas fases for constante, qual será a probabilidade de uma passagem com sucesso de uma única dessas fases?

3.6.11 Todo dia, uma página na internet seleciona aleatoriamente dez produtos para oferecer desconto. A impressora colorida de seu interesse está hoje com desconto.

- Qual é o número esperado de dias até que esse produto tenha novamente desconto?
- Qual é a probabilidade de que esse produto seja o primeiro a ter desconto novamente, exatamente daqui a dez dias?
- Se o produto não tiver desconto nos próximos cinco dias, qual será a probabilidade de que seja o primeiro a ter desconto novamente daqui a 15 dias?
- Qual é a probabilidade de que esse produto seja o primeiro a ter desconto novamente dentro de três ou menos dias?

3.6.12 Considere as pessoas que saem de um hospital sem serem atendidas por médicos (PNEM) da emergência no Exemplo 2.6. Suponha que as pessoas cheguem independentemente ao serviço do Hospital 1.

- Qual é a probabilidade de que a quinta pessoa seja a primeira a ser considerada como PNEM?
- Qual é a probabilidade de que a quinta ou a sexta pessoa seja a primeira a ser considerada como PNEM?
- Qual é a probabilidade de que a primeira pessoa considerada como PNEM esteja entre as quatro primeiras pessoas?
- Qual é o número esperado de atendimentos até que a terceira pessoa considerada como PNEM ocorra?

3.6.13 Considere o tempo de recarga do *flash* em câmeras de celulares, como no Exemplo 3.1. Admita que a probabilidade de uma câmera

passar no teste seja de 0,8 e que as câmeras atuem de forma independente. Determine o seguinte:

- A probabilidade de que a segunda falha ocorra na décima câmera testada.
- A probabilidade de que a segunda falha ocorra nos testes de quatro ou menos câmeras.
- O número esperado de câmeras testadas para obter a terceira falha.

3.6.14 Consumidores visitam um *site* da internet e existe probabilidade de 0,01 de uma compra ser realizada, se um consumidor vir cinco ou menos páginas. Entretanto, se um consumidor vir mais de cinco páginas, a probabilidade de uma compra é de 0,1. A probabilidade de um consumidor ver cinco ou mais páginas é de 0,25. Os consumidores se comportam de forma independente.

- O número de consumidores que visitam o *site* até que uma compra seja realizada é uma variável aleatória geométrica? Por que sim ou por que não?
- Qual é a probabilidade de que a primeira compra realizada seja do décimo consumidor a visitar o *site*?

3.6.15 Um arranjo de 30 lâmpadas de LED é usado nas luzes de automóveis. A probabilidade de uma lâmpada ser defeituosa é de 0,001 e lâmpadas defeituosas ocorrem independentemente. Determine o seguinte:

- A probabilidade de que uma luz automotiva tenha uma ou mais lâmpadas defeituosas.
- O número esperado de luzes automotivas de modo a se obter uma com duas ou mais lâmpadas defeituosas.

## Exercícios para a Seção 3.7

3.7.1 Suponha que  $X$  tenha uma distribuição hipergeométrica com  $N = 100$ ,  $n = 4$  e  $K = 20$ . Determine o seguinte:

- $P(X = 1)$
- $P(X = 6)$
- $P(X = 4)$
- Média e a variância de  $X$

3.7.2 Suponha que  $X$  tenha uma distribuição hipergeométrica com  $N = 10$ ,  $n = 3$  e  $K = 4$ . Esquematize a função de probabilidade de  $X$ . Determine a função de distribuição cumulativa para  $X$ .

3.7.3 Um estudo usa 800 homens com menos de 55 anos. Suponha que 30 % carreguem um marcador no cromossomo masculino, que indique um risco crescente de pressão sanguínea alta.

- Se dez homens forem selecionados aleatoriamente e testados em relação ao marcador, qual será a probabilidade de exatamente um homem ter esse marcador?
- Se dez homens forem selecionados aleatoriamente e testados em relação ao marcador, qual será a probabilidade de mais de um homem ter esse marcador?

3.7.4 Cartões de circuito integrado são verificados em um teste funcional depois de serem preenchidos com *chips* semicondutores. Um lote contém 140 cartões, e 20 são selecionados sem reposição para o teste funcional.

- Se 20 cartões forem defeituosos, qual será a probabilidade de no mínimo um cartão defeituoso estar na amostra?
- Se cinco cartões forem defeituosos, qual será a probabilidade de no mínimo um cartão defeituoso aparecer na amostra?

3.7.5 A análise dos resultados de um experimento de transmutação de uma folha (a folha se transforma em uma pétala) é resumida pelo tipo de transformação completada:

### Transformação Total da Textura

		Sim	Não
Transformação da cor total	Sim	243	26
	Não	13	18

Um naturalista seleciona aleatoriamente, sem reposição, três folhas desse conjunto. Determine as seguintes probabilidades:

- Exatamente uma folha sofreu ambos os tipos de transformação.
- No mínimo uma folha sofreu ambos os tipos de transformação.
- Exatamente uma folha sofreu um tipo, mas não ambos os tipos de transformação.
- Pelo menos uma folha sofreu no mínimo um tipo de transformação.

3.7.6 Um estado tem uma loteria em que seis números são selecionados aleatoriamente de 40, sem reposição. Um jogador escolhe seis números antes de a amostra da loteria ser selecionada.

- Qual é a probabilidade de que os seis números escolhidos pelo jogador coincidam com todos os seis números do resultado da loteria?
- Qual é a probabilidade de que cinco dos seis números escolhidos pelo jogador apareçam no resultado da loteria?
- Qual é a probabilidade de que quatro dos seis números escolhidos pelo jogador apareçam no resultado da loteria?
- Se um jogador for a uma loteria toda semana, qual é o número esperado de semanas até que o jogador ganhe na loteria?

3.7.7 Uma talhadeira contém 48 lâminas. Cinco lâminas são selecionadas ao acaso e avaliadas a cada dia em relação ao afiamento. Se alguma lâmina não afiada for encontrada, o arranjo será trocado por um novo conjunto de lâminas afiadas.

a. Se dez das lâminas em uma talhadeira não estiverem afiadas, qual será a probabilidade de que a talhadeira seja trocada no primeiro dia de sua avaliação?

b. Se dez das lâminas em uma talhadeira não estiverem afiadas, qual será a probabilidade de que a talhadeira não seja trocada até o terceiro dia de avaliação? [Sugestão: suponha que as decisões diárias sejam independentes. Use a distribuição geométrica.]

c. Considere que no primeiro dia de avaliação duas das lâminas não estejam afiadas; no segundo dia de avaliação, seis não estejam afiadas; e no terceiro dia de avaliação, dez não estejam afiadas. Qual é a probabilidade de que a talhadeira não seja trocada até o terceiro dia de avaliação? [Sugestão: suponha que as decisões diárias sejam independentes. No entanto, a probabilidade de troca muda a cada dia.]

3.7.8 a. Para o Exercício 3.7.1, calcule  $P(X = 1)$  e  $P(X = 4)$ , considerando que  $X$  tenha uma distribuição binomial, e compare esses resultados aos resultados derivados da distribuição hipergeométrica.

b. Use a aproximação da distribuição binomial para a distribuição hipergeométrica para aproximar as probabilidades no Exercício 3.7.4. Qual é a correção para a população finita nesse exercício?

3.7.9 Suponha que uma companhia de seguro-saúde selecione 20 pacientes aleatoriamente (sem reposição) entre 500 para avaliar a adesão a um horário de medicação. Suponha que 10 % dos 500 pacientes falhem nessa adesão ao horário. Determine o seguinte:

a. A probabilidade de que exatamente 10 % dos pacientes na amostra falhem em aderir.

b. A probabilidade de que menos de 10 % dos pacientes na amostra falhem em aderir.

c. A probabilidade de que mais de 10 % dos pacientes na amostra falhem em aderir.

d. Média e variância do número de pacientes na amostra que falharam em aderir.

3.7.10 Uma companhia de utilidades pode oferecer taxas elétricas, baseadas no consumo diário, de modo a aumentar o pico de demanda em um dia. Um número suficiente de clientes precisa aceitar o plano para que haja sucesso. Suponha que, entre os 50 principais clientes, 15 aceitassem o plano. A companhia seleciona aleatoriamente (sem reposição) os dez maiores clientes para entrar em contato e promover o plano.

a. Qual é a probabilidade de que exatamente dois dos principais clientes selecionados aceitem o plano?

b. Qual é a probabilidade de que no mínimo um dos clientes selecionados aceite o plano?

c. Em vez de 15 consumidores, qual é o número mínimo de principais clientes que necessitariam aceitar o plano de modo a encontrar o objetivo a seguir? A probabilidade de que no mínimo um cliente principal selecionado aceite o plano é maior do que ou igual a 0,95.

3.7.11 Suponha que lesões estejam presentes em cinco locais entre 50 em um paciente. Uma biópsia seleciona aleatoriamente oito locais (sem reposição).

a. Qual é a probabilidade de que lesões estejam presentes em, no mínimo, um local selecionado?

b. Qual é a probabilidade de que lesões estejam presentes em dois ou mais locais selecionados?

c. Em vez de oito locais, qual é o número mínimo de locais que necessitam ser selecionados para encontrar o objetivo seguinte? A probabilidade de que no mínimo um local tenha lesões presentes é maior do que ou igual a 0,9.

## Exercícios para a Seção 3.8

3.8.1 Suponha que  $X$  tenha uma distribuição de Poisson. Determine as seguintes probabilidades quando a média de  $X$  for quatro, e repita quando a média for 0,4:

a.  $P(X = 0)$                       b.  $P(X \leq 2)$

c.  $P(X = 4)$                       d.  $P(X = 8)$

3.8.2 O número de chamadas telefônicas que chegam a uma central é frequentemente modelado como uma variável aleatória de Poisson. Considere que, em média, haja dez chamadas por hora. Determine as seguintes probabilidades:

a. Exatamente cinco chamadas em uma hora.

b. Três ou menos chamadas em uma hora.

c. Exatamente 15 chamadas em duas horas.

d. Exatamente cinco chamadas em 30 minutos.

3.8.3 Suponha que o número de clientes que entrem em um banco em uma hora seja uma variável aleatória de Poisson. Suponha também que  $P(X = 0) = 0,05$ . Determine a média e a variância de  $X$ .

3.8.4 Dados provenientes de [www.centralhudsonlabs.com](http://www.centralhudsonlabs.com) determinaram que o número médio de fragmentos de insetos em uma barra de 225 gramas de chocolate foi igual a 14,4; porém, três marcas apresentaram contaminação por inseto maior que duas vezes a média. Veja o U.S. Food and Drug Administration – Center for Food Safety and Applied Nutrition for Defect Action Levels para

produtos alimentícios. Considere que o número de fragmentos (contaminantes) siga a distribuição de Poisson.

a. Se você consumir uma barra de 225 gramas de uma marca com um nível médio de contaminação, qual será a probabilidade de não ocorrer contaminação por insetos?

b. Suponha que você consuma uma barra que tem um quinto do tamanho testado (45 gramas) de uma marca com um nível médio de contaminação. Qual é a probabilidade de não ocorrer contaminação por insetos?

c. Se você consumir sete barras de 28,35 gramas, essa semana, de uma marca com um nível médio de contaminação, qual será a probabilidade de você consumir um ou mais fragmentos de inseto em mais de uma barra?

d. A probabilidade de uma contaminação ser maior do que duas vezes a média de 14,4 não é usual, ou pode ser considerada uma variação típica? Explique.

3.8.5 Astrônomos tratam o número de estrelas em determinado volume do espaço como uma variável aleatória de Poisson. A densidade na Via Láctea, na vizinhança de nosso sistema solar, é uma estrela por 16 anos-luz<sup>3</sup>.

a. Qual é a probabilidade de duas ou mais estrelas em 16 anos-luz<sup>3</sup>?

b. Quantos anos-luz<sup>3</sup> de espaço têm de ser estudados de modo que a probabilidade de uma ou mais estrelas exceda(m) 0,95?

3.8.6 Em uma seção de autoestrada, o número de buracos, que é bastante significativo para requerer reparo, pressupõe seguir uma distribuição de Poisson, com uma média de dois buracos por milha.

- Qual é a probabilidade de que não haja buracos que requeiram reparo em cinco milhas de autoestrada?
- Qual é a probabilidade de que no mínimo um buraco requeira reparo em 0,5 milha de autoestrada?
- Se o número de buracos estiver relacionado com a carga do veículo na autoestrada e algumas seções dessa autoestrada estiverem sujeitas a uma carga pesada de veículos, enquanto outras seções estiverem sujeitas a uma carga leve de veículos, como você se sente a respeito da suposição de distribuição de Poisson para o número de buracos que requerem reparo?

3.8.7 Em 1898, L. J. Bortkiewicz publicou um livro intitulado *The Law of Small Numbers* e empregou dados coletados ao longo de 20 anos, para mostrar que o número de soldados mortos por coices de cavalo em cada ano em cada corporação na cavalaria prussiana seguia uma distribuição de Poisson com média de 0,61.

- Qual é a probabilidade de mais de uma morte na corporação em um ano?
- Qual é a probabilidade de nenhuma morte na corporação ao longo de cinco anos?

3.8.8 O número de visualizações a uma página da internet segue a distribuição de Poisson, com uma média de 1,5 por minuto.

- Qual é a probabilidade de nenhuma visualização em um minuto?
- Qual é a probabilidade de duas ou menos visualizações em dez minutos?
- A resposta do item anterior depende do fato de o período de dez minutos ser um intervalo ininterrupto? Explique.
- Determine o intervalo de tempo de modo que a probabilidade de nenhuma visualização nesse intervalo seja de 0,001.

3.8.9 O número de falhas na superfície de painéis de plástico, usados no interior de automóveis, tem uma distribuição de Poisson, com uma média de 0,05 falha por pé quadrado de painel plástico. Considere que o interior de um automóvel contenha dez pés quadrados de painel plástico.

- Qual é a probabilidade de não haver falha na superfície do interior do automóvel?
- Se dez carros forem vendidos para uma companhia de aluguel de carros, qual será a probabilidade de que nenhum dos dez carros tenha qualquer falha na superfície?

c. Se dez carros forem vendidos para uma companhia de aluguel de carros, qual será a probabilidade de no máximo um carro ter qualquer falha na superfície?

3.8.10 *Inclusões* são defeitos em um metal vertido causados por contaminantes. O número de inclusões (grandes) em um ferro fundido segue uma distribuição de Poisson, com uma média de 2,5 por milímetro cúbico. Determine o seguinte:

- A probabilidade de no mínimo uma inclusão em um milímetro cúbico.
- A probabilidade de no mínimo cinco inclusões em 5,0 milímetros cúbicos.
- O volume de material para inspecionar, de modo que a probabilidade de existir pelo menos uma inclusão seja igual a 0,99.
- Em vez de uma média de 2,5 por milímetro cúbico, determine as inclusões médias por milímetro cúbico, de modo que a probabilidade de existir no mínimo uma inclusão seja igual a 0,95.

3.8.11 Pontos de táxis obedecem a um processo de Poisson, com uma média de cinco táxis por hora. Suponha que você saia do ponto às 18 horas. Determine o seguinte:

- A probabilidade de que você espere mais de dez minutos por um táxi.
- A probabilidade de que você espere menos de 20 minutos por um táxi.
- O número médio de táxis por hora, de modo que a probabilidade de espera de mais de dez minutos seja igual a 0,1.

3.8.12 O artigo "An Association Between Fine Particles and Asthma Emergency Department Visits for Children in Seattle" [*Environmental Health Perspectives*, junho 1999, Vol. 107(6)] usou modelos de Poisson para o número de atendimentos diários ao Departamento de Aasma (DA). Para os CEPs (Código de Endereçamento Postal) estudados, o número médio de atendimentos ao DA foi de 1,8 por dia. Determine o seguinte:

- Probabilidade de mais de cinco atendimentos em um dia.
- Probabilidade de menos de cinco atendimentos em uma semana.
- Número de dias, de modo que a probabilidade de existir pelo menos um atendimento seja igual a 0,99.
- Em vez de uma média de 1,8 por dia, determine o número médio de atendimentos diários, de modo que a probabilidade de haver mais de cinco atendimentos em um dia seja igual a 0,1.

## Exercícios Suplementares para o Capítulo 3

3.S13 Seja a variável aleatória  $X$  igualmente provável de ter valores iguais a  $1/8$ ,  $1/4$  e  $3/8$ . Determine a média e a variância de  $X$ .

3.S14 Seja  $X$  o número de *bits* recebidos com erro em um canal digital de comunicação e considere que  $X$  seja uma variável aleatória binomial com  $p = 0,001$ . Se 1000 *bits* forem transmitidos, determine o seguinte:

- $P(X = 1)$
- $P(X \geq 1)$
- $P(X \leq 2)$
- Média e variância de  $X$

3.S15 Um carregador automático de cartelas de ovos tem probabilidade de 1 % de quebrar um ovo, e o consumidor reclamará se mais de um ovo por dúzia for quebrado. Considere que cada carregamento de ovo seja um evento independente.

- Qual é a distribuição de ovos quebrados por dúzia? Inclua os valores dos parâmetros.
- Qual é a probabilidade de uma cartela de uma dúzia de ovos resultar em reclamação?

c. Quais são a média e o desvio-padrão do número de ovos quebrados em uma cartela de uma dúzia?

3.S16 Uma rede congestionada de computadores tem 1 % de chance de perder um bloco de dados que devem ser reenviados, e perdas de blocos são eventos independentes. Uma mensagem de *e-mail* requer 100 blocos.

- Qual é a distribuição de blocos de dados que devem ser reenviados? Inclua os valores dos parâmetros.
- Qual é a probabilidade de no mínimo um bloco ser reenviado?
- Qual é a probabilidade de dois ou mais blocos serem reenviados?
- Quais são a média e o desvio-padrão do número de blocos que são reenviados?
- Se há dez mensagens e cada uma contém 100 blocos, qual é a probabilidade de no mínimo uma mensagem requerer que dois ou mais blocos sejam reenviados?

3.S17 Em seu caminho matinal, você se aproxima de determinado sinal de trânsito, que está verde 20 % do tempo. Suponha que cada manhã represente uma tentativa independente.

- Qual é a probabilidade de na primeira manhã em que o sinal esteja verde ser a quarta manhã em que você se aproxime dele?
- Qual é a probabilidade de a luz não estar verde por dez manhãs consecutivas?

3.S18 A probabilidade de uma águia matar um coelho em um dia de caça é 10 %. Considere que resultados sejam independentes entre dias.

- Qual é a distribuição do número de dias até que a caça ao coelho tenha sucesso?
- Qual é a probabilidade de a águia ter de esperar cinco dias para a sua primeira caçada de sucesso?
- Qual é o número esperado de dias até que a caçada tenha sucesso?
- Se a águia puder sobreviver até dez dias sem alimento (isso requer uma caçada com sucesso no décimo dia), qual é a probabilidade de a águia ainda estar viva dez dias a partir de agora?

3.S19 Uma balança eletrônica, em uma operação automatizada de enchimento, interrompe a linha de produção depois que três embalagens abaixo do peso sejam detectadas. Suponha que a probabilidade de uma embalagem abaixo do peso seja de 0,001, e que cada enchimento seja independente.

- Qual é o número médio de enchimentos antes de a linha ser interrompida?
- Qual é o desvio-padrão do número de enchimentos antes de a linha parar?

3.S20 Um carregamento de compostos químicos chega em 15 depósitos. Três deles são selecionados ao acaso e sem reposição para uma inspeção de pureza. Se dois dos depósitos não satisfizerem os requerimentos de pureza, qual será a probabilidade de no mínimo um dos depósitos não conformes ser selecionado na amostra?

3.S21 Tráfego de carros é tradicionalmente modelado como uma distribuição de Poisson. Um engenheiro de tráfego monitora o fluxo de carros em um cruzamento que tem uma média de seis carros por minuto. Para estabelecer o tempo de um sinal, as seguintes probabilidades são usadas.

- Qual é a probabilidade de nenhum carro passar pelo cruzamento em 30 segundos?
- Qual é a probabilidade de três ou mais carros passarem pelo cruzamento em 30 segundos?
- Calcule o número mínimo de carros que passam pelo cruzamento, de modo que a probabilidade desse número ou menos de carros em 30 segundos seja, no mínimo, de 90 %.
- Se a variância do número de carros que passam pelo cruzamento por minuto for igual a 20, a distribuição de Poisson será apropriada? Explique.

3.S22 O número de mensagens enviadas para um boletim em um computador é uma variável aleatória de Poisson, com uma média de cinco mensagens por hora.

- Qual é a probabilidade de cinco mensagens chegarem em uma hora?
- Qual é a probabilidade de dez mensagens chegarem em 1,5 hora?
- Qual é a probabilidade de menos de duas mensagens chegarem em meia hora?

3.S23 A probabilidade de que sua ligação para uma linha de serviço seja respondida em menos de 30 segundos é de 0,75. Suponha que suas ligações sejam independentes.

- Se você ligar dez vezes, qual será a probabilidade de que exatamente nove de suas ligações sejam respondidas dentro de 30 segundos?

b. Se você ligar 20 vezes, qual será a probabilidade de que no mínimo 16 ligações sejam respondidas em menos de 30 segundos?

c. Se você ligar 20 vezes, qual será o número médio de ligações que serão respondidas em menos de 30 segundos?

d. Qual é a probabilidade de você ter de ligar quatro vezes para obter a primeira resposta em menos de 30 segundos?

e. Qual é o número médio de ligações até que você receba uma resposta em menos de 30 segundos?

3.S24 O número de erros em um livro-texto segue uma distribuição de Poisson, com uma média de 0,01 erro por página. Qual a probabilidade de haver três ou menos erros em 100 páginas?

3.S25 A probabilidade de um indivíduo se recuperar de uma doença em um período de uma semana, sem tratamento, é de 0,1. Suponha que 20 indivíduos independentes, sofrendo dessa doença, sejam tratados com uma droga e quatro se recuperem em um período de uma semana. Se a droga não tiver efeito, qual será a probabilidade de quatro ou mais pessoas se recuperarem no período de uma semana?

3.S26 A resposta de um paciente a um medicamento genérico para controlar dor é pontuada em uma escala de cinco pontos, em que o 5 indica alívio completo. Historicamente, a distribuição de pontos é

1	2	3	4	5
0,05	0,1	0,2	0,25	0,4

Dois pacientes, considerados independentes, são pontuados.

- Qual é a função de probabilidade da pontuação total?
- Qual é a função de probabilidade da pontuação média?

3.S27 Em um processo de fabricação que lamina várias camadas de cerâmica, 1 % dos arranjos tem defeitos. Considere que os arranjos sejam independentes.

- Qual é o número médio de arranjos que precisam ser verificados, de modo a se obterem cinco arranjos com defeitos?
- Qual é o desvio-padrão do número de arranjos que necessitam ser verificados, de modo a se obterem cinco arranjos com defeitos?
- Determine o número mínimo de arranjos que necessitam ser verificados, de modo que a probabilidade de no mínimo um arranjo com defeito exceda 0,95.

3.S28 Um fabricante de produtos eletrodomésticos espera que 2 % das unidades falhem durante o período de garantia. Uma amostra de 500 unidades independentes é rastreada para desempenho de garantia.

- Qual é a probabilidade de nenhuma unidade falhar durante o período de garantia?
- Qual é o número esperado de falhas durante o período de garantia?
- Qual é a probabilidade de mais de duas unidades falharem durante o período de garantia?

3.S29 Determine a função de probabilidade para a variável aleatória com a seguinte função de distribuição cumulativa:

$$F(x) = \begin{cases} 0 & x < 2 \\ 0,2 & 2 \leq x < 5,7 \\ 0,5 & 5,7 \leq x < 6,5 \\ 0,8 & 6,5 \leq x < 8,5 \\ 1 & 8,5 \leq x \end{cases}$$

3.S30 A variável aleatória  $X$  tem a seguinte distribuição de probabilidades:

Probabilidade	2	3	5	8
$x$	0,2	0,4	0,3	0,1

Determine o seguinte:

- $P(X \leq 3)$
- $P(X > 2,5)$
- $P(2,7 < X < 5,1)$
- $E(X)$
- $V(X)$

## 22 Exercícios Complementares

3.S31 Considere que o número de erros ao longo de uma superfície magnética gravadora seja uma variável aleatória de Poisson, com média de um erro a cada  $10^5$  bits. Um setor de dados consiste em 4096 bytes de oito bits.

- Qual é a probabilidade de mais de um erro em um setor?
- Qual é o número médio de setores até que um erro seja encontrado?

3.S32 Cada cápsula principal do mancal em um motor contém quatro parafusos. Esses parafusos são selecionados, ao acaso e sem reposição, de peças que contêm 30 parafusos de um fornecedor e 70 parafusos de outro.

- Qual é a probabilidade de que a cápsula principal contenha todos os parafusos provenientes do mesmo fornecedor?
- Qual é a probabilidade de que exatamente três parafusos sejam provenientes do mesmo fornecedor?

3.S33 De 500 consumidores, um grande fabricante de aparelhos eletrônicos selecionará aleatoriamente uma amostra sem reposição. A companhia estima que 25 % dos consumidores responderão à pesquisa. Se essa estimativa for correta, qual será a função de probabilidade do número de consumidores que responderão?

- Considere que a companhia amostra cinco consumidores.
- Considere que a companhia amostra dez consumidores.

3.S34 Um técnico de instalação de um sistema especializado de comunicação é enviado para uma cidade, somente quando existirem três ou mais ordens de serviço. Suponha que as ordens de serviço sigam

a distribuição de Poisson, com uma média de 0,25 por semana, para uma cidade com uma população de 100.000. Considere que sua cidade contém uma população de 800.000.

- Qual é a probabilidade de um técnico ser requisitado depois de um período de uma semana?
- Se você for o primeiro na cidade a solicitar uma ordem de serviço, qual será a probabilidade de você ter de esperar mais de duas semanas, a partir do tempo da solicitação da ordem de serviço, até que o técnico seja enviado?

3.S35 *Saguaro cacti* são cactos indígenas do sudoeste dos Estados Unidos e do México. Admita que o número de cactos em uma região siga uma distribuição de Poisson com uma média de 280 por quilômetro quadrado. Determine o seguinte:

- Número médio de cactos por 10.000 metros quadrados.
- A probabilidade de nenhum cacto em 10.000 metros quadrados.
- A área de uma região tal que a probabilidade de haver pelo menos dois cactos na região seja de 0,9.

3.S36 Suspeita-se que alguns dos reservatórios de produtos químicos comprados de um fornecedor excedam o conteúdo-padrão de umidade. Considere os reservatórios como independentes. Determine a proporção de reservatórios provenientes do fornecedor que têm de exceder o teor-padrão de umidade, de modo que a probabilidade seja 0,90 de, no mínimo, um reservatório falhar no teste em uma amostra com 30 reservatórios.

# Capítulo 4 Exercícios

## Exercícios para a Seção 4.1

4.1.1 Suponha que  $f(x) = e^{-x}$  para  $0 < x$ . Determine o seguinte:

- $P(1 < X)$
  - $P(1 < X < 2,5)$
  - $P(X = 3)$
  - $P(X < 4)$
  - $P(3 \leq X)$
- f. Determine  $x$  tal que  $P(x < X) = 0,10$ .  
g. Determine  $x$  tal que  $P(X \leq x) = 0,10$ .

4.1.2 O diâmetro de uma partícula de contaminação (em micrômetros) é modelado com a função densidade de probabilidade  $f(x) = 2/x^3$  para  $x > 1$ . Determine o seguinte:

- $P(X < 2)$
- $P(X > 5)$
- $P(4 < X < 8)$
- $P(X < 4 \text{ ou } X > 8)$
- $x$  tal que  $P(X < x) = 0,95$

4.1.3 Suponha que  $f(x) = 0,5 \cos x$  para  $-\pi/2 < x < \pi/2$ . Determine o seguinte:

- $P(X < 0)$
- $P(X < -\pi/4)$
- $P(-\pi/4 < X < \pi/4)$
- $P(X > -\pi/4)$
- $x$  tal que  $P(X < x) = 0,95$

4.1.4 Suponha que  $f(x) = x/8$  para  $3 < x < 5$ . Determine as seguintes probabilidades:

- $P(X < 4)$
- $P(X > 3,5)$
- $P(4 < X < 5)$
- $P(X < 4,5)$
- $P(X < 3,5 \text{ ou } X > 4,5)$

4.1.5 Suponha que  $f(x) = 1,5x^2$  para  $-1 < x < 1$ . Determine o seguinte:

- $P(0 < X)$
- $P(0,5 < X)$
- $P(-0,5 \leq X \leq 0,5)$
- $P(X < -2)$
- $P(X < 0 \text{ ou } X > -0,5)$
- $x$  tal que  $P(x < X) = 0,05$ .

4.1.6 A função densidade de probabilidade do peso líquido, em libras, de um pacote de herbicida químico é  $f(x) = 2,0$ , para  $49,75 < x < 50,25$  libras.

- Determine a probabilidade de um pacote pesar mais de 50 libras.
- Quanto herbicida químico está contido em 90 % de todos os pacotes?

4.1.7 A função densidade de probabilidade do comprimento de um bastão metálico é  $f(x) = 2$  para  $2,3 < x < 2,8$  metros.

- Se as especificações para esse processo forem de 2,25 a 2,75 metros, que proporção das barras não se ajustará às especificações?
- Suponha que a função densidade de probabilidade seja  $f(x) = 2$ , para um intervalo de comprimento de 0,5 metro. Sobre qual valor a densidade deveria ser centrada, de modo a atingir a maior proporção de bastões dentro das especificações?

4.1.8 Um artigo em *Electric Power Systems Research* ["Modeling Real-Time Balancing Power Demands in Wind Power Systems Using Stochastic Differential Equations" (2010, Vol. 80(8), pp. 966-974)] considerou um novo modelo probabilístico para equilibrar a demanda de energia com grande quantidade de energia eólica. Nesse modelo, a perda de energia causada por desligamentos é considerada ter uma distribuição triangular com função densidade de probabilidade

$$f(x) = \begin{cases} -5,56 \times 10^{-4} + 5,56 \times 10^{-6}x, & x \in [100, 500] \\ 4,44 \times 10^{-3} - 4,44 \times 10^{-6}x, & x \in [500, 1000] \\ 0, & \text{caso contrário} \end{cases}$$

Determine o seguinte:

- $P(X < 90)$
- $P(100 < X \leq 200)$
- $P(X > 800)$
- Valor excedido com probabilidade de 0,1.

4.1.9 O tempo (em horas) de espera para o atendimento em uma emergência de hospital segue uma distribuição cuja função de densidade de probabilidade é igual a  $f(x) = 0,5\exp(-0,5x)$  para  $0 < x$ . Determine o seguinte:

- a.  $P(X < 0,5)$     b.  $P(X > 2)$   
c. O valor de  $x$  (em horas) excedido com probabilidade igual a 0,05.

4.1.10 A distribuição de  $X$  é aproximada por uma função densidade de probabilidade triangular  $f(x) = 0,0025x - 0,075$  para  $30 < x < 50$  e  $f(x) = -0,0025x + 0,175$  para  $50 < x < 70$ . Determine o seguinte:

- a.  $P(X \leq 40)$     b.  $P(40 < X \leq 60)$   
c. O valor de  $x$  excedido com probabilidade igual a 0,99.

## Exercícios para a Seção 4.2

4.2.1 Suponha que a função de distribuição cumulativa da variável aleatória  $X$  seja

$$F(x) = \begin{cases} 0 & x < 0 \\ 0,25x & 0 \leq x < 5 \\ 1 & 5 \leq x \end{cases}$$

Determine o seguinte:

- a.  $P(X < 2,8)$     b.  $P(X > 1,5)$   
c.  $P(X < -2)$     d.  $P(X > 6)$

4.2.2 Determine a função de distribuição cumulativa para a distribuição no Exercício 4.1.1.

4.2.3 Determine a função de distribuição cumulativa para a distribuição no Exercício 4.1.3.

4.2.4 Determine a função de distribuição cumulativa para a distribuição no Exercício 4.1.2.

4.2.5 Determine a função de distribuição cumulativa para a distribuição no Exercício 4.1.4.

4.2.6 A função densidade de probabilidade do tempo (em minutos depois das 8 horas) em que clientes chegam a um terminal é  $f(x) = 0,1\exp(-0,1x)$  para  $0 < x$ . Determine a probabilidade de:

- a. Você chegar até às 9 horas.  
b. Você chegar entre 8h15min e 8h30min.  
c. Você chegar antes das 8h40min em dois ou mais dias de cinco dias. Considere que os seus tempos de chegada em diferentes dias sejam independentes.

d. Determine a função de distribuição cumulativa e use-a para indicar a probabilidade de você chegar entre 8h15min e 8h30min.

4.2.7 Determine a função de distribuição cumulativa para a distribuição no Exercício 4.1.7. Use a função de distribuição cumulativa para determinar a probabilidade de um comprimento exceder 2,7 metros.

Determine a função densidade de probabilidade para cada uma das seguintes funções de distribuição cumulativa.

4.2.8  $F(x) = 1 - e^{-2x} \quad x > 0$

4.2.9

$$F(x) = \begin{cases} 0 & x < 0 \\ 0,2x & 0 \leq x < 4 \\ 0,04x + 0,64 & 4 \leq x < 9 \\ 1 & 9 \leq x \end{cases}$$

4.2.10 Determine a função de distribuição cumulativa para a variável aleatória do Exercício 4.1.10. Use a função de distribuição cumulativa para determinar a probabilidade de a variável aleatória ser menor que 55.

4.2.11 Determine a função de distribuição cumulativa para a variável aleatória do Exercício 4.1.9. Use a função de distribuição cumulativa para determinar a probabilidade de o tempo de espera ser menor do que uma hora.

4.2.12 Determine a função de distribuição cumulativa para a variável aleatória do Exercício 4.1.8. Use a função de distribuição cumulativa para determinar a probabilidade de  $400 < X < 500$ .

## Exercícios para a Seção 4.3

4.3.1 Suponha que  $f(x) = 1,5x^2$  para  $-1 < x < 1$ . Determine a média e a variância de  $X$ .

4.3.2 Suponha que  $f(x) = 0,125x$  para  $0 < x < 4$ . Determine a média e a variância de  $X$ .

4.3.3 Determine a média e a variância da variável aleatória do Exercício 4.1.1.

4.3.4 Determine a média e a variância da variável aleatória do Exercício 4.1.8.

4.3.5 Determine a média e a variância da variável aleatória do Exercício 4.1.10.

4.3.6 Suponha que uma função densidade de probabilidade do comprimento de cabos de computadores seja  $f(x) = 0,1$ , de 1200 a 1210 milímetros.

- a. Determine a média e o desvio-padrão do comprimento do cabo.  
b. Se as especificações do comprimento forem  $1195 < x < 1205$  milímetros, que proporção dos cabos estará dentro das especificações?

4.3.7 A espessura, em micrômetros, de um revestimento condutivo tem uma função densidade de  $600x^{-2}$  para  $100 \mu\text{m} < x < 120 \mu\text{m}$ .

- a. Determine a média e a variância da espessura do revestimento.  
b. Se o revestimento custar US\$ 0,50 por micrômetro de espessura em cada peça, qual será o custo médio de revestimento por peça?

## Exercícios para a Seção 4.4

4.4.1 Suponha que  $X$  tenha uma distribuição contínua uniforme no intervalo  $[-1; 1]$ . Determine o seguinte:

- A média, a variância e o desvio-padrão de  $X$ .
- O valor de  $x$ , tal que  $P(-x < X < x) = 0,90$ .
- A função de distribuição cumulativa.

4.4.2 A espessura de um flange em um componente de espaçonavé é uniformemente distribuída entre 0,95 e 1,05 milímetro. Determine o seguinte:

- A função de distribuição cumulativa da espessura do flange.
- A proporção de flanges que excedem 1,02 milímetro.
- A espessura excedida por 90 % dos flanges.
- A média e a variância da espessura do flange.

4.4.3 A espessura de um filme fotorresistente aplicado a pastilhas na fabricação de semicondutores, em certa localização na pastilha, é uniformemente distribuída entre 0,2050 e 0,2150 micrômetro. Determine o seguinte:

- A função de distribuição cumulativa da espessura do filme fotorresistente.
- A proporção de pastilhas que excedem 0,2125 micrômetro na espessura do filme fotorresistente.
- A espessura excedida por 10 % das pastilhas.
- A média e a variância da espessura do filme fotorresistente.

4.4.4 Um adulto pode ganhar ou perder duas libras de água durante o dia. Considere que as mudanças na massa de água sejam distribuídas uniformemente entre menos duas e mais duas libras em um dia. Qual é o desvio-padrão do peso dessa pessoa ao longo do dia?

4.4.5 Um *show* de golfinhos está marcado para começar às 9 horas, 9h30min e 10 horas. Quando o *show* começa, o portão é fechado. Um visitante chegará ao portão no tempo uniformemente distribuído entre 8h30min e 10 horas. Determine o seguinte:

- A função de distribuição cumulativa do tempo (em minutos) entre a chegada do visitante e 8h30min.
- A média e a variância da distribuição do item anterior.

c. A probabilidade de o visitante esperar menos de dez minutos para o início do *show*.

d. A probabilidade de o visitante esperar mais de 20 minutos para o início do *show*.

4.4.6 O volume de um xampu dentro de um frasco está uniformemente distribuído entre 374 e 380 mililitros.

- Quais são a média e o desvio-padrão do volume de xampu?
- Qual é a probabilidade de o frasco estar cheio com menos do que o valor nominal de 375 mililitros?
- Qual é o volume de xampu que está em excesso em 95 % dos frascos?
- Cada mililitro de xampu custa US\$ 0,002 ao produtor. Mais de 375 mililitros representam um custo extra para ele. Qual é o custo extra médio?

4.4.7 Uma mensagem de *e-mail* chegará em um tempo uniformemente distribuído entre 9 e 11 horas. Você verifica seu *e-mail* às 9h15min e a cada meia hora depois desse tempo.

- Qual é o desvio-padrão do tempo de chegada (em minutos)?
- Qual é a probabilidade de que a mensagem chegue menos de dez minutos antes de você vê-la?
- Qual é a probabilidade de que a mensagem chegue mais de 15 minutos antes de você vê-la?

4.4.8 Um emissor de elétrons produz raios de elétrons com energia cinética variável que está uniformemente distribuída entre 3 e 7 joules. Suponha que seja possível ajustar o limite superior da energia cinética (correntemente estabelecido em 7 joules).

- Qual é a energia cinética média?
- Qual é a variância da energia cinética?
- Qual é a probabilidade de um raio de elétron ter uma energia cinética de exatamente 3,2 joules?
- Qual deve ser o limite superior de modo que a energia cinética média aumente para 8 joules?
- Qual deve ser o limite superior de modo que a variância da energia cinética diminua para 0,75 joule?

## Exercícios para a Seção 4.5

4.5.1 Use a Tabela III do Apêndice A para determinar as seguintes probabilidades para a variável aleatória normal padrão  $Z$ :

- $P(Z < 1,32)$
- $P(Z < 3,0)$
- $P(Z > 1,45)$
- $P(Z > -2,15)$
- $P(-2,34 < Z < 1,76)$

4.5.2 Suponha que  $Z$  tenha uma distribuição normal padrão. Use a Tabela III do Apêndice A para determinar o valor de  $z$  que resolve cada um dos seguintes itens:

- $P(-z < Z < z) = 0,95$
- $P(-z < Z < z) = 0,99$
- $P(-z < Z < z) = 0,68$
- $P(-z < Z < z) = 0,9973$

4.5.3 Suponha que  $X$  seja normalmente distribuída, com uma média de 10 e um desvio-padrão de 2. Determine o seguinte:

- $P(Z < 13)$
- $P(Z > 9)$
- $P(6 < X < 14)$
- $P(2 < X < 4)$
- $P(-2 < X < 8)$

4.5.4 Suponha que  $X$  seja normalmente distribuída, com uma média de dez e um desvio-padrão de dois. Determine o valor de  $x$  que resolve cada um dos seguintes itens:

- $P(X > x) = 0,5$
- $P(X > x) = 0,95$
- $P(x < X < 10) = 0$
- $P(-x < X - 10 < x) = 0,95$
- $P(-x < X - 10 < x) = 0,99$

4.5.5 Um artigo em *Knee Surgery Sports Traumatology Arthroscopy* ["Effect of Provider Volume on Resource Utilization for Surgical Procedures" (2005, Vol. 13, pp. 273-279)] mostrou um tempo médio de 129 minutos e um desvio-padrão de 14 minutos para a cirurgia de reconstrução ACL em hospitais com alto volume de cirurgias (com mais de 300 de tais cirurgias por ano).

- Qual é a probabilidade de sua cirurgia ACL, em um hospital com alto volume de cirurgias, requerer um tempo maior do que dois desvios-padrão acima da média?

- b. Qual é a probabilidade de sua cirurgia ACL em um hospital com alto volume de cirurgias ser completada em menos de 100 minutos?
- c. Em qual tempo a probabilidade de sua cirurgia ACL em um hospital com alto volume de cirurgias é igual a 0,95?
- d. Se sua cirurgia requer 199 minutos, o que você conclui acerca do volume de tais cirurgias em seu hospital? Explique.
- 4.5.6** O tempo de recarga, sob condições normais, de uma bateria de um *laptop* é distribuído normalmente, com uma média de 260 minutos e um desvio-padrão de 50 minutos.
- a. Qual é a probabilidade de a bateria durar mais de quatro horas?
- b. Quais são os quartis (os valores de 25 % e 75 %) da vida da bateria?
- c. Qual o valor da vida em minutos é excedido com 95 % de probabilidade?
- 4.5.7** O tempo de reação de um motorista para o estímulo visual é normalmente distribuído, com uma média de 0,4 segundo e um desvio-padrão de 0,05 segundo.
- a. Qual é a probabilidade de uma reação requerer mais de 0,5 segundo?
- b. Qual é a probabilidade de uma reação requerer entre 0,4 e 0,5 segundo?
- c. Qual o tempo de reação é excedido em 90 % do tempo?
- 4.5.8** Colesterol é uma substância gordurosa que é uma parte importante da ligação (membrana) externa das células no corpo de animais. Sua faixa normal para um adulto é de 120-240 mg/dL. O Instituto de Alimentos e Nutrição das Filipinas encontrou que o nível de colesterol total para adultos filipinos tem uma média de 159,2 mg/dL, e 84,1 % de adultos têm um nível de colesterol abaixo de 200 mg/dL ([www.fnri.dost.gov.ph/](http://www.fnri.dost.gov.ph/)). Suponha que o nível de colesterol total seja distribuído normalmente.
- a. Determine o desvio-padrão dessa distribuição.
- b. Quais são os quartis (os valores de 25 % e 75 %) dessa distribuição?
- c. Qual é o valor do nível de colesterol que excede 90 % da população?
- d. Um adulto tem um nível moderado de risco, se o nível de colesterol é mais de um, porém menos de dois desvios-padrão acima da média. Qual é a percentagem da população que tem risco moderado de acordo com esse critério?
- e. Um adulto tem alto risco, se seu nível de colesterol é mais de dois desvios-padrão acima da média. Qual é a percentagem da população que tem alto risco?
- f. Um adulto tem baixo risco, se seu nível de colesterol é menor do que um desvio-padrão abaixo da média. Qual é a percentagem da população que tem baixo risco?
- 4.5.9** Em um centro acelerador, um experimento necessita de um cilindro de alumínio, com 1,41 centímetro de espessura ([http://puhep1.princeton.edu/mumu/target/Solenoid\\_Coil.pdf](http://puhep1.princeton.edu/mumu/target/Solenoid_Coil.pdf)). Suponha que a espessura de um cilindro tenha uma distribuição normal, com média igual a 1,41 centímetro e um desvio-padrão igual a 0,01 centímetro.
- a. Qual é a probabilidade de a espessura ser maior do que 1,42 centímetro?
- b. Que espessura é excedida por 95 % das amostras?
- c. Se as especificações requerem que a espessura esteja entre 1,39 e 1,43 centímetro, que proporção das amostras satisfaz as especificações?
- 4.5.10** A altura média de uma mulher na faixa de 20-74 anos era de 64 polegadas em 2002, tendo aumentado aproximadamente uma polegada a partir de 1960 (<http://usgovinfo.about.com/od/healthcare>). Suponha que a altura de uma mulher seja normalmente distribuída, com um desvio-padrão de duas polegadas.
- a. Qual é a probabilidade de uma mulher, selecionada aleatoriamente dessa população, ter entre 58 e 70 polegadas?
- b. Quais são os quartis dessa distribuição?
- c. Determine a altura, simétrica em torno da média, que inclui 90 % dessa população.
- d. Qual é a probabilidade de a altura de cinco mulheres, selecionadas, ao acaso, dessa população, exceder 68 polegadas?
- 4.5.11** A vida de um semicondutor a *laser*, a uma potência constante, é normalmente distribuída com uma média de 7000 horas e desvio-padrão de 600 horas.
- a. Qual é a probabilidade de o *laser* falhar antes de 5000 horas?
- b. Qual é o tempo de vida em horas que 95 % dos *lasers* excedem?
- c. Se três *lasers* forem usados em um produto e se eles falharem independentemente, qual será a probabilidade de todos os três estarem ainda operando depois de 7000 horas?
- 4.5.12** A demanda por uso de água em Fênix, em 2003, alcançou um alto valor, de cerca de 442 milhões de galões por dia em 27 de junho de 2003 (<http://phoenix.gov/WATER/wtrfacts.html>). O consumo de água no verão é distribuído normalmente, com uma média de 310 milhões de galões por dia e um desvio-padrão de 45 milhões de galões por dia. Reservatórios da cidade têm uma capacidade combinada de armazenagem de aproximadamente 350 milhões de galões.
- a. Qual é a probabilidade de que um dia requeira mais água que aquela armazenada nos reservatórios da cidade?
- b. Que capacidade do reservatório é necessária para que a probabilidade de ela ser excedida seja 1 %?
- c. Qual é a intensidade de uso de água que é excedida com 95 % de probabilidade?
- d. Água é fornecida para aproximadamente 1,4 milhão de pessoas. Qual é o consumo médio diário por pessoa no qual a probabilidade de a demanda exceder a capacidade atual do reservatório seja igual a 1 %? Considere que o desvio-padrão da demanda continue o mesmo.
- 4.5.13** O peso de um sofisticado sapato de corrida é normalmente distribuído, com uma média de 12 onças e um desvio-padrão de 0,5 onça.
- a. Qual é a probabilidade de o sapato pesar mais de 13 onças?
- b. Qual deve ser o desvio-padrão do peso para que a companhia estabeleça que 99,9 % de seus sapatos sejam menores do que 13 onças?
- c. Se o desvio-padrão permanecer em 0,5 onça, qual deverá ser o peso médio para que a companhia estabeleça que 99,9 % de seus sapatos sejam menores que 13 onças?
- 4.5.14** O diâmetro do ponto produzido por uma impressora é normalmente distribuído, com uma média de 0,002 polegada e um desvio-padrão de 0,0004 polegada.
- a. Qual é a probabilidade de o diâmetro de um ponto exceder 0,0026 polegada?
- b. Qual é a probabilidade de um diâmetro estar entre 0,0014 e 0,0026 polegada?
- c. Que desvio-padrão do diâmetro é necessário para que a probabilidade do item (b) seja 0,995?
- 4.5.15** Suponha que uma variável aleatória seja normalmente distribuída, com uma média de 24 e um desvio-padrão de 2. Considere um intervalo de comprimento unitário que comece no valor  $a$ , de modo que o intervalo seja  $[a, a + 1]$ . Para que valor de  $a$  a probabilidade é do maior intervalo? O desvio-padrão afeta tal escolha de intervalo?

4.5.16 O período de permanência em uma emergência específica em Fênix, Arizona, em 2009, teve uma média de 4,6 horas, com um desvio-padrão de 2,9. Considere que o período de permanência seja normalmente distribuído.

- Qual é a probabilidade de um período de permanência ser maior do que dez horas?
- Que período de permanência é excedido por 25 % das pessoas?
- A partir do modelo de distribuição normal, qual é a probabilidade de um período de permanência ser menor do que zero hora? Comente a respeito da suposição de distribuição normal neste exemplo.

4.5.17 Um artigo em *Atmospheric Chemistry and Physics* [“Relationship Between Particulate Matter and Childhood Asthma – Basis of a Future Warning System for Central Phoenix” (2012, Vol. 12, pp. 2479-2490)] usou dados de qualidade do ar para PM10 (matéria particulada < 10  $\mu\text{m}$  em diâmetro), medidos por hora, usando sensores em Fênix, Arizona. A média diária de PM10 para um sensor localizado no centro foi de 50,9  $\mu\text{g}/\text{m}^3$ , com um desvio-padrão de 25,0. Considere que a média diária de PM10 seja normalmente distribuída.

- Qual é a probabilidade de uma média diária de PM10 ser maior do que 100  $\mu\text{g}/\text{m}^3$ ?
- Qual é a probabilidade de uma média diária de PM10 ser menor do que 25  $\mu\text{g}/\text{m}^3$ ?
- Que média diária do valor de PM10 é excedida com probabilidade de 5 %?

4.5.18 Um artigo em *Microelectronics Reliability* [“Advanced Electronic Prognostics Through System Telemetry and Pattern Recognition Methods” (2007, Vol. 47(12), pp. 1865-1873)] apresentou um exemplo de prognóstico eletrônico. O objetivo foi detectar falhas de modo a diminuir o tempo de inatividade do sistema e o número de reparos não planejados em sistemas de alta

confiabilidade. Medidas prévias do suprimento de potência indicaram que o sinal é normalmente distribuído com uma média de 1,5 V e um desvio-padrão de 0,02 V.

- Suponha que os limites inferior e superior das especificações predeterminadas sejam 1,45 V e 1,55 V, respectivamente. Qual será a probabilidade de que um sinal esteja dentro dessas especificações?
- Qual o valor do sinal que é excedido com 95 % de probabilidade?
- Qual é a probabilidade de que o valor do sinal exceda a média por dois ou mais desvios-padrão?

4.5.19 Um sinal em um canal de comunicação é detectado quando a voltagem é maior do que 1,5 volt em valor absoluto. Considere que a voltagem seja distribuída normalmente, com uma média de 0. Qual é o desvio-padrão da voltagem, tal que a probabilidade de um sinal falso seja 0,005?

4.5.20 Um artigo publicado no *Journal of Cardiovascular Magnetic Resonance* [“Right Ventricular Ejection Fraction Is Better Reflected by Transverse Rather Than Longitudinal Wall Motion in Pulmonary Hypertension” (2010, Vol. 12(35))] discutiu um estudo do movimento da região da parede transversal do ventrículo direito em pacientes com hipertensão pulmonar (HP). A fração de ejeção (FE) do ventrículo direito foi aproximadamente distribuída normalmente, com média e desvio-padrão de 36 e 12, respectivamente, para indivíduos com HP, e com média e desvio-padrão de 56 e 8, respectivamente, para indivíduos-controle.

- Com 5 % de probabilidade, qual é a FE para indivíduos com HP que é excedida?
- Qual é a probabilidade de que a FE de um indivíduo-controle seja menor do que o valor do item (a)?
- Comente sobre quão bem os indivíduos-controle e com HP podem ser distinguidos por medidas de FE.

## Exercícios para a Seção 4.6

4.6.1 Suponha que  $X$  seja uma variável aleatória binomial, com  $n = 200$  e  $p = 0,4$ . Aproxime as seguintes probabilidades:

- $P(X \leq 70)$
- $P(70 < X < 90)$
- $P(X = 80)$

4.6.2 Suponha que  $X$  seja uma variável aleatória de Poisson, com  $\lambda = 6$ .

- Calcule a probabilidade exata de  $X$  ser menor que quatro.
- Aproxime a probabilidade de  $X$  ser menor que quatro e compare-a ao resultado do item (a).
- Aproxime a probabilidade de  $8 < X < 12$ .

4.6.3 Havia 49,7 milhões de pessoas com algum tipo de condição ou incapacidade duradoura vivendo nos Estados Unidos em 2000. Isso representava 19,3 % da maioria de civis com cinco e mais de cinco anos de idade (<http://factfinder.census.gov>). Uma amostra de 1000 pessoas é selecionada ao acaso.

- Aproxime a probabilidade de mais de 200 pessoas na amostra terem uma incapacidade.
- Aproxime a probabilidade de 180 a 300 pessoas na amostra terem uma incapacidade.

4.6.4 A fabricação de *chips* semicondutores produz 2 % de *chips* defeituosos. Considere que os *chips* sejam independentes e que um lote contenha 1000 *chips*. Aproxime as seguintes probabilidades:

- Mais de 25 *chips* são defeituosos.
- Entre 20 e 30 *chips* são defeituosos.

4.6.5 Golpes em *sites* da internet, com alta consulta, seguem supostamente uma distribuição de Poisson, com uma média de 10.000 por dia. Aproxime cada uma das seguintes opções:

- A probabilidade de mais de 20.000 golpes em um dia.
- A probabilidade de menos de 9900 golpes em um dia.
- O valor tal que a probabilidade de o número de golpes em um dia exceder esse valor seja igual a 0,01.
- O número esperado de dias em um ano (365 dias) que excede 10.200 golpes.
- A probabilidade de que, ao longo de um ano (365 dias), mais de 15 dias tenham mais de 10.200 golpes por dia.

4.6.6 A água de Fênix (no Arizona) é fornecida para, aproximadamente, 1,4 milhão de pessoas, que são servidas por meio de mais de 362.000 contas (<http://phoenix.gov/WATER/wtrfacts.html>). Todas as contas são medidas e cobradas mensalmente. A probabilidade de uma conta conter um erro em um mês é 0,001, e contas podem ser consideradas independentes.

- Quais são a média e o desvio-padrão do número de contas com erro a cada mês?
- Aproxime a probabilidade de menos de 350 contas com erro em um mês.
- Aproxime um valor de modo que a probabilidade de o número de contas com erro exceder esse valor seja 0,05.

d. Aproxime a probabilidade de mais de 400 contatos com erro por mês nos próximos dois meses. Considere que os resultados entre os meses sejam independentes.

**4.6.7** Um artigo em *Atmospheric Chemistry and Physics* [“Relationship Between Particulate Matter and Childhood Asthma – Basis of a Future Warning System for Central Phoenix” (2012, Vol. 12, pp. 2479-2490)] relacionou qualidade do ar com casos de asma na infância. A região de estudo no centro de Fênix, Arizona, registrou 10.500 casos de asma em crianças em um período de 21 meses. Considere que o número de casos de asma segue uma distribuição de Poisson.

- Aproxime a probabilidade de mais de 550 casos de asma em um mês.
- Aproxime a probabilidade de 450 a 550 casos de asma em um mês.
- Aproxime o número de casos de asma excedidos com probabilidade de 5 %.
- Se o número de casos de asma fosse maior durante o inverno do que no verão, o que isso implicaria na suposição da distribuição de Poisson?

**4.6.8** Um artigo em *Biometrics* [“Integrative Analysis of Transcriptomic and Proteomic Data of *Desulfovibrio Vulgaris*: A Nonlinear Model to Predict Abundance of Undetected Proteins” (2009, Vol. 25(15), pp. 1905-1914)] encontrou que a abundância de proteína proveniente de um operon (um conjunto de genes biologicamente relacionados) estava menos dispersa do que a partir de genes selecionados aleatoriamente. Na pesquisa, 1000 conjuntos de genes foram construídos aleatoriamente e 75 % desses conjuntos estavam mais dispersos do que um operon específico. Se a probabilidade de um conjunto aleatório estar mais disperso do que esse operon for verdadeiramente 0,5, aproxime a probabilidade de que 750 ou mais conjuntos aleatórios excedam o operon. A partir desse resultado, o que você conclui acerca da dispersão no operon *versus* genes aleatórios?

**4.6.9** Um conjunto de 200 pacientes independentes tomam medicamento antiácido no começo dos sintomas e 80 têm alívio de moderado a substancial dentro de 90 minutos. Historicamente, 30 % dos pacientes experimentam alívio dentro de 90 minutos sem nenhum medicamento. Se o medicamento não tem efeito, aproxime a probabilidade de que 80 ou mais pacientes experimentem alívio de sintomas. O que você pode concluir sobre a eficiência desse medicamento?

**4.6.10** Táxis passam por seu ponto de acordo com um processo de Poisson, com uma média de cinco táxis por hora.

- Determine a média e o desvio-padrão do número de táxis por dia de dez horas.
- Aproxime a probabilidade de mais de 65 táxis passarem em um dia de dez horas.
- Aproxime a probabilidade de entre 50 e 65 táxis passarem em um dia de dez horas.
- Determine a taxa horária média, de modo que a probabilidade seja aproximadamente 0,95 de que 100 ou mais táxis passem em um dia de dez horas.

**4.6.11** O número de inclusões (grandes) em ferro fundido segue uma distribuição normal, com uma média de 2,5 milímetros cúbicos. Aproxime as seguintes probabilidades:

- Determine a média e o desvio-padrão do número de inclusões em um centímetro cúbico (cc).
- Aproxime a probabilidade de menos de 2600 inclusões ocorrerem em um cc.
- Aproxime a probabilidade de mais de 2400 inclusões ocorrerem em um cc.
- Determine o número médio de inclusões por milímetro cúbico, tal que a probabilidade de 500 ou menos inclusões ocorrerem em um cc seja de aproximadamente 0,9.

## Exercícios para a Seção 4.7

**4.7.1** Suponha que  $X$  tenha uma distribuição exponencial, com média igual a 10. Determine o seguinte:

- $P(X > 10)$
- $P(X > 20)$
- $P(X < 30)$
- Encontre o valor de  $x$  tal que  $P(X < x) = 0,95$ .

**4.7.2** Suponha que  $X$  tenha uma distribuição exponencial, com média igual a 10. Determine o seguinte:

- $P(X < 5)$
- $P(X < 15 | X > 10)$
- Compare os resultados dos itens (a) e (b) e comente sobre o papel da propriedade de falta de memória.

**4.7.3** Suponha que as conexões a uma rede de computadores sigam um processo de Poisson, com uma média de três contagens por minuto.

- Qual é o tempo médio entre as contagens?
- Qual é o desvio-padrão do tempo entre as contagens?
- Determine  $x$  tal que a probabilidade de no mínimo uma contagem ocorrer antes do tempo  $x$  minutos seja de 0,95.

**4.7.4** A vida de reguladores de voltagem para automóveis tem uma distribuição exponencial, com uma vida média de seis anos. Você compra um automóvel com seis anos de uso, com um regulador de voltagem funcionando, e planeja ficar com o carro por seis anos.

- Qual é a probabilidade de o regulador de voltagem falhar durante o período em que você possui o carro?
- Se o seu regulador falhar depois de você possuir o carro por três anos e se ele for trocado, qual é o tempo médio até a próxima falha?

**4.7.5** O tempo entre as chamadas para uma loja de suprimento de encanamentos é distribuído exponencialmente, com um tempo médio de 15 minutos entre as chamadas.

- Qual é a probabilidade de não haver chamadas dentro do intervalo de 30 minutos?
- Qual é a probabilidade de no mínimo uma chamada chegar dentro do intervalo de dez minutos?
- Qual é a probabilidade de a primeira chamada chegar dentro de cinco e dez minutos depois da loja aberta?
- Determine o comprimento de um intervalo de tempo, tal que exista probabilidade igual a 0,90 de haver no mínimo uma chamada no intervalo.

**4.7.6** O número de aparições de cegonhas em uma rota na Carolina do Sul segue um processo de Poisson, com uma média de 2,3 por ano.

- Qual é o tempo médio entre as aparições?
- Qual é a probabilidade de não haver aparições dentro de três meses (0,25 ano)?
- Qual é a probabilidade de que o tempo até a primeira aparição exceda seis meses?
- Qual é a probabilidade de não haver aparições dentro de três anos?

4.7.7 O tempo entre as chegadas de táxis a uma interseção movimentada é distribuído exponencialmente, com uma média de dez minutos.

- Qual é a probabilidade de você esperar mais de uma hora por um táxi?
- Suponha que você já estivesse esperando uma hora por um táxi. Qual será a probabilidade de um táxi chegar dentro dos próximos dez minutos?
- Determine  $x$  tal que a probabilidade de você esperar mais de  $x$  minutos seja 0,10.
- Determine  $x$  tal que a probabilidade de você esperar menos de  $x$  minutos seja 0,90.
- Determine  $x$  tal que a probabilidade de você esperar menos de  $x$  minutos seja 0,50.

4.7.8 A distância entre fraturas grandes em uma autoestrada segue uma distribuição exponencial, com uma média de cinco milhas.

- Qual é a probabilidade de não haver grandes fraturas em uma extensão de dez milhas de autoestrada?
- Qual é a probabilidade de haver duas grandes fraturas em uma extensão de dez milhas de autoestrada?
- Qual é o desvio-padrão da distância entre grandes fraturas?
- Qual é a probabilidade de que a primeira grande fratura ocorra entre 12 e 15 milhas a partir do começo da inspeção?
- Qual é a probabilidade de não haver grandes fraturas em duas extensões separadas por cinco milhas de autoestrada?
- Dado que não existem fraturas nas cinco primeiras milhas inspecionadas, qual é a probabilidade de não haver grandes fraturas nas próximas dez milhas inspecionadas?

4.7.9 De acordo com os resultados da análise de barras de chocolate no Capítulo 3, o número médio de fragmentos de insetos foi de 14,4 em 225 gramas. Considere que o número de fragmentos segue uma distribuição de Poisson.

- Qual é o número médio de gramas de chocolate até que um fragmento seja detectado?
- Qual é a probabilidade de não haver fragmentos em uma barra de chocolate de 28,35 gramas (uma onça)?
- Suponha que você consuma sete barras de uma onça (28,35 gramas) esta semana. Qual é a probabilidade de não haver fragmentos de insetos?

4.7.10 O tempo entre as chegadas de pequenas aeronaves a um aeroporto em um condado é distribuído exponencialmente, com média de uma hora.

- Qual é a probabilidade de mais de três aeronaves chegarem dentro de uma hora?
- Se 30 intervalos, de uma hora cada, forem escolhidos, qual será a probabilidade de nenhum intervalo conter mais de três chegadas?
- Determine o comprimento de um intervalo de tempo (em horas), tal que a probabilidade de nenhuma chegada ocorrer durante o intervalo seja 0,10.
- Qual é a probabilidade de a primeira ocorrência na qual o tempo entre duas chegadas excede 1,5 hora estar na terceira chegada?

4.7.11 O tempo entre as chamadas para o escritório de uma corporação é distribuído exponencialmente, com uma média de dez minutos.

- Qual é a probabilidade de que haja mais de três chamadas em meia hora?
- Qual é a probabilidade de não haver chamadas dentro de meia hora?
- Determine  $x$  tal que a probabilidade de nenhuma chamada ocorrer durante  $x$  horas seja igual a 0,01.
- Qual é a probabilidade de não haver chamadas em um intervalo de duas horas?
- Se quatro intervalos, não coincidentes, de meia hora forem selecionados, qual será a probabilidade de nenhum desses intervalos conter qualquer chamada?
- Qual é a probabilidade de a primeira ocorrência, na qual o tempo entre duas chamadas exceda dez minutos, ser a quinta chamada?

4.7.12 Deduza a fórmula para a média e a variância de uma variável aleatória exponencial.

4.7.13 Se a variável aleatória  $X$  tiver uma distribuição exponencial, com média  $\theta$ , determine o seguinte:

- $P(X > \theta)$
- $P(X > 2\theta)$
- $P(X > 3\theta)$
- Como os resultados dependem de  $\theta$ ?

4.7.14 O período de permanência em uma emergência hospitalar específica em Fênix, Arizona, teve uma média de 4,6 horas. Suponha que o período de permanência seja distribuído exponencialmente.

- Qual é o desvio-padrão do período de permanência?
- Qual é a probabilidade de um período de permanência ser maior do que dez horas?
- Que período de permanência é excedido por 25 % dos pacientes?

4.7.15 Um artigo publicado no *Journal of National Cancer Institute* ["Breast Cancer Screening Policies in Developing Countries: A Cost-Effectiveness Analysis for India" (2008, Vol. 100(18), pp. 1290-1300)] apresentou um modelo de análise exploratória de câncer de mama, baseado em dados da Índia. Nessa análise, o tempo no qual o caso de câncer de mama permanece em um estado pré-clínico é modelado por uma distribuição exponencial, com uma média dependente do estado. Por exemplo, o tempo em que um caso de câncer permanece no estado de TIC (tamanho do tumor de 11-20 mm) é distribuído exponencialmente com uma média de 1,48 ano.

- Qual é a probabilidade de que um caso de câncer de mama na Índia permaneça no estado de TIC por mais de dois anos?
- Qual é a proporção de casos de câncer de mama na Índia que leva no mínimo um ano no estado de TIC?
- Considere que uma pessoa na Índia seja diagnosticada no estado de TIC. Qual será a probabilidade de o paciente estar no mesmo estado seis meses depois?

4.7.16 Um artigo em *Ad Hoc Networks* ["Underwater Acoustic Sensor Networks: Target Size Detection and Performance Analysis" (2009, Vol. 7(4), pp. 803-808)] discutiu uma rede de sensores acústicos subaquáticos para monitorar determinada área em um oceano. A rede não usa cabos e não interfere em atividades de remessa. A chegada de grupos de sinais gerados pelo mesmo pulso é tido como um processo de Poisson, com uma média de  $\lambda$  por unidade de tempo. Suponha que para uma rede específica de sensores acústicos subaquáticos, esse processo de Poisson tenha uma taxa de 2,5 chegadas por unidade de tempo.

- Qual é o tempo médio entre duas chegadas consecutivas?
- Qual é a probabilidade de que não haja chegadas dentro de 0,3 unidade de tempo?
- Qual é a probabilidade de que o tempo até a primeira chegada exceda uma unidade de tempo?
- Determine a taxa média de chegada, tal que a probabilidade de haver chegadas em 0,3 unidade de tempo seja de 0,9.

4.7.17 Um artigo em *Vaccine* ["Modeling the Effects of Influenza Vaccination of Health Care Workers in Hospital Departments" (2009, Vol. 27(44), pp. 6261-6267)] considerou a imunização de agentes de saúde para reduzir a taxa de perigo de infecção pelo vírus influenza em pacientes de departamentos de hospitais regulares. Nessa análise, cada permanência do paciente no departamento é considerada como distribuída exponencialmente, com uma média de sete dias.

- Qual a probabilidade de um paciente permanecer no hospital por menos de 5,5 dias?
- Qual é a probabilidade de um paciente permanecer no hospital por mais de dez dias, se o paciente tiver correntemente permanecido por sete dias?
- Determine a permanência média, tal que a probabilidade de um paciente ficar no hospital menos de seis dias seja de 0,9.

## Exercícios para a Seção 4.8

4.8.1 Use as propriedades da função gama para avaliar o seguinte:

- a.  $\Gamma(6)$     b.  $\Gamma(5/2)$     c.  $\Gamma(9/2)$

4.8.2 Dada a função densidade de probabilidade  $f(x) = 0,01^3 x^2 e^{-0,01x} / \Gamma(3)$ , determine a média e a variância da distribuição.

4.8.3 Chamadas para o sistema telefônico seguem uma distribuição de Poisson, com uma média de cinco chamadas por minuto.

- a. Qual é o nome aplicado à distribuição e quais são os valores dos parâmetros do tempo até a décima chamada?  
 b. Qual é o tempo médio até a décima chamada?  
 c. Qual é o tempo médio entre a nona e a décima chamadas?  
 d. Qual é a probabilidade de, exatamente, quatro chamadas ocorrerem dentro de um minuto?  
 e. Se dez intervalos separados por um minuto forem escolhidos, qual será a probabilidade de que todos os intervalos contenham mais de duas chamadas?

4.8.4 Matérias-primas são estudadas em relação a contaminações. Suponha que o número de partículas de contaminação por libra de material seja uma variável aleatória de Poisson, com média de 0,01 partícula por libra.

- a. Qual é o número esperado de libras de material requerido para obter 15 partículas de contaminação?  
 b. Qual é o desvio-padrão das libras de materiais requeridas para obter 15 partículas de contaminação?

4.8.5 O tempo entre chegadas de usuários em um caixa eletrônico é uma variável aleatória exponencial, com média de cinco minutos.

- a. Qual é a probabilidade de mais de três usuários chegarem em dez minutos?  
 b. Qual é a probabilidade de o tempo, até que o quinto usuário chegue, ser menor que 15 minutos?

4.8.6 Em um sistema de comunicação de dados, várias mensagens que chegam a um nó são agrupadas em um bloco, antes de serem transmitidas ao longo da rede. Suponha que as mensagens cheguem ao nó de acordo com um processo de Poisson, com  $\lambda = 30$  mensagens por minuto. Cinco mensagens são usadas para formar um bloco.

- a. Qual é o tempo médio até que um bloco seja formado, isto é, até que cinco mensagens cheguem ao nó?  
 b. Qual é o desvio-padrão do tempo até que um bloco seja formado?  
 c. Qual é a probabilidade de um pacote ser formado em menos de dez segundos?  
 d. Qual é a probabilidade de um pacote ser formado em menos de cinco segundos?

4.8.7 Pacientes chegam a uma emergência de acordo com um processo de Poisson, com média de 6,5 por hora.

- a. Qual é o tempo médio até a décima chegada?  
 b. Qual é a probabilidade de que mais de 20 minutos sejam necessários para a terceira chegada?

4.8.8 Use a integração por partes para mostrar que  $\Gamma(r) = (r - 1)\Gamma(r - 1)$ .

4.8.9 Um artigo em *Sensors and Actuators A: Physical* ["Characterization and Simulation of Avalanche PhotoDiodes for Next-Generation Colliders" (2011, Vol. 172(1), pp. 181-188)] considerou um fotodiodo de avalanche (APD) para detectar partículas carregadas em uma foto. O número de chegadas em cada janela de detecção foi modelado com uma distribuição de Poisson, com uma média dependente da intensidade do feixe. Para uma intensidade de feixe, o número de elétrons que chegam em um APD segue uma distribuição de Poisson, com média de 1,74 partícula por janela de detecção de 200 nanossegundos.

- a. Qual é a média e a variância do tempo para 100 chegadas?  
 b. Qual é a probabilidade de o tempo até a chegada da quinta partícula ser maior do que 1,0 nanossegundo?

4.8.10 O tempo total de serviço de uma operação multietapas de fabricação tem uma distribuição gama com média de 18 minutos e desvio-padrão de 6.

- a. Determine os parâmetros  $\lambda$  e  $r$  da distribuição.  
 b. Considere que cada etapa tem a mesma distribuição para o tempo de serviço. Qual é a distribuição para cada etapa e quantas etapas produz essa distribuição gama do tempo total de serviço?

4.8.11 Um artigo em *Mathematical Biosciences* ["Influence of Delayed Viral Production on Viral Dynamics in HIV-1 Infected Patients" (1998, Vol. 152(2), pp. 143-163)] considerou o atraso no tempo entre a infecção inicial pelo tipo 1 do vírus de imunodeficiência (HIV-1) e a formação de células produtivamente infectadas. No modelo de simulação, o atraso no tempo é aproximado por uma distribuição gama com parâmetros  $r = 4$  e  $1/\lambda = 0,25$  dia. Determine o seguinte:

- a. Média e variância do atraso no tempo.  
 b. Probabilidade de o atraso no tempo ser mais do que a metade de um dia.  
 c. Probabilidade de o atraso no tempo estar entre metade de um dia e um dia todo.

## Exercícios para a Seção 4.9

4.9.1 Se  $X$  for uma variável aleatória de Weibull, com  $\beta = 1$  e  $\delta = 1000$  horas, qual será o outro nome para a distribuição de  $X$  e qual será a média de  $X$ ?

4.9.2 Suponha que  $X$  tenha uma distribuição de Weibull, com  $\beta = 0,2$  e  $\delta = 100$  horas. Determine o seguinte:

- a.  $P(X < 10.000)$     b.  $P(X > 5000)$     c.  $E(X)$  e  $V(X)$

4.9.3 A vida (em horas) de um aparelho de imagem por ressonância magnética (IRM) é modelada por uma distribuição de Weibull, com parâmetros  $\beta = 2$  e  $\delta = 500$  horas. Determine o seguinte:

- a. A vida média do aparelho de IRM.

b. A variância da vida do aparelho de IRM.

c. A probabilidade de o aparelho de IRM falhar antes de 250 horas.

4.9.4 Suponha que a vida de um mancal de rolamento siga uma distribuição de Weibull, com parâmetros  $\beta = 2$  e  $\delta = 10.000$  horas.

- a. Determine a probabilidade de um mancal durar no mínimo 8000 horas.  
 b. Determine o tempo médio até haver falha de um mancal.  
 c. Se dez mancais estiverem em uso e as falhas ocorrerem independentemente, qual será a probabilidade de que todos os dez mancais durem no mínimo 8000 horas?

4.9.5 Um artigo publicado no *Journal of Geophysical Research* ["Spatial and Temporal Distributions of U.S. of Winds and Wind Power at 80 m Derived from Measurements" (2003, vol. 108)] considerou a velocidade do vento em estações em todos os Estados Unidos. Para uma estação em Amarillo, Texas, a velocidade média do vento a 80 metros (a altura do ponto mais alto de grandes turbinas de vento) foi de 10,3 m/s, com um desvio-padrão de 4,9 m/s. Determine os parâmetros de forma e de escala de uma distribuição de Weibull com essas propriedades.

4.9.6 Um artigo publicado no *Journal of the Indian Geophysical Union*, intitulado "Weibull and Gamma Distributions for Wave Parameter Predictions" (2005, Vol. 9, pp. 55-64), usou a distribuição de Weibull para modelar as alturas das ondas do oceano. Considere que a média da altura da onda na estação de observação seja de 2,5 metros e o parâmetro de forma seja igual a 2. Determine o desvio-padrão da altura da onda.

4.9.7 Suponha que o tempo (em horas) de vida de um componente seja modelado com uma distribuição de Weibull, com  $\beta = 2$  e  $\delta = 4000$ . Determine o seguinte, quanto aos itens (a) e (b):

- $P(X > 5000)$
- $P(X > 8000 \mid X > 3000)$
- Comente a respeito das probabilidades dos itens prévios comparadas aos resultados para uma distribuição exponencial.

4.9.8 Suponha que  $X$  tenha uma distribuição de Weibull com  $\beta = 2$  e  $\delta = 8,6$ . Determine o seguinte:

- $P(X < 10)$
- $P(X > 9)$
- $P(8 < X < 11)$
- Valor para  $x$ , tal que  $P(X > x) = 0,9$ .

4.9.9 Um artigo em *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval* ["Understanding Web Browsing Behaviors Through Weibull Analysis of Dwell Time" (2010, pp. 379-386)] propôs que uma distribuição de Weibull fosse usada para modelar o tempo de permanência em uma página na internet (o comprimento de tempo que um visitante passa em uma página da Web). Para uma página específica da internet, os parâmetros de forma e de escala são 1 e 300 segundos, respectivamente. Determine o seguinte:

- Média e variância do tempo de permanência.
- Probabilidade de um usuário da internet gastar mais de quatro minutos nessa página.
- Tempo de permanência excedido com probabilidade de 0,25.

4.9.10 Um artigo publicado no *Electronic Journal of Applied Statistical Analysis* ["Survival Analysis of Dialysis Patients Under Parametric and Non-Parametric Approaches" (2012, Vol. 5(2), pp. 271-288)] modelou o tempo de sobrevivência de pacientes em diálise que são doentes renais crônicos, por meio da distribuição de Weibull. A média e o desvio-padrão do tempo de sobrevivência foram 16,01 e 11,66 meses, respectivamente. Determine o seguinte:

- Parâmetros de forma e de escala dessa distribuição de Weibull.
- Probabilidade de o tempo de sobrevivência ser maior do que 48 meses.
- Tempo de sobrevivência que excedeu uma probabilidade de 90 %.

4.9.11 Um artigo em *IEEE Transactions on Dielectrics and Electrical Insulation* ["Statistical Analysis of the AC Breakdown Voltages of Ester Based Transformer Oils" (2008, Vol. 15(4), pp. 1044-1050)] usou as distribuições de Weibull para modelar a voltagem de ruptura de isoladores. A voltagem de ruptura é a voltagem mínima conduzida pelo isolador. Para 1 mm de éster natural, a probabilidade de 1 % da voltagem de ruptura representa aproximadamente 26 kV, e a probabilidade de 7 % representa aproximadamente 31,6 kV. Determine os parâmetros  $\delta$  e  $\beta$  da distribuição de Weibull.

4.9.12 Um artigo em *Financial Markets Institutions and Instruments* ["Pricing Reinsurance Contracts on FDIC Losses" (2008, Vol. 17(3), pp. 225-247)] modelou as perdas anuais médias (em bilhões de dólares) de uma corporação de seguros (Federal Deposit Insurance Corporation, FDIC) por uma distribuição de Weibull, com parâmetros  $\delta = 1,9317$  e  $\beta = 0,8472$ . Determine o seguinte:

- Probabilidade de uma perda maior do que 2 bilhões.
- Probabilidade de uma perda entre 2 e 4 bilhões.
- Probabilidade de exceder o valor igual a 0,05.
- Média e desvio-padrão da perda.

## Exercícios para a Seção 4.10

4.10.1 Suponha que  $X$  tenha uma distribuição lognormal, com parâmetros  $\theta = 5$  e  $\omega^2 = 9$ . Determine o seguinte:

- $P(X < 13.300)$ .
- O valor para  $x$ , tal que  $P(X \leq x) = 0,95$ .
- A média e a variância de  $X$ .

4.10.2 Suponha que  $X$  tenha uma distribuição lognormal, com parâmetros  $\theta = 2$  e  $\omega^2 = 4$ . Determine o seguinte nos itens (a) e (b):

- $P(X < 500)$ .
- A probabilidade condicional de  $X < 1500$ , dado que  $X > 1000$ .
- O que a diferença entre as probabilidades dos itens (a) e (b) afeta os tempos de vida das variáveis aleatórias lognormais?

4.10.3 O período de tempo (em segundos) em que um usuário visualiza uma página na internet antes de mudar para outra é uma variável aleatória lognormal, com parâmetros  $\theta = 0,5$  e  $\omega^2 = 1$ .

- Qual é a probabilidade de a página ser vista por mais de dez segundos?

- Durante quanto tempo 50 % dos usuários se movem para outra página?
- Quais são a média e o desvio-padrão do tempo até que um usuário mude a página?

4.10.4 Suponha que  $X$  tenha uma distribuição lognormal e que a média e a variância de  $X$  sejam iguais a 100 e 85.000, respectivamente. Determine os parâmetros  $\theta$  e  $\omega^2$  da distribuição lognormal. [Sugestão: defina  $u = \exp(\theta)$  e  $v = \exp(\omega^2)$  e escreva duas equações em termos de  $u$  e  $v$ .]

4.10.5 O tempo de vida de um *laser* semiconductor tem uma distribuição lognormal, e é sabido que a média e o desvio-padrão do tempo de vida são 10.000 e 20.000, respectivamente.

- Calcule os parâmetros da distribuição lognormal.
- Determine a probabilidade de um tempo de vida exceder 10.000 horas.
- Determine o tempo de vida que é excedido por 90 % dos *lasers*.

**4.10.6** Um artigo em *Health and Population: Perspectives and Issues* (2000, Vol. 23, pp. 28-36) usou a distribuição lognormal para modelar a pressão sanguínea em humanos. A pressão sanguínea sistólica (PSS) média em homens com 17 anos foi de 120,87 mm Hg. Se o coeficiente de variação ( $100\% \times \text{desvio-padrão}/\text{média}$ ) for 9%, quais serão os valores dos parâmetros da distribuição lognormal?

**4.10.7** Suponha que o período (em horas) de permanência em uma emergência seja modelado com uma variável aleatória lognormal  $X$ , com  $\theta = 1,5$  e  $\omega = 0,4$ . Determine o seguinte nos itens (a) e (b):

- Média e variância.
- $P(X < 8)$ .
- Comente a respeito da diferença entre a probabilidade  $P(X < 0)$  calculada a partir dessa distribuição lognormal e da distribuição normal com a mesma média e variância.

**4.10.8** Um artigo publicado na revista *Journal of Hydrology* ["Use of a Lognormal Distribution Model for Estimating Soil Water Retention Curves from Particle-Size Distribution Data" (2006, Vol. 323(1), pp. 325-334)] considerou um modelo de distribuição lognormal para estimar as curvas de retenção de água para uma faixa de texturas do solo. A distribuição de tamanho de partícula (em centímetros) foi modelada como uma variável aleatória lognormal  $X$ , com  $\theta = -3,8$  e  $\omega = 0,7$ . Determine o seguinte:

- $P(X < 0,02)$ .
- O valor para  $x$ , tal que  $P(X \leq x) = 0,95$ .
- A média e a variância de  $X$ .

**4.10.9** Um artigo em *Applied Mathematics and Computation* ["Confidence Intervals for Steady State Availability of a System with Exponential Operating Time and Lognormal Repair Time" (2003, Vol. 137(2), pp. 499-509)] considerou a disponibilidade a longo prazo de um sistema com uma distribuição lognormal suposta para o tempo de reparo. Em determinado exemplo, o tempo de reparo segue uma distribuição lognormal com  $\theta = \omega = 1$ . Determine o seguinte:

- A probabilidade de o tempo de reparo ser maior do que cinco unidades de tempo.
- A probabilidade condicional de um tempo de reparo ser menor do que oito unidades de tempo, dado que ele é maior do que cinco unidades de tempo.
- A média e a variância do tempo de reparo.

**4.10.10** Um artigo em *Chemosphere* ["Statistical Evaluations Reflecting the Skewness in the Distribution of TCDD Levels in Human Adipose Tissue" (1987, Vol. 16(8), pp. 2135-2140)] concluiu que o nível de 2,3,7,8-TCDD (contaminantes ambientais incoloros e persistentes, sem odor distinguível em temperatura ambiente) em tecido adiposo humano tem distribuição lognormal. A média e a variância dessa distribuição lognormal nos Estados Unidos são 8 e 21, respectivamente. Seja  $X$  essa variável aleatória lognormal. Determine o seguinte:

- $P(2000 < X < 2500)$ .
- Valor excedido com probabilidade de 10%.
- A média e a variância de  $X$ .

## Exercícios para a Seção 4.11

**4.11.1** Suponha que  $X$  tenha uma distribuição beta com parâmetros  $\alpha = 2,5$  e  $\beta = 1$ . Determine o seguinte:

- $P(X < 0,25)$ .
- $P(0,25 < X < 0,75)$ .
- Média e variância.

**4.11.2** Suponha que  $X$  tenha uma distribuição beta com parâmetros  $\alpha = 2,5$  e  $\beta = 1$ . Esquematize um gráfico aproximado da função densidade de probabilidade. A função é simétrica?

**4.11.3** Um padrão europeu para uma janela de vidro de baixa emissão usa 0,59 como a proporção de energia solar que entra no ambiente. Suponha que a distribuição da proporção de energia solar que entra no ambiente seja uma variável aleatória beta.

- Calcule a moda, a média e a variância da distribuição para  $\alpha = 3$  e  $\beta = 1,4$ .
- Calcule a moda, a média e a variância da distribuição para  $\alpha = 10$  e  $\beta = 6,25$ .
- Comente a respeito da diferença na dispersão nas distribuições dos itens (a) e (b).

**4.11.4** O período de permanência na emergência de um hospital é a soma dos tempos de atendimento e de espera. Seja  $X$  a proporção do tempo gasto na espera. Considere uma distribuição beta com  $\alpha = 10$  e  $\beta = 1$ . Determine o seguinte:

- $P(X > 0,9)$ .
- $P(X < 0,5)$ .
- Média e variância.

**4.11.5** O tempo máximo para completar uma tarefa em um projeto é de 2,5 dias. Suponha que o tempo para completar essa tarefa, como uma proporção desse máximo, seja uma variável aleatória beta com  $\alpha = 2$  e  $\beta = 3$ . Qual é a probabilidade de necessitar de mais de dois dias para completar a tarefa?

**4.11.6** Um *alelo* é uma forma alternativa de um gene, e a proporção de alelos em uma população é de interesse em genética. Um artigo em *BMC Genetics* ["Calculating Expected DNA Remnants From Ancient Founding Events in Human Population Genetics" (2008, Vol. 9, p. 66)] usou uma distribuição beta com média 0,3 e desvio-padrão 0,17 para modelar proporções iniciais de alelos em uma simulação genética. Determine os parâmetros  $\alpha$  e  $\beta$  para essa distribuição beta.

**4.11.7** Suponha que a construção de uma estação solar seja iniciada. O tempo para completar o projeto não foi estabelecido por causa de incertezas nos recursos financeiros. O tempo para completar a primeira fase é modelado com uma distribuição beta, e os tempos mínimo, mais provável (moda), e máximo para completar a primeira fase são 1,0; 1,25 e 2,0 anos, respectivamente. Além disso, supõe-se que o tempo médio seja igual a  $\mu = 1 + 4(1,25) + 2/6 = 1,333$ . Determine o seguinte, com relação aos itens (a) e (b):

- Parâmetros  $\alpha$  e  $\beta$  da distribuição beta.
- O desvio-padrão da distribuição.
- Esquematize a função densidade de probabilidade.

## Exercícios Suplementares para o Capítulo 4

4.S8 A função densidade de probabilidade do tempo que um analista leva para contar células do sangue, de modo a completar um teste em uma amostra de sangue, é  $f(x) = 0,04$  para  $50 < x < 75$  segundos.

- Que percentagem de testes requer mais de 70 segundos para se completar?
- Que percentagem de testes requer menos de um minuto para se completar?
- Determine a média e a variância do tempo para completar um teste em uma amostra.

4.S9 O tempo que uma célula leva para se dividir (processo chamado de *mitose*) é normalmente distribuído, com um tempo médio de uma hora e um desvio-padrão de cinco minutos.

- Qual é a probabilidade de uma célula se dividir em menos de 45 minutos?
- Qual é a probabilidade de uma célula levar mais de 65 minutos para se dividir?
- Em que tempo, aproximadamente, 99 % de todas as células completaram a mitose?

4.S10 O período de falta ao trabalho em um mês por causa de doenças dos empregados de uma firma é normalmente distribuído, com uma média de 100 horas e desvio-padrão de 20 horas.

- Qual é a probabilidade de esse período no próximo mês estar entre 50 e 80 horas?
- Quanto tempo deveria ser orçado para esse período, se a quantidade orçada dovesse ser excedida com uma probabilidade de somente 10 %?

4.S11 A percentagem de pessoas expostas à bactéria que ficam doentes é 20 %. Considere que pessoas sejam independentes. Considere que 1000 pessoas sejam expostas à bactéria. Aproxime cada uma das seguintes questões:

- Probabilidade de mais de 225 pessoas ficarem doentes.
- Probabilidade de entre 175 e 225 pessoas ficarem doentes.
- Valor tal que seja de 0,01 a probabilidade de o número de pessoas que se tornam doentes exceder o valor.

4.S12 O tempo (em horas) de falha de um *laser* em uma máquina de citometria é modelado por uma distribuição exponencial com  $\lambda = 0,00004$ . Qual é a probabilidade de o tempo até a falha ser

- No mínimo 20.000 horas?
- No máximo 30.000 horas?
- Entre 20.000 e 30.000 horas?

4.S13 Quando um serviço de ônibus reduz tarifas, uma viagem particular de Nova York para Albany, Nova York, torna-se muito popular. Um pequeno ônibus pode transportar quatro passageiros. O tempo entre chamadas para bilhetes é exponencialmente distribuído, com uma média de 30 minutos. Considere que cada chamada compre um bilhete. Qual é a probabilidade de o ônibus lotar em menos de três horas a partir do momento da redução da tarifa?

4.S14 Sabe-se que o tamanho de partículas de prata em uma emulsão fotográfica tem distribuição lognormal, com média de 0,001 mm e um desvio-padrão de 0,002 mm.

- Determine os valores dos parâmetros para a distribuição lognormal.
- Qual é a probabilidade de um tamanho de partícula ser maior do que 0,005 mm?

4.S15 A vida de uma bomba de recirculação segue uma distribuição de Weibull, com parâmetros  $\beta = 2$  e  $\delta = 700$  horas. Determine para os itens (a) e (b):

- A vida média de uma bomba.
- A variância da vida de uma bomba.
- Qual é a probabilidade de uma bomba durar mais do que sua média?

4.S16 O tempo entre chamadas é distribuído exponencialmente, com um tempo médio entre chamadas igual a dez minutos.

- Qual é a probabilidade de o tempo até a primeira chamada ser menor do que cinco minutos?
- Qual é a probabilidade de o tempo até a primeira chamada estar entre 5 e 15 minutos?
- Sendo 0,90 a probabilidade de existir no mínimo uma chamada no intervalo, determine o comprimento de um intervalo de tempo.
- Se não houvesse uma chamada em dez minutos, qual seria a probabilidade de o tempo, até a próxima chamada, ser menor que cinco minutos?
- Qual é a probabilidade de não haver chamadas nos intervalos de 10h a 10h5min, de 11h30min a 11h35min e de 14h a 14h5min?
- Qual é a probabilidade de o tempo até a terceira chamada ser maior do que 30 minutos?
- Qual é o tempo médio até a quinta chamada?

4.S17 Suponha que  $f(x) = 0,5x - 1$  para  $2 < x < 4$ . Determine o seguinte:

- $P(X < 2,5)$
- $P(X > 3)$
- $P(2,5 < X < 3,5)$
- Determine a função de distribuição cumulativa da variável aleatória.
- Determine a média e a variância da variável aleatória.

4.S18 Suponha que  $X$  tenha uma distribuição lognormal, com parâmetros  $\theta = 0$  e  $\omega^2 = 4$ . Determine o seguinte:

- $P(10 < X < 50)$ .
- O valor para  $x$ , tal que  $P(X < x) = 0,05$ .
- A média e a variância de  $X$ .

4.S19 Suponha que  $X$  tenha uma distribuição lognormal e que a média e a variância de  $X$  sejam 50 e 4000, respectivamente. Determine o seguinte:

- Os parâmetros  $\theta$  e  $\omega^2$  da distribuição lognormal.
- A probabilidade de  $X$  ser menor do que 150.

4.S20 Fibras de asbestos, em uma amostra de poeira, são identificadas, por meio de um microscópio eletrônico, depois da preparação da amostra. Suponha que o número de fibras seja uma variável aleatória de Poisson e que o número médio de fibras por centímetro quadrado de poeira na superfície seja 100. Uma amostra, de 800 centímetros quadrados de poeira, é analisada. Considere que determinada célula de amostragem sob o microscópio represente 1/160.000 da amostra.

- Qual é a probabilidade de no mínimo uma fibra ser visível na célula de amostragem?
- Qual é a média do número de células de amostragem que necessitam ser visualizadas para observar dez que contenham fibras?
- Qual é o desvio-padrão do número de células de amostragem que necessitam ser visualizadas para observar dez que contenham fibras?

4.S21 O tempo de espera para um serviço em uma emergência hospitalar segue uma distribuição exponencial, com uma média de três horas. Determine o seguinte:

- Tempo de espera é maior do que quatro horas.
- Tempo de espera é maior do que seis horas, dado que você já esperou duas horas.
- O valor de  $x$  (em horas) excedido com probabilidade de 0,25.

4.S22 O diâmetro de um ponto produzido por uma impressora é normalmente distribuído, com diâmetro médio de 0,002 polegada.

a. Suponha que as especificações requeiram que o diâmetro do ponto esteja entre 0,0014 e 0,0026 polegada. Se a probabilidade de um ponto obedecer às especificações for de 0,9973, qual será o desvio-padrão necessário?

b. Considere que o desvio-padrão do tamanho de um ponto seja 0,0004 polegada. Se a probabilidade de um ponto obedecer às especificações for igual a 0,9973, que especificações serão necessárias? Suponha que as especificações sejam escolhidas simetricamente em torno da média de 0,002.

4.S23 Uma polegada quadrada de um tapete contém 50 fibras de tapete. A probabilidade de existir uma fibra danificada é 0,0001. Considere que as fibras danificadas ocorrem independentemente.

a. Aproxime a probabilidade de uma ou mais fibras danificadas em uma jarda quadrada de tapete.

b. Aproxime a probabilidade de quatro ou mais fibras danificadas em uma jarda quadrada de tapete.

4.S24 Uma companhia aérea faz 200 reservas para um voo que pode conter 185 passageiros. A probabilidade de um passageiro chegar para o voo é 0,9 e os passageiros são considerados independentes.

a. Aproxime a probabilidade de todos os passageiros que chegam poderem sentar-se.

b. Aproxime a probabilidade de que haja assentos vazios.

c. Aproxime o número de reservas que a companhia deveria fazer, de modo que 0,95 seja a probabilidade de todos os passageiros que chegam poderem sentar-se. [Sugestão: tente, sucessivamente, valores para o número de reservas.]

4.S25 Suponha que a construção de uma estação solar seja iniciada. O tempo para completar o projeto não foi estabelecido por causa de incertezas nos recursos financeiros. A proporção de conclusão dentro de um ano tem uma distribuição beta, com parâmetros  $\alpha = 1$  e  $\beta = 5$ . Determine o seguinte:

a. A média e a variância da proporção concluída dentro de um ano.

b. A probabilidade de que mais da metade do projeto seja concluída dentro de um ano.

c. A proporção do projeto que foi concluída dentro de um ano com probabilidade de 0,9.

4.S26 Um artigo publicado no *IEEE Journal on Selected Areas in Communications* ["Impulse Response Modeling of Indoor Radio Propagation Channels" (1993, Vol. 11(7), pp. 967-978)] indicou que o sucesso de um projeto de sistemas internos de comunicação requer caracterização de propagação de rádio. A distribuição da amplitude de componentes individuais multicaminhos foi bem modelada por uma distribuição lognormal. Para uma configuração de teste (com atrasos de 100 ns), a amplitude média foi -24 dB (a partir do pico) com um desvio-padrão de 4,1 dB. A amplitude diminuiu aproximadamente de forma linear, com o aumento de atrasos em excesso. Determine o seguinte:

a. A probabilidade de a amplitude exceder -20 dB.

b. A amplitude excedida com probabilidade de 0,05.

4.S27 Considere o movimento da região da parede transversal do ventrículo direito em pacientes com hipertensão pulmonar (HP). A fração de ejeção (FE) do ventrículo direito é aproximadamente distribuída normalmente, com desvio-padrão de 12 para indivíduos com HP, e com média e desvio-padrão de 56 e 8, respectivamente, para indivíduos-controle.

a. Qual é o valor de FE para indivíduos-controle que é excedido com probabilidade de 99 %?

b. Qual é a média para indivíduos com HP, tal que a probabilidade de seja 1 % de que o FE de um indivíduo com HP seja maior do que o valor no item (a)?

c. Comente quão bem os indivíduos-controle e com HP [com a média determinada no item (b)] podem ser distinguidos por medidas de FE.

4.S28 Forneça esquemas aproximados para as funções densidades de probabilidade beta com os parâmetros a seguir. Comente quaisquer simetrias e mostre quaisquer picos nas funções densidades de probabilidade nos esquemas.

a.  $\alpha = \beta < 1$       b.  $\alpha = \beta = 1$       c.  $\alpha = \beta > 1$

4.S29 Entre os proprietários de casa em uma área metropolitana, 25 % do papel é reciclado a cada semana. Uma companhia que gerencia o lixo serve 10.000 casas (consideradas independentes). Aproxime as seguintes probabilidades:

a. Mais de 2600 papéis reciclados em uma semana.

b. Entre 2400 e 2600 papéis reciclados em uma semana.

c. Número de consumidores que reciclam papel em uma semana que é excedido com uma probabilidade de aproximadamente 0,05.

4.S30 Um artigo publicado no *Journal of Theoretical Biology* ["Computer Model of Growth Cone Behavior and Neuronal Morphogenesis" (1995, Vol. 174(4), pp. 381-389)] desenvolveu um modelo para morfogênese neuronal em que cones de crescimento neuronal têm uma função significativa no desenvolvimento do sistema nervoso. Esse modelo considera que o intervalo de tempo entre a formação de *filopodium* (um processo no comportamento de crescimento de cones) é distribuído exponencialmente, com uma média de seis unidades de tempo. Determine o seguinte:

a. A probabilidade de a formação requerer mais do que nove unidades de tempo.

b. A probabilidade de a formação ocorrer dentro de seis a sete unidades de tempo.

c. O tempo de formação excedido com probabilidade de 0,9.

4.S31 Um artigo em *Electric Power Systems Research* ["On the Self-Scheduling of a Power Producer in Uncertain Trading Environments" (2008, Vol. 78(3), pp. 311-317)] considerou uma abordagem de autoagendamento para um produtor de energia. Além do preço e dos desligamentos forçados, outra incerteza ocorreu em razão de realocações de geração para gerenciar congestionamentos. A realocação de geração foi modelada como  $110X - 60$  (com amplitude [-60,50] MW/h), em que  $X$  tem uma distribuição beta com parâmetros  $\alpha = 3,2$  e  $\beta = 2,8$ . Determine a média e a variância de realocações de geração.

4.S32 Um artigo publicado no *Electronic Journal of Applied Statistical Analysis* ["Survival Analysis of Acute Myocardial Infarction Patients Using Non-Parametric and Parametric Approaches" (2009, Vol. 2(1), pp. 22-36)] descreveu o uso de uma distribuição de Weibull para modelar o tempo de sobrevivência de pacientes que sofreram um infarto agudo do miocárdio (IAM) em um estudo baseado em uma retrospectiva de um hospital. Os parâmetros de forma e de escala para uma distribuição de Weibull no modelo foram 1,16 e 0,25 ano, respectivamente. Determine o seguinte:

a. A média e o desvio-padrão do tempo de sobrevivência.

b. A probabilidade de um paciente sobreviver por mais de um ano.

c. O tempo de sobrevivência que é excedido com uma probabilidade de 0,9.

# Capítulo 5 Exercícios

## Exercícios para a Seção 5.1

5.1.1 Mostre que a seguinte função satisfaz as propriedades de uma função de probabilidade conjunta.

$x$	$y$	$f_{XY}(x, y)$
1,0	1	1/4
1,5	2	1/8
1,5	3	1/4
2,5	4	1/4
3,0	5	1/8

Determine o seguinte:

- a.  $P(X < 2,5, Y < 3)$
- b.  $P(X < 2,5)$
- c.  $P(Y < 3)$
- d.  $P(X > 1,8, Y > 4,7)$
- e.  $E(X), E(Y), V(X)$  e  $V(Y)$
- f. A distribuição de probabilidades marginais de  $X$ .

5.1.2 Mostre que a seguinte função satisfaz as propriedades de uma função de probabilidade conjunta.

$x$	$y$	$f_{XY}(x, y)$
-1,0	-2	1/8
-0,5	-1	1/4
0,5	1	1/2
1,0	2	1/8

Determine o seguinte:

- a.  $P(X < 0,5, Y < 1,5)$
- b.  $P(X < 0,5)$
- c.  $P(Y < 1,5)$
- d.  $P(X > 0,25, Y < 4,5)$
- e.  $E(X), E(Y), V(X)$  e  $V(Y)$
- f. A distribuição de probabilidades marginais de  $X$ .

5.1.3 Na transmissão de informação digital, a probabilidade de um *bit* ter alta, moderada e baixa distorção é 0,01, 0,04 e 0,95, respectivamente. Suponha que três *bits* sejam transmitidos e que a quantidade de distorção de cada *bit* seja considerada independente. Sejam  $X$  e  $Y$  o número de *bits* com alta e moderada distorção, respectivamente. Determine:

- a.  $f_{XY}(x, y)$
- b.  $f_X(x)$
- c.  $E(X)$

5.1.4 Um *site* da internet de pequenos negócios contém 100 páginas, em que 60 %, 30 % e 10 % das páginas contêm baixo, moderado e alto conteúdo gráfico, respectivamente. Uma amostra de quatro páginas é selecionada, sem reposição, e  $X$  e  $Y$  denotam o número de páginas com moderada e alta saída gráfica na amostra. Determine:

- a.  $f_{XY}(x, y)$
- b.  $f_X(x)$
- c.  $E(X)$

5.1.5 Uma fábrica usa dois dispositivos de inspeção para amostrar uma fração de sua saída, com o objetivo de controlar a qualidade. O primeiro monitor de inspeção é capaz de detectar acuradamente 99,3 % dos itens defeituosos que a fábrica recebe, enquanto o segundo dispositivo é capaz de detectar 99,7 % dos casos. Considere que quatro itens defeituosos são produzidos e enviados para inspeção. Sejam  $X$  e  $Y$  o número de itens que serão identificados como defeituosos pelos dispositivos de inspeção 1 e 2, respectivamente. Suponha que os dispositivos sejam independentes. Determine:

- a.  $f_{XY}(x, y)$
- b.  $f_X(x)$
- c.  $E(X)$

5.1.6 Um artigo publicado no *Journal of Database Management* [“Experimental Study of a Self-Tuning Algorithm for DBMS Buffer Pools” (2005, Vol. 16, pp. 1-20)] forneceu a carga de trabalho usada no TPC-C OLTP (Transaction Processing Performance Council’s Version C On-Line Transaction Processing) no termo de referência, que simula uma aplicação típica de ordem de entrada. Veja a tabela a seguir. A frequência de cada tipo de transação (na segunda coluna) pode ser usada como a porcentagem de cada tipo de transação. Sejam  $X$  e  $Y$  o número médio de operações *seleciona* e *atualiza*, respectivamente, requerido para cada tipo de transação. Determine o seguinte:

- a.  $P(X < 5)$
- b.  $E(X)$

5.1.7 Uma classe de estatística para engenheiros tem 40 estudantes, sendo 60 % de engenharia elétrica, 10 % de engenharia industrial e 30 % de engenharia mecânica. Uma amostra de quatro estudantes é selecionada aleatoriamente, sem reposição, para um grupo de projeto. Sejam  $X$  e  $Y$  o número de engenheiros industriais e mecânicos, respectivamente. Determine o seguinte:

- a.  $f_{XY}(x, y)$
- b.  $f_X(x)$
- c.  $E(X)$

5.1.8 Determine o valor de  $c$  que torna a função  $f(x, y) = c(x + y)$  uma função densidade de probabilidade conjunta ao longo da faixa  $0 < x < 3$  e  $x < y < x + 2$ . Determine o seguinte:

- a.  $P(X < 1, Y < 2)$
- b.  $P(1 < X < 2)$
- c.  $P(Y > 1)$
- d.  $P(X < 2, Y < 2)$
- e.  $E(X)$
- f.  $V(X)$
- g. Distribuição de probabilidades marginais de  $X$ .

5.1.9 Determine o valor de  $c$ , tal que a função  $f(x, y) = cxy$  para  $0 < x < 3$  e  $0 < y < 3$  satisfaça as propriedades de uma função densidade de probabilidade conjunta. Determine o seguinte:

- a.  $P(X < 2, Y < 3)$
- b.  $P(X < 2,5)$
- c.  $P(1 < Y < 2,5)$
- d.  $P(X > 1,8, 1 < Y < 2,5)$
- e.  $E(X)$
- f.  $P(X < 0, Y < 4)$
- g. Distribuição de probabilidades marginais de  $X$ .

### Frequências Médias e Operações em TPC-C

Transação	Frequência	Seleciona	Atualiza	Inserir	Apaga	Não Apenas Seleciona	Conecta
Novo pedido	43	23,0	11	12	0	0	0
Pagamento	44	4,2	3	1	0	0,6	0
Status do pedido	4	11,4	0	0	0	0,6	0
Entrega	5	130,0	120	0	10	0	0
Nível do estoque	4	0	0	0	0	0	1

**5.1.10** Determine o valor de  $c$  que torna a função  $f(x, y) = ce^{-2x-3y}$  uma função densidade de probabilidade conjunta ao longo da faixa  $0 < x < y < 2$ . Determine o seguinte:

- a.  $P(X < 1, Y < 2)$       b.  $P(1 < X < 2)$

- c.  $P(Y > 3)$       d.  $P(X < 2, Y < 2)$   
 e.  $E(X)$       f.  $E(Y)$   
 g. Distribuição de probabilidades marginais de  $X$ .

## Exercícios para a Seção 5.2

**5.2.1** A distribuição de probabilidades condicionais de  $Y$ , dado  $X = x$ , é  $f_{Y|X}(y) = xe^{-xy}$  para  $y > 0$ , e a distribuição de probabilidades marginais de  $X$  é uma distribuição contínua uniforme de 0 a 10.

- a. Plote  $f_{Y|X}(y) = xe^{-xy}$  para  $y > 0$  para vários valores de  $x$ .

Determine:

- b.  $P(Y < 2 | X = 2)$       c.  $E(Y | X = 2)$   
 d.  $E(Y | X = x)$       e.  $f_{XY}(x, y)$       f.  $f_Y(y)$

**5.2.2** Considere a distribuição conjunta no Exercício 5.1.1. Determine o seguinte:

- a. Distribuição de probabilidade condicional de  $Y$ , dado que  $X = 1,5$ .  
 b. Distribuição de probabilidade condicional de  $X$ , dado que  $Y = 2$ .  
 c.  $E(Y | X = 1,5)$ .  
 d.  $X$  e  $Y$  são independentes?

**5.2.3** Considere a distribuição conjunta no Exercício 5.1.4. Determine o seguinte:

- a.  $f_{Y|X}(y)$       b.  $E(Y | X = 3)$       c.  $V(Y | X = 3)$

**5.2.4** Um fabricante de lâmpadas de eletroluminescência sabe que a quantidade de tinta luminescente depositada em um de seus produtos é normalmente distribuída, com uma média de 1,2 grama e um desvio-padrão de 0,03 grama. Qualquer lâmpada com menos de 1,14 grama de tinta luminescente não satisfará as especificações dos consumidores. Uma amostra aleatória de 25 lâmpadas é coletada e a massa de tinta luminescente é medida.

- a. Qual é a probabilidade de no mínimo uma lâmpada não satisfazer as especificações?  
 b. Qual é a probabilidade de cinco lâmpadas ou menos não satisfazerem as especificações?  
 c. Qual é a probabilidade de todas as lâmpadas satisfazerem as especificações?  
 d. Por que a distribuição de probabilidades conjuntas das 25 lâmpadas não é necessária para responder às questões prévias?

**5.2.5** Os comprimentos dos eixos menor e maior são usados para resumir partículas de poeira que são aproximadamente elípticas na forma. Sejam  $X$  e  $Y$  os comprimentos (em micrômetros) dos eixos menor e maior, respectivamente. Suponha que  $f_X(x) = \exp(-x)$ ,  $0 < x$  e a distribuição condicional  $f_{Y|X}(y) = \exp[-(y-x)]$ ,  $x < y$ . Responda ou determine o seguinte:

- a. Mostre que  $f_{Y|X}(y)$  é uma função densidade de probabilidade para qualquer valor de  $x$ .  
 b.  $P(X < Y)$ .  
 c. A função densidade de probabilidade conjunta  $f_{XY}(x, y)$ .  
 d. A função densidade de probabilidade condicional de  $X$ , dado  $Y = y$ .  
 e.  $P(Y < 2 | X = 1)$       f.  $E(Y | X = 1)$   
 g.  $P(X < 1, Y < 2)$       h.  $P(Y < 2)$   
 i. c, tal que  $P(Y < c) = 0,9$       j.  $X$  e  $Y$  são independentes?

**5.2.6** Um artigo em *Health Economics* ["Estimation of the Transition Matrix of a Discrete-Time Markov Chain" (2002, Vol. 11, pp. 33-42)] considerou as mudanças na contagem CD4 das células brancas do sangue de um mês para outro. A contagem CD4 é uma medida clínica importante para determinar a severidade de infecções HIV. A contagem CD4 foi agrupada em três categorias distintas: 0-49, 50-74 e  $\geq 75$ . Sejam  $X$  e  $Y$  a contagem CD4 (categoria mínima) em um mês e no mês seguinte, respectivamente. As probabilidades condicionais para  $Y$ , dados valores para  $X$ , foram fornecidas por uma matriz de probabilidade de transição, mostrada na tabela a seguir.

$X$	$Y$		
	0	50	75
0	0,9819	0,0122	0,0059
50	0,1766	0,7517	0,0717
75	0,0237	0,0933	0,8830

Essa tabela é interpretada como segue. Por exemplo,  $P(Y = 50 | X = 75) = 0,0933$ , Suponha também que a distribuição de probabilidade para  $X$  seja  $P(X = 75) = 0,9$ ,  $P(X = 50) = 0,08$ ,  $P(X = 0) = 0,02$ . Determine o seguinte:

- a.  $P(Y \leq 50 | X = 50)$   
 b.  $P(X = 0, Y = 75)$   
 c.  $E(Y | X = 50)$       d.  $f_Y(y)$   
 e.  $f_{XY}(x, y)$       f.  $X$  e  $Y$  são independentes?

**5.2.7** Um artigo em *Clinical Infectious Diseases* ["Strengthening the Supply of Routinely Administered Vaccines in the United States: Problems and Proposed Solutions" (2006, Vol. 42(3), pp. S97-S103)] reportou que vacinas recomendadas para infantes e crianças estavam periodicamente indisponíveis ou escassas nos Estados Unidos. Embora o número de doses demandadas a cada mês seja uma variável aleatória discreta, as grandes demandas podem ser aproximadas por uma distribuição contínua de probabilidade. Suponha que as demandas mensais para duas dessas vacinas, ou seja, sarampo-rubéola-caxumba (SRC) e varicela (para catapora), sejam distribuídas normal e independentemente, com médias de 1,1 e 0,55 milhão de doses e desvios-padrão de 0,3 e 0,1 milhão de doses, respectivamente. Suponha também que os níveis registrados no começo de determinado mês para as vacinas SRC e varicela sejam 1,2 e 0,6 milhão de doses, respectivamente.

- a. Qual é a probabilidade de que não haja falta de nenhuma das vacinas em um mês sem nenhuma produção de vacina?  
 b. Que níveis de estoque devem ser estabelecidos de modo que a probabilidade seja de 90 % para que não haja falta de nenhuma das vacinas em um mês sem produção? Pode existir mais de uma resposta? Explique.

**5.2.8** Os valores das pressões sanguíneas sistólica e diastólica (em mm Hg) são as pressões quando o músculo do coração se contrai e relaxa (denotadas por  $X$  e  $Y$ , respectivamente). Para um conjunto de

indivíduos, a distribuição de pressão diastólica é normal, com média 73 e desvio-padrão 8. A pressão sistólica é distribuída normal e condicionalmente, com média  $1,6x$ , quando  $X = x$  e desvio-padrão de 10. Determine o seguinte:

- a. A função densidade de probabilidade condicional de  $Y$ , dado  $X = 73$ .  
b.  $P(Y < 115 | X = 73)$

c.  $E(Y | X = 73)$

d. Reconheça a distribuição  $f_{XY}(x, y)$  e identifique a média e a variância de  $Y$  e a correlação entre  $X$  e  $Y$ .

5.2.9 Considere a distribuição conjunta no Exercício 5.1.10. Determine o seguinte:

- a.  $f_{Y2}(y)$                       b.  $E(Y | X = 2)$   
c.  $V(Y | X = 2)$                 d.  $X$  e  $Y$  são independentes?

## Exercícios para a Seção 5.3

5.3.1 Para a referência da *Transaction Processing Performance Council* no Exercício 5.1.6, considere  $X$ ,  $Y$  e  $Z$  como o número médio de operações para *selecionar*, *atualizar* e *inserir*, requerido para cada tipo de transação, respectivamente. Calcule o seguinte:

- a.  $f_{XYZ}(x, y, z)$   
b. Função de probabilidade condicional para  $X$  e  $Y$ , dado que  $Z = 0$ .  
c.  $P(X < 6, Y < 6 | Z = 0)$   
d.  $E(X | Y = 0, Z = 0)$

5.3.2 Suponha que as variáveis aleatórias  $X$ ,  $Y$  e  $Z$  tenham a seguinte distribuição de probabilidades conjuntas:

$x$	$y$	$z$	$f(x, y, z)$
1	1	1	0,05
1	1	2	0,10
1	2	1	0,15
1	2	2	0,20
2	1	1	0,20
2	1	2	0,15
2	2	1	0,10
2	2	2	0,05

Determine o seguinte:

- a.  $P(X = 2)$                       b.  $P(X = 1, Y = 2)$   
c.  $P(Z < 1,5)$                   d.  $P(X = 1 \text{ ou } Z = 2)$   
e.  $E(X)$                           f.  $P(X = 1 | Y = 1)$   
g.  $P(X = 1, Y = 1 | Z = 2)$     h.  $P(X = 1 | Y = 1, Z = 2)$   
i. Distribuição de probabilidade conjunta de  $X$ , dado que  $Y = 1$  e  $Z = 2$ .

5.3.3 Suponha que as variáveis aleatórias  $X$ ,  $Y$  e  $Z$  tenham a distribuição de probabilidades conjuntas  $f(x, y, z) = 8xyz$  para  $0 < x < 1$ ,  $0 < y < 1$  e  $0 < z < 1$ . Determine o seguinte:

- a.  $P(X < 0,5)$                   b.  $P(X < 0,5, Y < 0,5)$   
c.  $P(Z < 2)$                   d.  $P(X < 0,5 \text{ ou } Z < 2)$   
e.  $E(X)$                         f.  $P(X < 0,5 | Y = 0,5)$   
g.  $P(X < 0,5, Y < 0,5 | Z = 0,8)$   
h. Distribuição de probabilidade condicional de  $X$ , dado que  $Y = 0,5$  e  $Z = 0,8$ .  
i.  $P(X < 0,5 | Y = 0,5, Z = 0,8)$

5.3.4 Suponha que as variáveis aleatórias  $X$ ,  $Y$  e  $Z$  tenham a distribuição de probabilidades conjuntas  $f_{XYZ}(x, y, z) = c$  sobre o cilindro  $x^2 + y^2 < 4$  e  $0 < z < 4$ . Determine a constante  $c$ , de modo que  $f_{XYZ}(x, y, z)$  seja uma função densidade de probabilidade. Determine o seguinte:

- a.  $P(X^2 + Y^2 < 2)$               b.  $P(Z < 2)$   
c.  $E(X)$                           d.  $P(X < 1 | Y = 1)$   
e.  $P(X^2 + Y^2 < 1 | Z = 1)$   
f. Distribuição de probabilidade condicional de  $Z$ , dado que  $X = 1$  e  $Y = 1$ .

## Exercícios para a Seção 5.4

5.4.1 Determine a covariância e a correlação para a seguinte distribuição de probabilidade conjunta:

$x$	1	1	2	4
$y$	3	4	5	6
$f_{XY}(x, y)$	1/8	1/4	1/2	1/8

5.4.2 Determine o valor de  $c$ , a covariância e a correlação para a função de probabilidade conjunta  $f_{XY}(x, y) = c(x + y)$  para  $x = 1, 2, 3$  e  $y = 1, 2, 3$ .

5.4.3 Um tratamento à base de drogas é administrado a pacientes que são então avaliados. Os sintomas de melhoria, piora ou permanência no mesmo estado têm probabilidades de 0,4, 0,1 e 0,5, respectivamente. Considere que quatro pacientes sejam

tratados e que  $X$  e  $Y$  sejam o número de pacientes que melhoram ou pioram, respectivamente.  $X$  e  $Y$  são independentes? Calcule a covariância e a correlação entre  $X$  e  $Y$ .

5.4.4 Para o termo de referência do Conselho de Desempenho do Processamento de Transação do Exercício 5.1.6, sejam  $X$ ,  $Y$  e  $Z$  o número médio das operações de *seleciona*, *atualiza* e *insere* requerido para cada tipo de transação, respectivamente. Calcule o seguinte:

- a. Correlação entre  $X$  e  $Y$ .  
b. Correlação entre  $X$  e  $Z$ .

5.4.5 Indique o valor de  $c$ , a covariância e a correlação para a função densidade de probabilidade conjunta  $f_{XY}(x, y) = cxy$  para a faixa de  $0 < x < 3$  e  $0 < y < x$ .

5.4.6 Determine a covariância e a correlação para a função densidade de probabilidade conjunta,  $f_{XY}(x, y) = e^{-x-y}$  para a faixa de  $0 < x$  e  $0 < y$ .

5.4.7 A distribuição de probabilidade conjunta é

$x$	-1	0	0	1
$y$	0	-1	1	0
$f_{XY}(x, y)$	1/4	1/4	1/4	1/4

Mostre que a correlação entre  $X$  e  $Y$  é zero, porém  $X$  e  $Y$  não são independentes.

5.4.8 Determine a covariância e a correlação para as contagens CD4 em um mês e no mês seguinte no Exercício 5.2.6.

5.4.9 Determine a covariância e a correlação para os comprimentos do menor e do maior eixo no Exercício 5.2.5.

5.4.10 Suponha que a correlação entre  $X$  e  $Y$  seja  $\rho$ . Para as constantes  $a, b, c$  e  $d$ , qual é a correlação entre as variáveis aleatórias  $U = aX + b$  e  $V = cY + d$ ?

## Exercícios para a Seção 5.5

5.5.1 Com base no número de vazios, uma placa de óxido de ferro é classificada como alta, média ou baixa. Historicamente, 5 % das placas são classificadas como altas, 85 % como médias e 10 % como baixas. Uma amostra de 20 placas é selecionada para teste. Sejam  $X$ ,  $Y$  e  $Z$  o número de placas que são independentemente classificadas como altas, médias e baixas, respectivamente.

- Quais são o nome e os valores dos parâmetros da distribuição de probabilidades de  $X$ ,  $Y$  e  $Z$ ?
  - Qual é a faixa da distribuição de probabilidades conjuntas de  $X$ ,  $Y$  e  $Z$ ?
  - Quais são o nome e os valores dos parâmetros da distribuição de probabilidades marginais de  $X$ ?
  - Determine  $E(X)$  e  $V(X)$ .
- Determine o seguinte:
- $P(X = 1, Y = 17, Z = 3)$
  - $P(X \leq 1, Y = 17, Z = 3)$
  - $P(X \leq 1)$
  - $E(Y)$
  - $P(X = 2, Z = 3 | Y = 17)$
  - $P(X = 2 | Y = 17)$
  - $E(X | Y = 17)$

5.5.2 Um *site* da internet usa anúncios para guiar visitantes para uma de quatro páginas de destino (*landing pages* – páginas a que se chega mediante um *site*, quando se clica em um resultado de busca ou em um anúncio de patrocinadores). As probabilidades para cada página de destino são iguais. Considere 20 visitantes independentes e sejam as variáveis aleatórias  $W$ ,  $X$ ,  $Y$  e  $Z$  o número de visitantes guiados para cada página. Calcule o seguinte:

- $P(W = 5, X = 5, Y = 5, Z = 5)$
- $P(W = 5, X = 5, Y = 5, Z = 5)$
- $P(W = 7, X = 7, Y = 6 | Z = 3)$
- $P(W = 7, X = 7, Y = 3 | Z = 3)$
- $P(W \leq 2)$
- $E(W)$
- $P(W = 5, X = 5)$
- $P(W = 5 | X = 5)$

5.5.3 Suponha que  $X$  e  $Y$  tenham uma distribuição normal bivariada, com  $\sigma_x = 0,04$ ,  $\sigma_y = 0,08$ ,  $\mu_x = 3,00$ ,  $\mu_y = 7,70$  e  $\rho = 0$ .

Determine o seguinte:

- $P(2,95 < X < 3,05)$
- $P(7,60 < Y < 7,80)$
- $P(2,95 < X < 3,05, 7,60 < Y < 7,80)$

5.5.4 Sejam  $X$  e  $Y$  a concentração e a viscosidade de um produto químico. Suponha que  $X$  e  $Y$  tenham distribuição normal bivariada, com  $\sigma_x = 4$ ,  $\sigma_y = 1$ ,  $\mu_x = 2$ ,  $\mu_y = 1$ . Desenhe, aproximadamente, uma curva

de nível da função densidade de probabilidade conjunta para cada um dos seguintes valores de  $\rho$ :

- $\rho = 0$
- $\rho = 0,8$
- $\rho = -0,8$

5.5.5 Na fabricação de lâmpadas de eletroluminescência, várias camadas diferentes de tinta são depositadas em um substrato plástico. A espessura dessas camadas é crítica se especificações relativas à cor final e à intensidade de luz devem ser obedecidas. Sejam  $X$  e  $Y$  a espessura de duas camadas diferentes de tinta. Sabe-se que  $X$  é normalmente distribuída, com uma média de 0,1 milímetro e um desvio-padrão de 0,00031 milímetro, e  $Y$  é normalmente distribuída, com uma média de 0,23 milímetro e desvio-padrão de 0,00017 milímetro. O valor de  $\rho$  para essas variáveis é igual a zero. As especificações exigem que uma lâmpada tenha uma espessura de tinta correspondendo a  $X$  na faixa de 0,099535 a 0,100465 milímetro e  $Y$  na faixa de 0,22966 a 0,23034 milímetro. Qual é a probabilidade de uma lâmpada selecionada aleatoriamente obedecer às especificações?

5.5.6 Em uma titulação ácido-base, uma base ou um ácido é gradualmente adicionada(o) ao outro até que eles estejam completamente neutralizados. Sejam  $X$  e  $Y$  os mililitros de ácido e de base necessários para a equivalência, respectivamente. Considere que  $X$  e  $Y$  têm uma distribuição normal bivariada, com  $\sigma_x = 5$  mL,  $\sigma_y = 2$  mL,  $\mu_x = 120$  mL,  $\mu_y = 100$  mL e  $\rho = 0,6$ .

Determine o seguinte:

- Covariância entre  $X$  e  $Y$ .
- Distribuição de probabilidades marginais de  $X$ .
- $P(X < 116)$
- Distribuição de probabilidades condicionais de  $X$ , dado que  $Y = 102$ .
- $P(X < 116 | Y = 102)$

5.5.7 Se  $X$  e  $Y$  têm uma distribuição normal bivariada com  $\rho = 0$ , mostre que  $X$  e  $Y$  são independentes.

5.5.8 As probabilidades de pacientes que tomam medicamentos como terapia melhorarem, piorarem ou permanecerem no mesmo estado são iguais a 0,5, 0,4 e 0,1, respectivamente. Considere que 20 pacientes (supostos independentes) sejam tratados com essa terapia. Sejam  $X_1$ ,  $X_2$  e  $X_3$  o número de pacientes que melhoram, permanecem no mesmo estado ou pioram, respectivamente. Determine o seguinte:

- $X_1, X_2$  e  $X_3$  são independentes?
- $P(X_1 = 10)$
- $P(X_1 = 10, X_2 = 8, X_3 = 2)$
- $P(X_1 = 5 | X_2 = 12)$
- $E(X_1)$

## Exercícios para a Seção 5.6

5.6.1  $X$  e  $Y$  são variáveis aleatórias normais e independentes, com  $E(X) = 2$ ,  $V(X) = 5$ ,  $E(Y) = 6$  e  $V(Y) = 8$ .

Determine o seguinte:

- a.  $E(3X + 2Y)$                       b.  $V(3X + 2Y)$   
 c.  $P(3X + 2Y < 18)$                   d.  $P(3X + 2Y < 28)$

5.6.2 Suponha que a variável aleatória  $X$  represente o comprimento, em centímetros, de uma peça perfurada. Seja  $Y$  o comprimento da peça em milímetros. Se  $E(X) = 5$  e  $V(X) = 0,25$ , quais serão os valores da média e da variância de  $Y$ ?

5.6.3 Na fabricação de lâmpadas de eletroluminescência, várias camadas diferentes de tinta são depositadas em um substrato plástico. A espessura dessas camadas será crítica se especificações relativas à cor final e à intensidade de luz devem ser obedecidas. Sejam  $X$  e  $Y$  as espessuras de duas camadas diferentes de tinta. Sabe-se que  $X$  é normalmente distribuída, com média de 0,1 mm e desvio-padrão de 0,00031 mm, e  $Y$  é também normalmente distribuída, com média de 0,23 mm e desvio-padrão de 0,00017 mm. Considere que essas variáveis sejam independentes.

a. Se determinada lâmpada for composta somente por essas duas tintas, qual será a probabilidade de a espessura total de tinta ser menor que 0,2337 mm?

b. Uma lâmpada com uma espessura total excedendo 0,2405 mm não tem uma uniformidade de cor requerida pelo consumidor. Encontre a probabilidade de uma lâmpada selecionada de forma aleatória não satisfazer as especificações do consumidor.

5.6.4 A confecção de potes de cerâmica geralmente é realizada em duas etapas: a elaboração do pote nos tornos de oleiro e o cozimento. O tempo de giro do torno e do cozimento são variáveis aleatórias normalmente distribuídas, com médias de 40 minutos e 60 minutos e desvios-padrão de dois minutos e três minutos, respectivamente.

a. Qual é a probabilidade de uma peça de cerâmica ser confeccionada em 95 minutos?

b. Qual é a probabilidade de a peça ser confeccionada em mais de 110 minutos?

5.6.5 Um artigo publicado em *Knee Surgery Sports Traumatology, Arthroscopy* ["Effect of Provider Volume on Resource Utilization for Surgical Procedures" (2005, Vol. 13, pp. 273-279)] mostrou um tempo médio de 129 minutos e um desvio-padrão de 14 minutos para a cirurgia de reconstrução ACL em hospitais com alto volume de cirurgias (com mais de 300 de tais cirurgias por ano). Se um hospital com um grande volume de cirurgias necessitar programar dez cirurgias, qual será a média e qual a variância do tempo total para completar essas cirurgias? Considere os tempos de cirurgias como independentes e normalmente distribuídos.

5.6.6 Latas de refrigerantes são cheias por uma máquina de enchimento automático, sendo o desvio-padrão igual a 0,5 onça fluida. Considere que os volumes de enchimento das latas sejam variáveis aleatórias normais independentes.

a. Qual é o desvio-padrão do volume médio de enchimento de 100 latas?

b. Se o volume médio de enchimento for 12,1 onças, qual será a probabilidade de o volume médio de enchimento das 100 latas estar abaixo de 12 onças fluidas?

c. Sendo 0,005 a probabilidade de que a média de 100 latas esteja abaixo de 12 onças fluidas, qual deve ser o volume médio de enchimento?

d. Se o volume médio de enchimento for 12,1 onças, qual deverá ser o desvio-padrão do volume de enchimento, para que a

probabilidade de a média de 100 latas estar abaixo de 12 onças fluidas seja igual a 0,005?

e. Determine o número de latas que precisamos medir, para que a probabilidade de a média de 100 latas estar abaixo de 12 onças fluidas seja igual a 0,01.

5.6.7 Considere que os pesos de indivíduos sejam independentes e normalmente distribuídos, com uma média de 160 libras e um desvio-padrão de 30 libras. Suponha que 25 pessoas se espremam em um elevador que seja projetado para suportar 4300 libras.

a. Qual é a probabilidade de a carga (peso total) exceder o limite de projeto?

b. Sendo a probabilidade igual a 0,0001 de que o limite seja excedido, qual é o limite de projeto excedido por 25 ocupantes?

5.6.8 Um componente em forma de U deve ser formado a partir de três peças  $A$ ,  $B$  e  $C$ . O esquema está mostrado na Figura 5.14. O comprimento de  $A$  é normalmente distribuído, com uma média de 10 mm e um desvio-padrão de 0,1 mm. A espessura das peças  $B$  e  $C$  é distribuída normalmente, com uma média de 2 mm e desvio-padrão de 0,05 mm. Suponha que todas as dimensões sejam independentes.

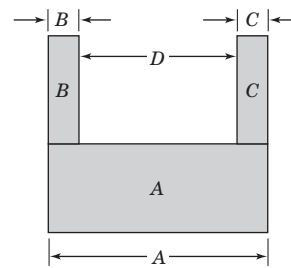


FIGURA 5.14

Ilustração para o componente em forma de U.

a. Determine a média e o desvio-padrão do comprimento da abertura  $D$ .

b. Qual é a probabilidade de a abertura  $D$  ser menor do que 5,9 milímetros?

5.6.9 Considere o perímetro de uma peça no Exemplo 5.27. Sejam  $X_1$  e  $X_2$  o comprimento e a largura de uma peça com desvios-padrão de 0,1 e 0,2 centímetro, respectivamente. Suponha que a covariância entre  $X_1$  e  $X_2$  seja de 0,02. Determine a variância do perímetro  $Y = 2X_1 + 2X_2$  de uma peça. Compare e comente o resultado aqui e no exemplo.

5.6.10 A taxa de retorno de um ativo é a variação no preço dividida pelo preço inicial (denotada por  $r$ ). Suponha que R\$ 10.000,00 sejam usados para comprar cotas em três ações com taxas de retorno iguais a  $X_1$ ,  $X_2$  e  $X_3$ . Inicialmente, R\$ 2500,00, R\$ 3000,00 e R\$ 4500,00 são alocados em cada ação, respectivamente. Depois de um ano, a distribuição da taxa de retorno para cada ação é normalmente distribuída com os seguintes parâmetros:  $\mu_1 = 0,12$ ,  $\sigma_1 = 0,14$ ,  $\mu_2 = 0,04$ ,  $\sigma_2 = 0,02$ ,  $\mu_3 = 0,07$ ,  $\sigma_3 = 0,08$ .

a. Admita que essas taxas de retorno sejam independentes. Determine a média e a variância da taxa de retorno, depois de um ano, para o investimento inteiro de R\$ 10.000,00.

b. Considere que  $X_1$  seja independente de  $X_2$  e  $X_3$ , mas que a covariância entre  $X_2$  e  $X_3$  seja  $-0,005$ . Repita o item (a).

c. Compare as médias e as variâncias obtidas nos itens (a) e (b) e comente sobre quaisquer benefícios das covariâncias negativas entre os ativos.

**5.6.11** No Exercício 5.2.7, a demanda mensal para vacina SRC foi admitida como tendo uma distribuição aproximadamente normal, com média e desvio-padrão de 1,1 e 0,3 milhão de doses, respectivamente. Suponha que as demandas para meses diferentes sejam independentes e que  $Z$  seja a demanda para um ano (em milhões de doses). Determine o seguinte:

- Média, variância e distribuição de  $Z$ .
- $P(Z < 13,2)$
- $P(11 < Z < 15)$
- Valor para  $c$  tal que  $P(Z < c) = 0,99$ .

## Exercícios para a Seção 5.7

**5.7.1** Suponha que  $X$  seja uma variável aleatória com distribuição de probabilidades

$$f_X(x) = 1/4, \quad x = 1, 2, 3, 4$$

Determine a distribuição de probabilidades de  $Y = 2X + 1$ .

**5.7.2** Suponha que  $X$  seja uma variável aleatória contínua com distribuição de probabilidades

$$f_X(x) = \frac{x}{18}, \quad 0 \leq x \leq 6$$

- Determine a distribuição de probabilidades da variável aleatória  $Y = 2X + 10$ .
  - Determine o valor esperado de  $Y$ .
- 5.7.3** Uma variável aleatória  $X$  tem a distribuição de probabilidades

$$f_X(x) = e^{-x}, \quad x \geq 0$$

Determine a distribuição de probabilidades para o seguinte:

- $Y = X^2$
- $Y = X^{1/2}$
- $Y = \ln X$

**5.7.4** A velocidade de uma partícula em um gás é uma variável aleatória  $V$ , com distribuição de probabilidades

$$f_V(v) = av^2 e^{-bv}, \quad v > 0$$

em que  $b$  é uma constante que depende da temperatura do gás e da massa da partícula.

- Determine o valor da constante  $a$ .
  - A energia cinética da partícula é  $W = mV^2/2$ . Determine a distribuição de probabilidades de  $W$ .
- 5.7.5** Suponha que  $X$  tenha uma distribuição de probabilidades

$$f_X(x) = 1, \quad 1 \leq x \leq 2$$

Determine a distribuição de probabilidades da variável aleatória  $Y = e^X$ .

**5.7.6** Deduza a função densidade de probabilidade para uma variável aleatória lognormal  $Y$  a partir da relação  $Y = \exp(W)$  para uma variável aleatória normal  $W$  com média  $\theta$  e variância  $\omega^2$ .

**5.7.7** Um avião está voando a uma altitude constante com velocidade igual a  $r_1$  (relativa ao ar) e ângulo  $\theta_1$  (em um sistema bidimensional de coordenadas). A magnitude e a direção do vento são  $r_2$  e  $\theta_2$ , respectivamente. Suponha que o ângulo do vento seja uniformemente distribuído entre 10 e 20 graus e que todos os outros parâmetros sejam constantes. Determine a função densidade de probabilidade da magnitude do vetor resultante  $r = [r_1^2 + r_2^2 + 2r_1r_2(\cos \theta_1 - \cos \theta_2)]^{0,5}$ .

**5.7.8** Medidores de potência capacitam ciclistas a obterem medidas de potência aproximadamente de forma contínua. Os medidores também calculam a potência média gerada ao longo de um intervalo de tempo. Corredores profissionais podem gerar 6,6 watts por quilograma de massa de corpo por períodos estendidos de tempo. Alguns medidores calculam uma medida de potência normalizada de modo a ajustar o esforço fisiológico requerido quando a potência varia frequentemente. Seja  $X$  a variável aleatória que denota a potência em um tempo de medida e suponha que  $X$  tenha uma distribuição lognormal com parâmetros  $\theta = 5,2933$  e  $\omega^2 = 0,00995$ . A potência normalizada é calculada como a quarta potência da média de  $Y = X^4$ . Determine o seguinte:

- Média e desvio-padrão de  $X$ .
- $f_Y(y)$ .
- Média e variância de  $Y$ .
- Quarta raiz da média de  $Y$ .
- Compare  $[E(X^4)]^{1/4}$  com  $E(X)$  e comente.

## Exercícios para a Seção 5.8

**5.8.1** Uma variável aleatória  $X$  tem a distribuição discreta uniforme

$$f(x) = \frac{1}{m}, \quad x = 1, 2, \dots, m$$

- Mostre que a função geradora de momento é

$$M_X(t) = \frac{e^t(1 - e^{tm})}{m(1 - e^t)}$$

- Use  $M_X(t)$  para encontrar a média e a variância de  $X$ .

**5.8.2** Uma variável aleatória  $X$  tem a distribuição de Poisson

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, \dots$$

- Mostre que a função geradora de momento é

$$M_X(t) = e^{\lambda(e^t - 1)}$$

- Use  $M_X(t)$  para encontrar a média e a variância da variável aleatória de Poisson.

5.8.3 A variável aleatória geométrica  $X$  tem a distribuição de probabilidade

$$f(x) = (1 - p)^{x-1}p, \quad x = 1, 2, \dots$$

a. Mostre que a função geradora de momento é

$$M_X(t) = \frac{pe^t}{1 - (1 - p)e^t}$$

b. Use  $M_X(t)$  para encontrar a média e a variância de  $X$ .

5.8.4 A variável aleatória qui-quadrado, com  $k$  graus de liberdade, tem a função geradora de momento  $M_X(t) = (1 - 2t)^{-k/2}$ . Suponha que  $X_1$  e  $X_2$  sejam variáveis aleatórias independentes do tipo qui-quadrado, com  $k_1$  e  $k_2$  graus de liberdade, respectivamente. Qual é a distribuição de  $Y = X_1 + X_2$ ?

5.8.5 Uma variável aleatória contínua  $X$  tem a seguinte função de probabilidade:

$$f(x) = 4xe^{-2x}, \quad x > 0$$

a. Encontre a função geradora de momento para  $X$ .

b. Encontre a média e a variância de  $X$ .

5.8.6 Uma variável aleatória contínua uniforme  $X$  tem função de densidade

$$f(x) = \frac{1}{\beta - \alpha}, \quad \alpha \leq x \leq \beta$$

a. Mostre que a função geradora de momento é

$$M_X(t) = \frac{e^{t\beta} - e^{t\alpha}}{t(\beta - \alpha)}$$

b. Use  $M_X(t)$  para encontrar a média e a variância de  $X$ .

5.8.7 Uma variável aleatória  $X$  tem a distribuição exponencial

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0$$

a. Mostre que a função geradora de momento de  $X$ , para  $t < \lambda$ , é

$$M_X(t) = \left(1 - \frac{t}{\lambda}\right)$$

b. Encontre a média e a variância de  $X$ .

5.8.8 Uma variável aleatória  $X$  tem a distribuição gama

$$f(x) = \frac{\lambda^r}{\Gamma(r)} (\lambda x)^{r-1} e^{-\lambda x}, \quad x > 0$$

a. Mostre que a função geradora de momento de  $X$ , para  $t < \lambda$ , é

$$M_X(t) = \left(1 - \frac{t}{\lambda}\right)^{-r}$$

b. Encontre a média e a variância de  $X$ .

5.8.9 Sejam  $X_1, X_2, \dots, X_r$  variáveis aleatórias independentes exponenciais com parâmetro  $\lambda$ .

a. Encontre a função geradora de momento de  $Y = X_1 + X_2 + \dots + X_r$ .

b. Qual é a distribuição da variável aleatória  $Y$ ?

5.8.10 Suponha que  $X_i$  tenha uma distribuição normal com média  $\mu_i$  e variância,  $i = 1, 2$ . Sejam  $X_1$  e  $X_2$  independentes.

a. Encontre a função geradora de momento de  $Y = X_1 + X_2$ .

b. Qual é a distribuição da variável aleatória  $Y$ ?

## Exercícios Suplementares para o Capítulo 5

5.S11 Mostre que a seguinte função satisfaz as propriedades de uma função de probabilidade conjunta:

$x$	$y$	$f(x, y)$
0	0	1/4
0	1	1/8
1	0	1/8
1	1	1/4
2	2	1/4

Determine o seguinte:

- a.  $P(X < 0,5, Y < 1,5)$       b.  $P(X \leq 1)$   
 c.  $P(X < 1,5)$                   d.  $P(X > 0,5, Y < 1,5)$

e. Determine  $E(X)$ ,  $E(Y)$ ,  $V(X)$  e  $V(Y)$ .

f. A distribuição de probabilidades marginais da variável aleatória  $X$ .

g. A distribuição de probabilidades condicionais de  $Y$ , dado que  $X = 1$ .

h.  $E(Y | X = 1)$ .

i.  $X$  e  $Y$  são independentes? Por que sim ou por que não?

j. A correlação entre  $X$  e  $Y$ .

5.S12 A percentagem de pessoas que tomam um medicamento contra reumatismo e que com isso sofrem efeitos colaterais severos, moderados ou menores é igual a 10, 20 e 70 %, respectivamente. Considere

que as pessoas reagem de forma independente e que 20 pessoas tomam medicamento. Determine o seguinte:

a. A probabilidade de 2, 4 e 14 pessoas sofrerem efeitos colaterais severos, moderados e menores, respectivamente.

b. A probabilidade de ninguém sofrer efeitos colaterais severos.

c. A média e a variância do número de pessoas que sofrerão efeitos colaterais severos.

d. A distribuição de probabilidades condicionais do número de pessoas que sofrem efeitos colaterais severos, dado que 19 sofrem efeitos colaterais menores.

e. A média condicional do número de pessoas que sofrem efeitos colaterais severos, dado que 19 sofrem efeitos colaterais menores.

5.S13 Determine o valor de  $c$ , tal que a função  $f(x, y) = cx^2y$ , para  $0 < x < 3$  e  $0 < y < 2$ , satisfaz as propriedades de uma função densidade de probabilidade conjunta.

Determine o seguinte:

- a.  $P(X < 1, Y < 1)$                   b.  $P(X < 2,5)$   
 c.  $P(1 < Y < 2,5)$                   d.  $P(X > 2, 1 < Y < 1,5)$   
 e.  $E(X)$                                   f.  $E(Y)$

g. A distribuição de probabilidades marginais da variável aleatória  $X$ .

h. A distribuição de probabilidades condicionais de  $Y$ , dado que  $X = 1$ .

i. A distribuição de probabilidades condicionais de  $X$ , dado que  $Y = 1$ .

5.S14 A distribuição conjunta das variáveis aleatórias contínuas  $X$ ,  $Y$  e  $Z$  é constante ao longo da região  $x^2 + y^2 \leq 1$ ,  $0 < z < 4$ .

Determine o seguinte:

- a.  $P(X^2 + Y^2 \leq 0,5)$       b.  $P(X^2 + Y^2 \leq 0,5, Z < 2)$   
 c. A função densidade de probabilidade condicional conjunta de  $X$  e  $Y$ , considerando que  $Z = 1$ .  
 d. A função densidade de probabilidade marginal de  $X$ .  
 e. A média condicional de  $Z$ , dado que  $X = 0$  e  $Y = 0$ .  
 f. A média condicional de  $Z$ , dado que  $X = x$  e  $Y = y$ .

5.S15 Problemas de contaminação na fabricação de semicondutores podem resultar em um defeito funcional, em um defeito menor, ou em nenhum defeito no produto final. Suponha que 20 %, 50 % e 30 % dos problemas de contaminação resultem em defeitos funcionais, em defeitos menores e em nenhum defeito, respectivamente. Considere que os defeitos de dez problemas de contaminação sejam independentes.

- a. Qual é a probabilidade de os dez problemas de contaminação resultarem em dois defeitos funcionais e cinco defeitos menores?  
 b. Qual é a distribuição do número de problemas de contaminação que resultam em nenhum defeito?  
 c. Qual é o número esperado de problemas de contaminação que resultam em nenhum defeito?

5.S16 O peso de tijolos de barro para construção é distribuído normalmente, com uma média de três libras e um desvio-padrão de 0,25 libra. Suponha que os pesos dos tijolos sejam independentes e que uma amostra aleatória de 25 tijolos seja escolhida.

- a. Qual é a probabilidade de o peso médio da amostra ser menor que 2,95 libras?  
 b. Sendo a probabilidade igual a 0,99, que valor o peso médio excederá?

5.S17 O tempo para um sistema automatizado em um depósito localizar uma peça é distribuído normalmente, com uma média de 45 segundos e desvio-padrão de 30 segundos. Suponha que pedidos independentes sejam feitos para dez peças.

- a. Qual é a probabilidade de que o tempo médio para localizar dez peças exceda 60 segundos?  
 b. Qual é a probabilidade de que o tempo total para localizar dez peças exceda 600 segundos?

5.S18 O peso de um pequeno bombom é normalmente distribuído, com uma média de 0,1 onça e um desvio-padrão de 0,01 onça. Suponha que 16 bombons sejam colocados em um pacote e que os pesos sejam independentes.

- a. Quais são a média e a variância do peso líquido do pacote?  
 b. Qual é a probabilidade de o peso líquido de um pacote ser menor que 1,6 onça?  
 c. Qual valor o peso médio excederá com probabilidade de 0,99?

5.S19 Suponha que  $X$  e  $Y$  tenham distribuição normal bivariada, com  $\sigma_x = 4$ ,  $\sigma_y = 1$ ,  $\mu_x = 2$ ,  $\mu_y = 4$  e  $\rho = -0,2$ . Desenhe uma curva de nível aproximada da função densidade de probabilidade conjunta.

5.S20 Um arranjo mecânico, usado em um motor de automóvel, contém quatro componentes principais. Os pesos dos componentes são independentes e normalmente distribuídos, com as seguintes médias e desvios-padrão (em onças):

Componente	Média	Desvio-Padrão
Capa esquerda	4,0	0,4
Capa direita	5,5	0,5
Conjunto de rolamentos	10,0	0,2
Conjunto de parafusos	8,0	0,5

- a. Qual é a probabilidade de o peso de um arranjo exceder 29,5 onças?  
 b. Qual é a probabilidade de o peso médio de oito arranjos independentes exceder 29 onças?

5.S21 A permeabilidade de uma membrana, utilizada como barreira de umidade em uma aplicação biológica, depende da espessura de duas camadas integradas. As camadas são normalmente distribuídas, com médias de 0,5 e 1 milímetro, respectivamente. Os desvios-padrão da espessura das camadas são 0,1 e 0,2 milímetro, respectivamente. A correlação entre as camadas é 0,7.

- a. Determine a média e a variância da espessura total das duas camadas.  
 b. Qual é a probabilidade de a espessura total ser menor do que 1 milímetro?  
 c. Sejam  $X_1$  e  $X_2$  a espessura das camadas 1 e 2, respectivamente. Uma medida do desempenho da membrana é uma função  $2X_1 + 3X_2$  da espessura. Determine a média e a variância dessa medida de desempenho.

5.S22 Uma pequena companhia deve decidir que investimentos usar para o dinheiro gerado a partir de operações financeiras. Cada investimento tem uma média e um desvio-padrão associado ao ganho percentual. O primeiro investimento tem um ganho percentual médio de 5 %, com um desvio-padrão de 2 %, e o segundo investimento tem a mesma média de 5 %, com um desvio-padrão de 4 %. Os investimentos têm uma correlação de  $-0,5$ . Se a companhia investe dois milhões de dólares, metade em cada investimento, quais são a média e o desvio-padrão do percentual de retorno? Compare o desvio-padrão dessa estratégia com aquele que investe os dois milhões de dólares somente no primeiro investimento.

5.S23 Uma companhia de *marketing* realizou uma análise de risco para um fabricante de fibras sintéticas e concluiu que novos competidores não apresentam risco 13 % das vezes (na maioria, por causa da diversidade de fibras fabricadas), risco moderado 72 % das vezes (alguma sobreposição de produtos) e risco muito alto (competidores fabricam exatamente os mesmos produtos) 15 % das vezes. Sabe-se que 12 companhias internacionais estão planejando abrir novas fábricas para produzir fibras sintéticas, dentro dos próximos três anos. Considere as companhias como independentes. Sejam  $X$ ,  $Y$  e  $Z$  o número de novos competidores que representarão risco nulo, moderado e muito alto para a companhia interessada, respectivamente.

Determine o seguinte:

- a. Faixa da distribuição de probabilidades conjuntas de  $X$ ,  $Y$  e  $Z$ .  
 b.  $P(X = 1, Y = 3, Z = 1)$       c.  $P(Z \leq 2)$   
 d.  $P(Z = 2 | Y = 1, X = 10)$       e.  $P(Z \leq 1 | X = 10)$   
 f.  $P(Y \leq 1, Z \leq 1 | X = 10)$       g.  $E(Z | X = 10)$

5.S24 Suponha que  $X$  tenha uma distribuição lognormal com parâmetros  $\theta$  e  $\omega$ . Determine a função densidade de probabilidade e os valores dos parâmetros para  $Y = X^\gamma$  para uma constante  $\gamma > 0$ . Qual é o nome dessa distribuição?

5.S25 A potência em um circuito DC é  $P = I/R$ , em que  $I$  e  $R$  denotam a corrente e a resistência, respectivamente. Suponha que  $I$  seja normalmente distribuída com média de 200 mA e desvio-padrão de 0,2 mA, sendo  $R$  uma constante. Determine a função densidade de probabilidade da potência.

5.S26 A intensidade ( $\text{mW}/\text{mm}^2$ ) de um feixe de *laser* em uma superfície segue teoricamente uma distribuição normal bivariada com intensidade máxima no centro, igual variância  $\sigma$  nas direções  $x$  e  $y$ , e covariância zero. Existem várias definições para a largura do feixe. Uma definição é o diâmetro no qual a intensidade é 50 % de seu pico. Suponha que a largura do feixe seja de 1,6 mm sob essa definição. Determine  $\sigma$ . Determine também a largura do feixe quando ele é definido como o diâmetro em que a intensidade é igual a  $1/e^2$  do pico.

# Capítulo 6 Exercícios

## Exercícios para a Seção 6.1

6.1.1 A média da amostra sempre corresponderá a uma das observações na amostra?

6.1.2 Exatamente metade das observações em uma amostra cairá abaixo da média?

6.1.3 A média da amostra sempre será o valor que ocorre com mais frequência na amostra?

6.1.4 Para qualquer conjunto de valores dados, é possível que o desvio-padrão da amostra seja maior do que a média da amostra? Dê um exemplo.

6.1.5 O desvio-padrão pode ser igual a zero? Se sim, dê um exemplo.

6.1.6 Em *Applied Life Data Analysis* (Wiley, 1982), Wayne Nelson apresenta o tempo de esgotamento de um fluido isolante entre eletrodos a 34 kV. Os tempos, em minutos, são: 0,19; 0,78; 0,96; 1,31; 2,78; 3,16; 4,15; 4,67; 4,85; 6,50; 7,35; 8,01; 8,27; 12,06; 31,75; 32,52; 33,91; 36,71 e 72,89. Calcule a média e o desvio-padrão da amostra.

6.1.7 Um artigo em *Human Factors* ["Visual Performance on CRT Screens and Hard-Copy Displays" (1989, Vol. 31(3), pp. 247-257)] apresentou dados sobre acomodação visual (uma função do movimento do olho), quando reconhecendo um padrão de mancha em um vídeo CRT de alta resolução. Os dados são: 36,45; 67,90; 38,77; 42,18; 26,72; 50,77; 39,30 e 49,71. Calcule a média e o desvio-padrão da amostra. Construa um diagrama de pontos dos dados.

6.1.8 Prevenir a propagação de trinca de fadiga em estruturas de aviões é um importante elemento de segurança em aeronaves. Um estudo de engenharia para investigar a trinca de fadiga em  $n = 9$  asas carregadas ciclicamente reportou os seguintes comprimentos (em mm) de trinca: 2,13; 2,96; 3,02; 1,82; 1,15; 1,37; 2,04; 2,47; 2,60. Calcule a média e o desvio-padrão da amostra. Prepare um diagrama de pontos dos dados.

6.1.9 Um artigo publicado no *Journal of Physiology* ["Response of Rat Muscle to Acute Resistance Exercise Defined by Transcriptional and Translational Profiling" (2002, Vol. 545, pp. 27-41)] estudou a expressão do gene como uma função do exercício de resistência. Dados de expressão (medidas de atividade dos genes) de um gene são mostrados na tabela seguinte. Um grupo de ratos foi exercitado durante seis horas, enquanto outro não recebeu exercícios. Calcule a média e o desvio-padrão da amostra do grupo que se exercitou e do que não se exercitou, separadamente. Construa um diagrama de pontos para o grupo que se exercitou e para o que não se exercitou, separadamente. Comente qualquer diferença entre os grupos.

Seis Horas de Exercício	Seis Horas de Exercício	Nenhum Exercício	Nenhum Exercício
425,313	208,475	485,396	406,921
223,306	286,484	159,471	335,209
388,793	244,242	478,314	
139,262	408,099	245,782	
212,565	157,743	236,212	
324,024	436,37	252,773	

6.1.10 O Exercício 6.1.7 descreve os dados de um artigo em *Human Factors* sobre a acomodação visual de um experimento envolvendo um monitor CRT de alta resolução.

Dados de um segundo experimento, usando um monitor de baixa resolução, foram também reportados no artigo. Eles são: 8,85; 35,80; 26,53; 64,63; 9,00; 15,38; 8,14 e 8,24. Prepare um diagrama de pontos para essa segunda amostra e compare-o com o da primeira amostra.

O que você pode concluir acerca da resolução CRT nessa situação?

6.1.11 Os seguintes dados são as temperaturas ( $^{\circ}\text{F}$ ) das juntas dos *O-rings* para cada teste de incêndio ou de lançamento real de espaçonaves (de *Presidential Commission on the Space Shuttle Challenger Accident*, Vol. 1, pp. 129-131): 84, 49, 61, 40, 83, 67, 45, 66, 70, 69, 80, 58, 68, 60, 67, 72, 73, 70, 57, 63, 70, 78, 52, 67, 53, 67, 75, 61, 70, 81, 76, 79, 75, 76, 58, 31.

a. Calcule a média e o desvio-padrão da amostra e construa o diagrama de pontos para os dados de temperatura.

b. Não considere a menor observação ( $31^{\circ}\text{F}$ ) e recalcule as grandezas do item (a). Comente o que achou. Quão "diferentes" as outras temperaturas são desse último valor?

6.1.12 Semeadura de nuvens, um processo em que químicos, tais como iodeto de prata e dióxido de carbono congelado, são introduzidos nas nuvens por avião para provocar chuva, foi amplamente usada no século XX. Uma recente pesquisa tem questionado sua eficiência ["Reassessment of Rain Enhancement Experiments and Operations in Israel Including Synoptic Considerations", *Journal of Atmospheric Research* (2010, Vol. 97(4), pp. 513-525)]. Um experimento foi feito estabelecendo aleatoriamente que 52 nuvens seriam semeadas ou não. A quantidade de chuva gerada foi então medida em acre-pé. A seguir, os dados de nuvens semeadas e não semeadas:

Não semeadas:

81,2 26,1 95,0 41,1 28,6 21,7 11,5 68,5 345,5 321,2  
1202,6 1,0 4,9 163,0 372,4 244,3 47,3 87,0 26,3 24,4  
830,1 4,9 36,6 147,8 17,3 29,0

Semeadas:

274,7 302,8 242,5 255,0 17,5 115,3 31,4 703,4 334,1  
1697,8 118,3 198,6 129,6 274,7 119,0 1656,0 7,7 430,0  
40,6 92,4 200,7 32,7 4,1 978,0 489,1 2745,6

Encontre a média, o desvio-padrão e a amplitude de chuva da amostra para

a. Todas as 52 nuvens.

b. As nuvens não semeadas.

c. As nuvens semeadas.

6.1.13 Construa diagramas de pontos das nuvens semeadas e não semeadas e compare suas distribuições em algumas sentenças.

6.1.14 Nos Jogos Olímpicos de Sydney em 2000, um programa especial iniciado pelo presidente do COI permitiu aos países em desenvolvimento enviar atletas para as Olimpíadas sem o usual procedimento de qualificação. Aqui estão os 71 tempos (em segundos) para a primeira rodada dos 100 metros de nado masculino.

60,39 49,93 53,40 51,82 50,46 51,34 50,28 50,19 52,14  
50,56 52,72 50,95 49,74 49,16 52,57 52,53 52,09 52,40  
49,75 54,06 53,50 50,63 51,93 51,62 52,58 53,55 51,07  
49,76 49,73 50,90 59,26 49,29 52,78 112,72 49,79 49,83  
52,43 51,28 52,22 49,76 49,70 52,90 50,19 54,33 62,45  
51,93 52,24 52,82 50,96 48,64 51,11 50,87 52,18 54,12  
50,49 49,84 52,91 52,52 50,32 51,52 52,0 52,85 52,24  
49,45 51,28 49,09 58,79 49,74 49,32 50,62 49,45

a. Encontre a média e o desvio-padrão da amostra para esses tempos de natação.

b. Construa um diagrama de pontos para esses dados.

c. Comente sobre qualquer coisa não usual que você tenha visto.

## Exercícios para a Seção 6.2

**6.2.1** Um diagrama de ramo e folhas comparativo (*back-to-back stem-and-leaf*) é feito colocando os dados em ambos os lados dos mesmos ramos. Eis um diagrama de ramo e folhas comparativo para os dados de semeadura de nuvens do Exercício 6.1.12, mostrando as nuvens não semeadas à esquerda e as nuvens semeadas à direita.

```

6509875443332221000 | 0 | 01233492223
                        | 2 | 00467703
                        | 4 | 39
                        | 6 | 10
3 | 8 | 8
  | 10 |
0 | 12 |
  | 14 |
  | 16 | 60
  | 18 |
  | 20 |
  | 22 |
  | 24 |
  | 26 | 5

```

Como o diagrama de ramo e folhas comparativo mostra as diferenças no conjunto de dados de uma forma que o diagrama de pontos não pode mostrar?

**6.2.2** Quando a mediana de uma amostra será igual à média da amostra?

**6.2.3** Quando a mediana de uma amostra será igual à moda?

**6.2.4** Um artigo em *Technometrics* ["Validation of Regression Models: Methods and Examples" (1977, Vol. 19(4), p. 425)] apresentou os seguintes dados sobre taxas de octanagem de combustível para motor, de várias misturas de gasolina:

88,5	98,8	89,6	92,2	92,7	88,4	87,5	90,9
94,7	88,3	90,4	83,4	87,9	92,6	87,8	89,9
84,3	90,4	91,6	91,0	93,0	93,7	88,3	91,8
90,1	91,2	90,7	88,2	94,4	96,5	89,2	89,7
89,0	90,6	88,6	88,5	90,4	84,3	92,3	92,2
89,8	92,2	88,3	93,3	91,2	93,2	88,9	
91,6	87,7	94,2	87,4	86,7	88,6	89,8	
90,3	91,1	85,3	91,1	94,2	88,7	92,7	
90,0	86,7	90,1	90,5	90,8	92,7	93,3	
91,5	93,4	89,3	100,3	90,1	89,3	86,7	
89,9	96,1	91,1	87,6	91,8	91,0	91,0	

Construa um diagrama de ramo e folhas para esses dados. Calcule a mediana e os quartis desses dados.

**6.2.5** Os seguintes dados são os números de ciclos, até falhar, de corpos de prova de alumínio, sujeitos a uma tensão alternada repetida, de 21.000 psi e 18 ciclos por segundo.

1115	865	1015	885	1594	1000	1416	1501
1310	2130	845	1223	2023	1820	1560	1238
1540	1421	1674	375	1315	1940	1055	990
1502	1109	1016	2265	1269	1120	1764	1468
1258	1481	1102	1910	1260	910	1330	1512
1315	1567	1605	1018	1888	1730	1608	1750
1085	1883	706	1452	1782	1102	1535	1642
798	1203	2215	1890	1522	1578	1781	
1020	1270	785	2100	1792	758	1750	

Construa um diagrama de ramo e folhas para esses dados. Calcule a mediana e os quartis desses dados. Você acha que o corpo de prova "sobreviverá" além de 2000 ciclos? Justifique sua resposta.

**6.2.6** Calcule a mediana, a moda e a média dos dados do Exercício 6.2.4. Explique como essas três medidas de localização descrevem características diferentes dos dados.

**6.2.7** Calcule a mediana, a moda e a média dos dados do Exercício 6.2.5. Explique como essas três medidas de localização descrevem características diferentes dos dados.

**6.2.8** Alunas de um curso de graduação em engenharia reportaram suas alturas aproximadas em polegadas. Os dados estão apresentados a seguir. Construa um diagrama de ramo e folhas para os dados de altura e comente importantes características que você observe. Calcule a média, o desvio-padrão e a mediana da altura.

62	64	61	67	65	68	61	65	60	65	64	63	59
68	64	66	68	69	65	67	62	66	68	67	66	65
69	65	69	65	67	67	65	63	64	67	65		

**6.2.9** As resistências ao cisalhamento de 100 soldas de ponto em uma liga de titânio são dadas a seguir. Construa um diagrama de ramo e folhas para os dados da resistência da solda e comente importantes características que você observe. Qual é o nonagésimo quinto percentil da resistência?

5408	5431	5475	5442	5376	5388	5459	5422	5416	5435
5420	5429	5401	5446	5487	5416	5382	5357	5388	5457
5407	5469	5416	5377	5454	5375	5409	5459	5445	5429
5463	5408	5481	5453	5422	5354	5421	5406	5444	5466
5399	5391	5477	5447	5329	5473	5423	5441	5412	5384
5445	5436	5454	5453	5428	5418	5465	5427	5421	5396
5381	5425	5388	5388	5378	5481	5387	5440	5482	5406
5401	5411	5399	5431	5440	5413	5406	5342	5452	5420
5458	5485	5431	5416	5431	5390	5399	5435	5387	5462
5383	5401	5407	5385	5440	5422	5448	5366	5430	5418

**6.2.10** Uma característica importante de qualidade da água é a concentração de material sólido suspenso. Em seguida, são apresentadas 60 medidas de sólidos suspensos de certo lago. Construa um diagrama de ramo e folhas para esses dados e comente importantes características que você observe. Calcule a média, o desvio-padrão e a mediana da amostra. Qual é o nonagésimo percentil da concentração?

42,4	65,7	29,8	58,7	52,1	55,8	57,0	68,7	67,3	67,3
54,3	54,0	73,1	81,3	59,9	56,9	62,2	69,9	66,9	59,0
56,3	43,3	57,4	45,3	80,1	49,7	42,8	42,4	59,6	65,8
61,4	64,0	64,2	72,6	72,5	46,1	53,1	56,1	67,2	70,7
42,6	77,4	54,7	57,1	77,3	39,3	76,4	59,3	51,1	73,8
61,4	73,1	77,3	48,5	89,8	50,7	52,0	59,6	66,1	31,6

**6.2.11** Um grupo de entusiastas do vinho testou um vinho *pinot noir* de Oregon. A avaliação foi graduar o vinho em uma escala de 0 a 100 pontos. Os resultados são dados a seguir. Construa um diagrama de ramo e folhas para esses dados e comente importantes características que você observe. Calcule a média, o desvio-padrão e a mediana da amostra. Um vinho avaliado acima de 90 é considerado verdadeiramente excepcional. Que proporção de produtores considerou esse particular *pinot noir* como verdadeiramente excepcional?

#### 44 Exercícios Complementares

94	90	92	91	91	86	89	91	91	90
90	93	87	90	91	92	89	86	89	90
88	95	91	88	89	92	87	89	95	92
85	91	85	89	88	84	85	90	90	83

6.2.12 Em seu livro *Introduction to Linear Regression Analysis* (5ª edição, Wiley, 2012), Montgomery, Peck e Vining apresentam medidas de concentração de  $\text{NbOCl}_3$  em um experimento em um reator tubular. Os dados, em  $\text{gmol/L} \times 10^{-3}$ , são fornecidos a seguir. Construa um diagrama de ramo e folhas para esses dados e comente importantes características que você observe. Calcule a média, o desvio-padrão e a mediana da amostra.

450	450	473	507	457	452	453	1215	1256
1145	1085	1066	1111	1364	1254	1396	1575	1617
1733	2753	3186	3227	3469	1911	2588	2635	2725

6.2.13 No Exercício 6.2.8, apresentamos os dados de altura que foram reportados por alunas de graduação em engenharia. Na mesma turma, os alunos reportaram suas próprias alturas, dadas a seguir. Construa um diagrama comparativo de ramo e folhas, listando os ramos no centro do gráfico e então colocando as folhas femininas no lado esquerdo e as folhas masculinas no lado direito. Comente importantes características observadas por você.

69	67	69	70	65	68	69	70	71	69	66	67	69	75	68	67	68
69	70	71	72	68	69	69	70	71	68	72	69	69	68	69	73	70
73	68	69	71	67	68	65	68	68	69	70	74	71	69	70	69	

### Exercícios para a Seção 6.3

6.3.1 Construa uma distribuição de frequências e um histograma usando os dados de fratura do Exercício 6.2.5.

6.3.2 Construa uma distribuição de frequências e um histograma para os dados de octanagem do combustível para motor do Exercício 6.2.4. Use oito intervalos de classe.

6.3.3 Construa histogramas com 8 e 16 intervalos de classe, para os dados do Exercício 6.2.5. Compare os histogramas. Os dois histogramas apresentam informação similar?

6.3.4 Construa uma distribuição de frequências e histogramas, com 8 e 16 intervalos de classe, para os dados de octanagem do combustível para motor do Exercício 6.2.4. Compare os histogramas. Os dois histogramas apresentam informação similar?

6.3.5 Construa um histograma para os dados de resistência da solda de ponto do Exercício 6.2.9. Comente a forma do histograma. Ele contém a mesma informação que o diagrama de ramo e folhas?

6.3.6 Construa um histograma para os dados de altura das alunas do Exercício 6.2.8.

6.3.7 Construa um histograma para os dados da avaliação do vinho *pinot noir* do Exercício 6.2.11. Comente a forma do histograma. Ele contém a mesma informação que o diagrama de ramo e folhas?

6.3.8 Construa um histograma para os dados de qualidade da água do Exercício 6.2.10. Comente a forma do histograma. Ele contém a mesma informação que o diagrama de ramo e folhas?

6.3.9 **Diagrama de Pareto.** Uma variação importante de um histograma para dados categóricos é o diagrama de Pareto. Esse gráfico é largamente empregado nos esforços de melhoria de qualidade, em que as categorias geralmente representam tipos diferentes de defeitos, modos de falha ou problemas nos produtos/processos. As categorias são ordenadas de modo que a categoria com a maior frequência ficará à esquerda, seguida pela categoria com a segunda maior frequência, e assim por diante. Esses diagramas têm o nome do economista italiano Vilfredo Pareto, e geralmente exibem a “Lei de Pareto”; ou seja, a maioria dos defeitos pode ser creditada apenas a poucas categorias. Suponha que a seguinte informação sobre defeitos estruturais nas portas dos automóveis seja obtida: quatro arranhões, quatro buracos, seis peças arrumadas fora de sequência, 21 peças subapareladas, oito fendas perdidas, cinco peças não lubrificadas, 30 peças fora de contorno e três peças com rebarbas. Construa um diagrama de Pareto e interprete-o.

### Exercícios para a Seção 6.4

6.4.1 Usando os dados do Exercício 6.1.12 sobre a semeadura de nuvens,

- Encontre a mediana e os quartis dos dados de nuvens não semeadas.

- Encontre a mediana e os quartis dos dados de nuvens semeadas.

- Faça dois diagramas de caixa, lado a lado, um para cada grupo, no mesmo gráfico.

- Compare as distribuições a partir do que você pode ver nos diagramas de caixa lado a lado.

6.4.2 Usando os dados do Exercício 6.1.14 sobre os tempos de nado,

- Encontre a mediana e os quartis para os dados.

- Faça um diagrama de caixa para os dados.

- Repita os itens (a) e (b) para os dados sem o *outlier* extremo e comente.

- Compare a distribuição dos dados com e sem o *outlier* extremo.

6.4.3 O “tempo de ignição fria” de um motor de carro está sendo investigado por um fabricante de gasolina. Os seguintes tempos (em segundos) foram obtidos em um veículo de teste: 1,75; 1,92; 2,62; 2,35; 3,09; 3,15; 2,53 e 1,91.

- Calcule a média, a variância e o desvio-padrão da amostra.

- Construa um diagrama de caixa dos dados.

6.4.4 Um artigo em *Transactions of the Institution of Chemical Engineers* (1956, Vol. 34, pp. 280-293) reportou dados sobre um

experimento investigando o efeito de muitas variáveis de processos na oxidação, em fase vapor, de naftaleno. Uma amostra da conversão percentual molar de naftaleno em anidrido maleico resulta em: 4,2; 4,7; 4,7; 5,0; 3,8; 3,6; 3,0; 5,1; 3,1; 3,8; 4,8; 4,0; 5,2; 4,3; 2,8; 2,0; 2,8; 3,3; 4,8 e 5,0.

- Calcule a média, a variância e o desvio-padrão da amostra.
- Construa um diagrama de caixa dos dados.

**6.4.5** O Exercício 6.1.11 apresentou as temperaturas (°F) na junção de *O-rings* para cada lançamento teste ou real do foguete espacial. Naquele exercício, desejava-se encontrar a média e o desvio-padrão da amostra em relação à temperatura.

- Encontre a mediana e os quartis superior e inferior da temperatura.
- Isole a menor observação (31 °F) e recalcule as grandezas do item (a). Comente o que encontrou. Quão “diferentes” são as outras temperaturas desse menor valor?
- Construa, a partir dos dados, um diagrama de caixa e comente a possível presença de *outliers*.

**6.4.6** Reconsidere os dados de qualidade da água do Exercício 6.2.10. Construa um diagrama de caixa dos dados e escreva uma interpretação do gráfico. Como o diagrama de caixa se compara, em valor interpretativo, ao diagrama original de ramo e folhas?

**6.4.7** Reconsidere os dados de resistência da solda do Exercício 6.2.9. Construa um diagrama de caixa dos dados e escreva uma interpretação do gráfico. Como o diagrama de caixa se compara, em valor interpretativo, ao diagrama original de ramo e folhas?

**6.4.8** No Exercício 6.4.3, dados foram apresentados sobre o tempo de ignição fria de uma gasolina particular usada em um veículo de teste. Uma segunda formulação da gasolina foi testada no mesmo veículo, com os seguintes tempos (em segundos): 1,83; 1,99; 3,13; 3,29; 2,65; 2,87; 3,40; 2,46; 1,89 e 3,35. Use esses novos dados, juntamente com os tempos de ignição fria reportados no Exercício 6.4.3, para construir diagramas de caixa comparativos. Escreva uma interpretação da informação que você vê nesses diagramas.

**6.4.9** Um artigo em *Nature Genetics* [“Treatment-specific Changes in Gene Expression Discriminate in Vivo Drug Response in Human

Leukemia Cells” (2003, Vol. 34(1), pp. 85-90)] estudou a expressão do gene como uma função de tratamentos para leucemia. Um grupo recebeu uma alta dose da droga, enquanto o grupo de controle não recebeu tratamento. Dados da expressão (medidas de atividade do gene) de um gene são mostrados na tabela seguinte. Construa um diagrama de caixa para cada grupo de pacientes. Escreva uma interpretação para comparar a informação nesses diagramas.

Alta Dose	Controle	Controle	Controle
16,1	297,1	25,1	131,1
134,9	491,8	820,1	166,5
52,7	1332,9	82,5	2258,4
14,4	1172	713,9	497,5
124,3	1482,7	785,6	263,4
99	335,4	114	252,3
24,3	528,9	31,9	351,4
16,3	24,1	86,3	678,9
15,2	545,2	646,6	3010,2
47,7	92,9	169,9	67,1
12,9	337,1	20,2	318,2
72,7	102,3	280,2	2476,4
126,7	255,1	194,2	181,4
46,4	100,5	408,4	2081,5
60,3	159,9	155,5	424,3
23,5	168	864,6	188,1
43,6	95,2	355,4	563
79,4	132,5	634	149,1
38	442,6	2029,9	2122,9
58,2	15,8	362,1	1295,9
26,5	175,6		

## Exercícios para a Seção 6.5

**6.5.1** Em seu livro *Time Series Analysis, Forecasting, and Control* (Prentice Hall, 1994), G. E. P. Box, G. M. Jenkins e G. C. Reinsel apresentam leituras de concentração de um processo químico, feitas a cada duas horas. Alguns desses dados são mostrados a seguir (leia para baixo e então da esquerda para a direita).

17,0	16,7	17,1	17,5	17,6
16,6	17,4	17,4	18,1	17,5
16,3	17,2	17,4	17,5	16,5
16,1	17,4	17,5	17,4	17,8
17,1	17,4	17,4	17,4	17,3
16,9	17,0	17,6	17,1	17,3
16,8	17,3	17,4	17,6	17,1

17,4	17,2	17,3	17,7	17,4
17,1	17,4	17,0	17,4	16,9
17,0	16,8	17,8	17,8	17,3

Construa e interprete um gráfico digiponto ou um diagrama de ramo e folhas desses dados.

**6.5.2** A tabela a seguir mostra as importações americanas de petróleo, como percentagens do total, e as importações do Golfo Pérsico, como uma percentagem de todas as importações, por ano, desde 1973 (fonte: página da internet do Departamento Americano de Energia, [www.eia.doe.gov/](http://www.eia.doe.gov/)). Construa e interprete um gráfico digiponto ou diagramas separados de ramo e folhas e de séries temporais dos dados.

<b>Ano</b>	<b>Importação de Petróleo (milhares de barris por dia)</b>	<b>Importação Total de Petróleo, dada em Percentagem de Produtos Fornecidos de Petróleo</b>	<b>Importação de Petróleo do Golfo Pérsico, dada em Percentagem de Importação Total de Petróleo</b>
1973	6256	36,1	13,5
1974	6112	36,7	17,0
1975	6055	37,1	19,2
1976	7313	41,8	25,1
1977	8807	47,7	27,8
1978	8363	44,3	26,5
1979	8456	45,6	24,4
1980	6909	40,5	21,9
1981	5996	37,3	20,3
1982	5113	33,4	13,6
1983	5051	33,1	8,7
1984	5437	34,5	9,3
1985	5067	32,2	6,1
1986	6224	38,2	14,6
1987	6678	40,0	16,1
1988	7402	42,8	20,8
1989	8061	46,5	23,0
1990	8018	47,1	24,5
1991	7627	45,6	24,1
1992	7888	46,3	22,5
1993	8620	50,0	20,6
1994	8996	50,7	19,2
1995	8835	49,8	17,8
1996	9478	51,7	16,9
1997	10,162	54,5	17,2
1998	10,708	56,6	19,9
1999	10,852	55,5	22,7
2000	11,459	58,1	21,7
2001	11,871	60,4	23,2
2002	11,530	58,3	19,6
2003	12,264	61,2	20,3
2004	13,145	63,4	18,9
2005	13,714	65,9	17,0
2006	13,707	66,3	16,1
2007	13,468	65,1	16,1
2008	12,915	66,2	18,4
2009	11,691	62,3	14,4
2010	11,793	61,5	14,5
2011	11,436	60,6	16,3
2012	10,627	57,3	20,3
2013	9859	52,0	20,4
2014	9241	48,4	20,3
2015	9449	48,4	15,9

6.5.3 Os 100 números anuais Wolfer da mancha solar, de 1770 a 1869, são mostrados a seguir. (Para uma análise interessante e interpretação desses números, veja o livro de Box, Jenkins e Reinsel, referenciado no Exercício 6.5.1. A análise deles requer algum conhecimento avançado de estatística e construção de modelo estatístico.) Leia para baixo e então da esquerda para a direita. O resultado de 1869 é 74. Construa e interprete um gráfico digiponto ou diagramas de ramo e folhas e de séries temporais desses dados.

101	31	154	38	83	90
82	7	125	23	132	67
66	20	85	10	131	60
35	92	68	24	118	47
41	10	16	8	62	94
21	8	7	13	98	96
16	2	4	57	124	77
6	0	2	122	96	59
4	1	8	138	66	44
7	5	17	103	64	47
14	12	36	86	54	30

34	14	50	63	39	16
45	35	62	37	21	7
43	46	67	24	7	37
48	41	71	11	4	74
42	30	48	15	23	
28	24	28	40	55	

6.5.4 A tabela a seguir contém as anomalias na temperatura mundial média do ar na superfície e a concentração mundial de CO<sub>2</sub> para os anos de 1880-2004 (fonte: <http://data.giss.nasa.gov/gistemp/>). A temperatura é medida em um número de localizações ao redor do mundo, sendo a média obtida anualmente e subtraída de uma média de um período de referência (1951-1980); o resultado é reportado como uma anomalia.

- Construa um gráfico de série temporal a partir dos dados das anomalias na temperatura mundial média do ar na superfície e comente quaisquer características que você observe.
- Construa um gráfico de série temporal a partir dos dados de concentração de CO<sub>2</sub> e comente quaisquer características que você observe.
- Sobreponha os dois gráficos no mesmo conjunto de eixos e comente o gráfico resultante.

Ano	Anomalia, °C	CO <sub>2</sub> , ppmv	Ano	Anomalia, °C	CO <sub>2</sub> , ppmv	Ano	Anomalia, °C	CO <sub>2</sub> , ppmv
1880	-0,11	290,7	1922	-0,09	303,8	1964	-0,25	319,2
1881	-0,13	291,2	1923	-0,16	304,1	1965	-0,15	320,0
1882	-0,01	291,7	1924	-0,11	304,5	1966	-0,07	321,1
1883	-0,04	292,1	1925	-0,15	305,0	1967	-0,02	322,0
1884	-0,42	292,6	1926	0,04	305,4	1968	-0,09	322,9
1885	-0,23	293,0	1927	-0,05	305,8	1969	0,00	324,2
1886	-0,25	293,3	1928	0,01	306,3	1970	0,04	325,2
1887	-0,45	293,6	1929	-0,22	306,8	1971	-0,10	326,1
1888	-0,23	293,8	1930	-0,03	307,2	1972	-0,05	327,2
1889	0,04	294,0	1931	0,03	307,7	1973	0,18	328,8
1890	-0,22	294,2	1932	0,04	308,2	1974	-0,06	329,7
1891	-0,55	294,3	1933	-0,11	308,6	1975	-0,02	330,7
1892	-0,40	294,5	1934	0,05	309,0	1976	-0,21	331,8
1893	-0,39	294,6	1935	-0,08	309,4	1977	0,16	333,3
1894	-0,32	294,7	1936	0,01	309,8	1978	0,07	334,6
1895	-0,32	294,8	1937	0,12	310,0	1979	0,13	336,9
1896	-0,27	294,9	1938	0,15	310,2	1980	0,27	338,7
1897	-0,15	295,0	1939	-0,02	310,3	1981	0,40	339,9
1898	-0,21	295,2	1940	0,14	310,4	1982	0,10	341,1
1899	-0,25	295,5	1941	0,11	310,4	1983	0,34	342,8
1900	-0,05	295,8	1942	0,10	310,3	1984	0,16	344,4
1901	-0,05	296,1	1943	0,06	310,2	1985	0,13	345,9
1902	-0,30	296,5	1944	0,10	310,1	1986	0,19	347,2
1903	-0,35	296,8	1945	-0,01	310,1	1987	0,35	348,9
1904	-0,42	297,2	1946	0,01	310,1	1988	0,42	351,5
1905	-0,25	297,6	1947	0,12	310,2	1989	0,28	352,9
1906	-0,15	298,1	1948	-0,03	310,3	1990	0,49	354,2
1907	-0,41	298,5	1949	-0,09	310,5	1991	0,44	355,6
1908	-0,30	298,9	1950	-0,17	310,7	1992	0,16	356,4
1909	-0,31	299,3	1951	-0,02	311,1	1993	0,18	357,0
1910	-0,21	299,7	1952	0,03	311,5	1994	0,31	358,9

(continua)

(continuação)

Ano	Anomalia, °C	CO <sub>2</sub> , ppmv	Ano	Anomalia, °C	CO <sub>2</sub> , ppmv	Ano	Anomalia, °C	CO <sub>2</sub> , ppmv
1911	-0,25	300,1	1953	0,12	311,9	1995	0,47	360,9
1912	-0,33	300,4	1954	-0,09	312,4	1996	0,36	362,6
1913	-0,28	300,8	1955	-0,09	313,0	1997	0,40	363,8
1914	-0,02	301,1	1956	-0,18	313,6	1998	0,71	366,6
1915	0,06	301,4	1957	0,08	314,2	1999	0,43	368,3
1916	-0,20	301,7	1958	0,10	314,9	2000	0,41	369,5
1917	-0,46	302,1	1959	0,05	315,8	2001	0,56	371,0
1918	-0,33	302,4	1960	-0,02	316,6	2002	0,70	373,1
1919	-0,09	302,7	1961	0,10	317,3	2003	0,66	375,6
1920	-0,15	303,0	1962	0,05	318,1	2004	0,60	377,4
1921	-0,04	303,4	1963	0,03	318,7			

## Exercícios para a Seção 6.6

6.6.1 Um artigo publicado no *Journal of Pharmaceuticals Sciences* ["Statistical Analysis of the Extended Hansen Method Using the Bootstrap Technique" (1991, Vol. 80, pp. 971-977)] apresentou dados sobre a solubilidade observada, expressa em fração molar, de um soluto a uma temperatura constante. Os parâmetros de solubilidade de Hansen são divididos em três contribuições: dispersão, ligação de hidrogênio e momento dipolo. Os dados são mostrados na tabela a seguir, em que  $y$  é o logaritmo da solubilidade, expressa em fração molar,  $x_1$  é a solubilidade parcial para a contribuição da dispersão,  $x_2$  é a solubilidade parcial para a contribuição do momento dipolo e  $x_3$  é a solubilidade parcial para a contribuição da ponte de hidrogênio.

- Construa uma matriz de gráficos de dispersão para essas variáveis.
- Comente sobre as relações aparentes entre  $y$  e as outras três variáveis.

6.6.2 Considere as importações de petróleo no Exercício 6.5.2.

- Construa uma matriz de gráficos de dispersão para as três variáveis e calcule os coeficientes de correlação.
- Comente sobre as relações aparentes entre as variáveis.

Número da Observação	$y$	$x_1$	$x_2$	$x_3$
1	0,22200	7,3	0,0	0,0
2	0,39500	8,7	0,0	0,3
3	0,42200	8,8	0,7	1,0
4	0,43700	8,1	4,0	0,2

Número da Observação	$y$	$x_1$	$x_2$	$x_3$
5	0,42800	9,0	0,5	1,0
6	0,46700	8,7	1,5	2,8
7	0,44400	9,3	2,1	1,0
8	0,37800	7,6	5,1	3,4
9	0,49400	10,0	0,0	0,3
10	0,45600	8,4	3,7	4,1
11	0,45200	9,3	3,6	2,0
12	0,11200	7,7	2,8	7,1
13	0,43200	9,8	4,2	2,0
14	0,10100	7,3	2,5	6,8
15	0,23200	8,5	2,0	6,6
16	0,30600	9,5	2,5	5,0
17	0,09230	7,4	2,8	7,8
18	0,11600	7,8	2,8	7,7
19	0,07640	7,7	3,0	8,0
20	0,43900	10,3	1,7	4,2
21	0,09440	7,8	3,3	8,5
22	0,11700	7,1	3,9	6,6
23	0,07260	7,7	4,3	9,5
24	0,04120	7,4	6,0	10,9
25	0,25100	7,3	2,0	5,2
26	0,00002	7,6	7,8	20,7

## Exercícios para a Seção 6.7

6.7.1 Construa um gráfico de probabilidade normal dos dados de acomodação visual do Exercício 6.1.7. Parece razoável supor que a acomodação visual seja normalmente distribuída?

6.7.2 Construa um gráfico de probabilidade normal dos dados de taxa de octanagem do Exercício 6.2.4. Parece razoável supor que a taxa de octanagem seja normalmente distribuída?

6.7.3 Construa um gráfico de probabilidade normal dos dados de ciclos de falha do Exercício 6.2.5. Parece razoável supor que os ciclos de falha sejam normalmente distribuídos?

6.7.4 Construa um gráfico de probabilidade normal dos dados de concentração de sólidos suspensos do Exercício 6.2.10. Parece razoável supor que a concentração de sólidos suspensos na água proveniente de determinado lago seja normalmente distribuída?

**6.7.5** Construa um gráfico de probabilidade normal dos dados de altura nos Exercícios 6.2.8 e 6.2.13. Faça um gráfico para os dados de alunas e de alunos nos mesmos eixos. A altura parece ser normalmente distribuída para ambos os grupos de estudantes? Se as duas populações têm a mesma variância, os dois gráficos de probabilidade normal devem ter as mesmas inclinações. Que conclusões você pode tirar acerca das alturas dos dois grupos de estudantes, a partir do exame visual dos gráficos de probabilidade normal?

**6.7.6** É possível obter uma estimativa “rápida e grosseira” da média de uma distribuição normal a partir do valor do quinquagésimo percentil em um gráfico de probabilidade normal. Argumente a razão disso. É possível também obter uma estimativa do desvio-padrão de uma distribuição normal, subtraindo o valor do octogésimo quarto percentil do valor do quinquagésimo percentil. Argumente a razão disso.

## Exercícios Suplementares para o Capítulo 6

**6.S7** A tabela a seguir mostra os dados de desempregos para os Estados Unidos, que são ajustados sazonalmente. Construa um gráfico de séries temporais desses dados e comente qualquer característica (fonte: página na internet do Escritório Americano do Trabalho, <http://data.bls.gov>).

**6.S8** Uma amostra com seis resistores resultou nas seguintes resistências (ohms):  $x_1 = 45$ ,  $x_2 = 38$ ,  $x_3 = 47$ ,  $x_4 = 41$ ,  $x_5 = 35$  e  $x_6 = 43$ .

- Calcule a variância e o desvio-padrão da amostra.
- Subtraia 35 de cada uma das medidas originais de resistências e calcule  $s^2$  e  $s$ . Compare seus resultados com aqueles obtidos no item (a) e explique sua resposta.
- Se as resistências fossem 450; 380; 470; 410; 350 e 430 ohms, você poderia usar os resultados dos itens anteriores deste problema para encontrar  $s^2$  e  $s$ ?

**6.S9** Considere as duas amostras dadas abaixo:

Amostra 1: 10; 9; 8; 7; 8; 6; 10; 6.

Amostra 2: 10; 6; 10; 6; 8; 10; 8; 6.

- Calcule a amplitude para as duas amostras. Você concluiria que as duas amostras exibem a mesma variabilidade? Explique.
- Calcule o desvio-padrão para as duas amostras. Essas grandezas indicam que as duas amostras têm a mesma variabilidade? Explique.
- Escreva um curto texto contrastando a amplitude da amostra com o desvio-padrão da amostra, como uma medida de variabilidade.

**6.S10** Os seguintes dados são as temperaturas de efluente, em dias consecutivos, na descarga de uma estação de tratamento de esgoto:

43	47	51	48	52	50	46	49
45	52	46	51	44	49	46	51
49	45	44	50	48	50	49	50

- Calcule a média, a mediana, a variância e o desvio-padrão da amostra.
- Construa um diagrama de caixa dos dados e faça comentários a respeito da informação obtida nesse diagrama.

**6.S11** O consumo americano líquido total de eletricidade, por ano, de 1980 a 2007 (em bilhões de quilowatts-hora), é dado na tabela a seguir (fonte: *site* do Departamento Americano de Energia, [www.eia.doe.gov/emeu/international/contents.html#InternationalElectricity](http://www.eia.doe.gov/emeu/international/contents.html#InternationalElectricity)). O consumo líquido exclui a energia consumida pelas unidades geradoras.

1980	2094,4	1981	2147,1	1982	2086,4	1983	2151,0
1984	2285,8	1985	2324,0	1986	2368,8	1987	2457,3
1988	2578,1	1989	2755,6	1990	2837,1	1991	2886,1
1992	2897,2	1993	3000,7	1994	3080,9	1995	3164,0
1996	3253,8	1997	3301,8	1998	3425,1	1999	3483,7
2000	3592,4	2001	3557,1	2002	3631,7	2003	3662,0
2004	3715,9	2005	3811,0	2006	3816,8	2007	3891,7

Construa um gráfico de séries temporais desses dados. Construa e interprete um diagrama de ramo e folhas para esses dados.

Ano	Jan.	Fev.	Mar.	Abr.	Maió	Jun.	Jul.	Ago.	Set.	Out.	Nov.	Dez.
1999	4,3	4,4	4,2	4,3	4,2	4,3	4,3	4,2	4,2	4,1	4,1	4,0
2000	4,0	4,1	4,0	3,8	4,0	4,0	4,0	4,1	3,9	3,9	3,9	3,9
2001	4,2	4,2	4,3	4,4	4,3	4,5	4,6	4,9	5,0	5,3	5,5	5,7
2002	5,7	5,7	5,7	5,9	5,8	5,8	5,8	5,7	5,7	5,7	5,9	6,0
2003	5,8	5,9	5,9	6,0	6,1	6,3	6,2	6,1	6,1	6,0	5,8	5,7
2004	5,7	5,6	5,8	5,6	5,6	5,6	5,5	5,4	5,4	5,5	5,4	5,4
2005	5,2	5,4	5,2	5,2	5,1	5,1	5,0	4,9	5,0	5,0	5,0	4,8
2006	4,7	4,8	4,7	4,7	4,7	4,6	4,7	4,7	4,5	4,4	4,5	4,4
2007	4,6	4,5	4,4	4,5	4,5	4,6	4,7	4,7	4,7	4,8	4,7	4,9
2008	4,9	4,8	5,1	5,0	5,5	5,6	5,8	6,2	6,2	6,6	6,8	7,2
2009	7,6	8,1	8,5	8,9	9,4	9,5	9,4	9,7	9,8			

6.S12 Transformações. Em alguns conjuntos de dados, uma transformação por alguma função matemática aplicada aos dados originais, tais como  $\sqrt{y}$  ou  $\log(y)$ , pode resultar em dados que sejam mais fáceis de trabalhar, estatisticamente, do que os dados originais. Para ilustrar o efeito de uma transformação, considere os seguintes dados, que representam ciclos de falhas de um tecido: 675; 3650; 175; 1150; 290; 2000; 100; 375.

- a. Construa um diagrama de caixa e comente a forma da distribuição dos dados.
- b. Transforme os dados usando logaritmos; isto é, seja  $y^*$  (novo valor) =  $\log y$  (valor antigo). Construa um gráfico de probabilidade normal dos dados transformados e faça comentários a respeito do efeito da transformação.

6.S13 Em 1879, A. A. Michelson fez 100 determinações da velocidade da luz no ar, usando uma modificação de um método proposto pelo físico francês Léon Foucault. Michelson fez as medidas em cinco tentativas de 20 medidas cada. As observações (em quilômetros por segundo) são apresentadas na tabela a seguir. Cada valor tem 299.000 subtraído dele.

Tentativa 1				
850	900	930	950	980
1000	930	760	1000	960
740	1070	850	980	880
980	650	810	1000	960

Tentativa 2				
960	960	880	850	900
830	810	880	800	760
940	940	800	880	840
790	880	830	790	800

Tentativa 3				
880	880	720	620	970
880	850	840	850	840
880	860	720	860	950
910	870	840	840	840

Tentativa 4				
890	810	800	760	750
910	890	880	840	850
810	820	770	740	760
920	860	720	850	780

Tentativa 5				
890	780	760	790	820
870	810	810	950	810
840	810	810	810	850
870	740	940	800	870

7,87	9,43	7,16	8,67	12,31
11,12	13,43	9,07	6,94	10,28
8,58	8,00	5,98	15,24	8,54
12,91	10,35	9,60	9,58	9,83
12,28	7,23	2,97	8,81	9,27
14,24	11,43	10,28	13,60	5,94
11,62	11,21	10,95	7,62	10,40
7,73	8,37	7,29	10,49	8,69
7,15	12,69	13,38	13,11	10,50

A velocidade verdadeira, correntemente aceita, da luz no vácuo é de 299.792,5 quilômetros por segundo. Stigler [“Do Robust Estimators Work with Real Data?” *Annals of Statistics* (1977, Vol. 5(6), pp. 1055-1098)] reporta que o valor “verdadeiro” para comparação com essas medidas é 734,5. Construa diagramas de caixa comparativos dessas medidas. Parece que todas as cinco tentativas são consistentes com relação à variabilidade das medidas? Todas as cinco tentativas estão centradas no mesmo valor? Como cada grupo de tentativas se compara ao valor verdadeiro? Poderia ter havido efeitos de “partida” no experimento que Michelson realizou? Poderia ter havido tendência no instrumento de medida?

6.S14 Em 1789, Henry Cavendish estimou a massa específica da Terra, usando uma balança de torção. Suas 29 medidas estão apresentadas a seguir, expressas como um múltiplo da massa específica da água.

5,50	5,30	5,47	5,10	5,29	5,65
5,55	5,61	5,75	5,63	5,27	5,44
5,57	5,36	4,88	5,86	5,34	5,39
5,34	5,53	5,29	4,07	5,85	5,46
5,42	5,79	5,62	5,58	5,26	

- a. Calcule a média, o desvio-padrão e a mediana da amostra dos dados de Cavendish para a densidade.
- b. Construa um gráfico de probabilidade normal dos dados. Comente o gráfico. Parece haver um *outlier* “baixo” nos dados?
- c. A mediana da amostra seria a melhor estimativa da densidade da Terra do que a média da amostra? Por quê?

6.S15 Em seu livro *Introduction to Time Series Analysis and Forecasting* (Wiley, 2008), Montgomery, Jennings e Kulahci apresentaram os dados de taxa de afogamento de crianças entre um e quatro anos por 100.000 pessoas no Arizona, de 1970 a 2004. Os dados são: 19,9; 16,1; 19,5; 19,8; 21,3; 15,0; 15,5; 16,4; 18,2; 15,3; 15,6; 19,5; 14,0; 13,1; 10,5; 11,5; 12,9; 8,4; 9,2; 11,9; 5,8; 8,5; 7,1; 7,9; 8,0; 9,9; 8,5; 9,1; 9,7; 6,2; 7,2; 8,7; 5,8; 5,7 e 5,2.

- a. Faça uma análise gráfica apropriada dos dados.
- b. Calcule e interprete os resumos numéricos apropriados.
- c. Note que a taxa parece começar a diminuir dramaticamente em torno de 1990. Discuta algumas razões potenciais para isso ter acontecido.
- d. Se tiver havido uma mudança real na taxa de afogamento começando em 1990, que impacto isso terá sobre as estatísticas calculadas no item (b)?

6.S16 O consumo de energia para 90 casas aquecidas a gás durante uma temporada de aquecimento no inverno é dado na tabela a seguir. A variável reportada é Btu/número de dias de aquecimento.

- a. Calcule a média da amostra e o desvio-padrão do uso de energia.
- b. Construa um histograma dos dados do uso de energia e comente sobre a forma da distribuição dos dados.
- c. Construa um diagrama de ramo e folhas das leituras do uso de energia.
- d. Que proporção dos dados do uso de energia está acima do uso médio mais dois desvios-padrão?

9,84	16,90	10,04	12,62	7,62
9,37	7,93	13,96	6,80	4,00
11,09	11,70	12,71	6,78	9,82
9,52	18,26	10,64	6,62	5,20
11,29	8,29	9,96	10,30	16,06
10,36	6,85	6,72	10,21	8,61
12,92	15,12	13,47	8,47	11,70
8,26	7,69	12,19	5,56	9,76
14,35	13,42	6,35	9,83	12,16

**6.S17** Um dos autores (DCM) tem um Mercedes-Benz 500 SL Roadster. É um modelo 2003, que tem um consumo razoavelmente baixo (atualmente, 60.112 milhas no odômetro). Ele está interessado em aprender como o consumo de seu carro se compara ao consumo de carros similares SL. A tabela a seguir contém o consumo de 100 carros Mercedes-Benz SL, modelos 2003-2009, consultados no *site* Cars.com.

- Calcule a média da amostra e o desvio-padrão das leituras do odômetro.
- Construa um histograma das leituras do odômetro e comente sobre a forma da distribuição dos dados.

2020	8905	1698	17.971	6207	22.643	4977	17.656	8940	11.508
7893	10.327	37.687	15.000	4166	9056	19.842	15.598	33.745	22.168
19.000	31.668	33.512	28.522	5824	18.327	31.845	30.015	2171	36.161
15.984	16.903	37.789	28.958	40.944	18.498	40.057	15.272	28.968	30.487
32.271	36.889	21.564	31.000	42.915	19.377	19.634	26.313	43.049	30.396
38.277	72.272	3800	21.218	29.250	48.648	29.216	44.944	49.125	33.065
32.524	38.139	62.940	51.326	54.126	4100	45.540	26.235	46.505	34.420
15.972	41.218	43.382	15.879	13.500	77.809	25.708	29.000	58.006	51.071
63.249	58.526	66.325	49.489	32.800	67.000	60.499	63.260	60.449	27.422
60.583	83.500	56.314	67.072	62.500	47.603	51.936	65.195	64.473	85.475

- Construa um diagrama de ramo e folhas das leituras do odômetro.
- Qual é o percentil do consumo do carro de DCM?

**6.S18** Considere a anomalia da temperatura média global do ar na superfície e os dados de concentração global de CO<sub>2</sub> mostrados originalmente no Exercício 6.5.4.

- Construa um gráfico de dispersão da anomalia da temperatura média global do ar na superfície *versus* a concentração global de CO<sub>2</sub>. Comente sobre o gráfico.
- Qual é o coeficiente simples de correlação entre essas duas variáveis?

## Capítulo 7 Exercícios

### Exercícios para a Seção 7.2

**7.2.1** Considere os dados de resistência à compressão na Tabela 6.2. Que proporção dos corpos de prova exibe uma resistência compressiva de no mínimo 200 psi? Qual é o parâmetro da população estimado aqui?

**7.2.2** Considere a fibra sintética do exercício prévio. Como varia o desvio-padrão da média da amostra, quando o tamanho da amostra aumenta de  $n = 6$  para  $n = 49$ ?

**7.2.3** Uma fibra sintética, usada na fabricação de carpete, tem uma resistência à tração que é normalmente distribuída, com média de 75,5 psi e desvio-padrão de 3,5 psi. Encontre a probabilidade de uma amostra aleatória de  $n = 6$  corpos de prova de fibra ter uma resistência média amostral à tração que exceda 75,75 psi.

**7.2.4** Suponha que  $X$  seja uma variável aleatória e tenha uma distribuição uniforme contínua

$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{caso contrário} \end{cases}$$

Suponha que uma amostra aleatória de  $n = 12$  observações seja selecionada dessa população. Qual é a distribuição de probabilidades de  $\bar{X} - 6$ ? Encontre a média e a variância dessa grandeza.

**7.2.5** Uma população normal tem média igual a 100 e variância igual a 25. Quão grande tem de ser a amostra aleatória, se quisermos que o erro-padrão da média amostral seja igual a 1,5?

**7.2.6** A quantidade de tempo que um consumidor gasta esperando no balcão de *check-in* de um aeroporto é uma variável aleatória, com média de 8,2 minutos e desvio-padrão de 1,5 minuto. Suponha que uma amostra aleatória de  $n = 49$  consumidores seja observada. Encontre a probabilidade de que o tempo médio de espera na fila para esses consumidores seja

- Menor do que dez minutos
- Entre cinco e dez minutos
- Menor do que seis minutos

**7.2.7** Uma amostra aleatória de tamanho  $n_1 = 16$  é selecionada de uma população normal, com média de 75 e desvio-padrão de 8. Uma segunda amostra aleatória de tamanho  $n_2 = 9$  é retirada de outra população normal, com média 70 e desvio-padrão 12. Sejam  $\bar{X}_1$  e  $\bar{X}_2$  duas médias amostrais. Encontre:

- A probabilidade de  $\bar{X}_1 - \bar{X}_2$  exceder 4.
- A probabilidade de  $3,5 \leq \bar{X}_1 - \bar{X}_2 \leq 5,5$ .

**7.2.8** Cientistas do Hopkins Memorial Forest, no oeste de Massachusetts, coletaram dados meteorológicos e ambientais da floresta por mais de 100 anos. Nos últimos anos, o teor médio de sulfato em amostras de água de Birch Brook foi 7,48 mg/L, com um desvio-padrão de 1,60 mg/L.

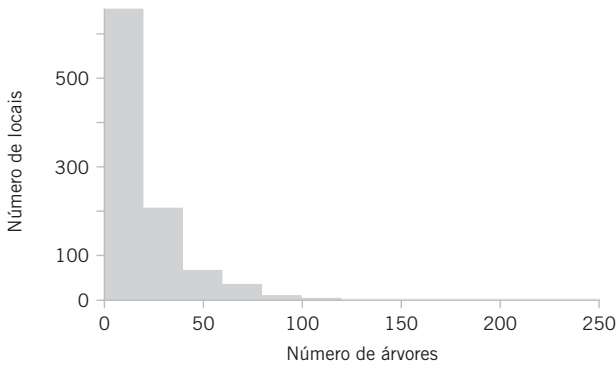
- Qual é o erro-padrão do sulfato em uma coleção com dez amostras de água?
- Se dez estudantes medirem o sulfato em suas amostras, qual será a probabilidade de que o teor médio de sulfato esteja entre 6,49 e 8,47 mg/L?
- O que você precisa considerar para que a probabilidade calculada em (b) seja acurada?

**7.2.9** Dados do pH da chuva em Ingham County, em Michigan, são mostrados:

5,47 5,37 5,38 4,63 5,37 3,74 3,71 4,96 4,64 5,11 5,65 5,39  
4,16 5,62 4,57 4,64 5,48 4,57 4,57 4,51 4,86 4,56 4,61 4,32  
3,98 5,70 4,15 3,98 5,65 3,10 5,04 4,62 4,51 4,34 4,16 4,64  
5,12 3,71 4,64

Qual é a proporção das amostras que têm pH abaixo de 5,0?

7.2.10 Pesquisadores do Hopkins Memorial Forest (veja o Exercício 7.2.8) também contaram o número de árvores de carvalho (gênero *acer*) em locais ao longo de toda a floresta. A seguir, é apresentado um histograma do número de árvores vivas em 1002 locais amostrados ao longo dos últimos 20 anos. O número médio de árvores por local foi de 19,86 árvores, com um desvio-padrão de 23,65 árvores.



- a. Se tomássemos a média de uma amostra de oito locais, qual seria o erro-padrão da média?
- b. Usando o teorema central do limite, qual é a probabilidade de que a média dos oito locais estivesse dentro de um erro-padrão da média?
- c. Por que você poderia pensar que a probabilidade que você calculou em (b) não poderia ser acurada?

7.2.11 Assim como furacões e terremotos, tempestades geomagnéticas são perigos naturais com possíveis impactos severos sobre a Terra. Tempestades severas podem causar colapsos nas comunicações e nas utilidades, levando a possíveis apagões. O órgão americano *National Oceanic and Atmospheric Administration* tem vários dados de fluxo de elétrons e de prótons em várias faixas de energia para várias estações na Terra, de modo a ajudar possíveis perturbações de previsão. A seguir, as 25 leituras do fluxo de próton na faixa de 47-68 keV [as unidades estão em p/(cm<sup>2</sup>-s-ster-Mev)] na noite de 28 de dezembro de 2011:

2310 2320 2010 10800 2190 3360 5640 2540 3360 11800 2010  
 3430 10600 7370 2160 3200 2020 2850 3500 10200 8550 9500  
 2260 7730 2250

- a. Encontre uma estimativa pontual do fluxo médio de prótons nesse período de tempo.
- b. Encontre uma estimativa pontual do desvio-padrão do fluxo médio de prótons nesse período de tempo.
- c. Encontre uma estimativa do erro-padrão da estimativa do item (a).
- d. Encontre uma estimativa pontual para a mediana do fluxo de prótons nesse período de tempo.
- e. Encontre uma estimativa pontual para a proporção das leituras que são menores do que 5000 p/(cm<sup>2</sup>-s-ster-Mev).

7.2.12 Wayne Collier projetou um experimento para medir a eficiência de combustível do carro de sua família sob diferentes pressões dos pneus. Para cada corrida, ele estabeleceu a pressão do pneu e então mediu as milhas que ele dirigiu em uma autoestrada até que os dois litros de combustível, colocados de cada vez, tivessem acabado. Para fazer isso, foram feitas algumas alterações no escoamento normal de gasolina para o motor. Eis as palavras de Wayne: “Eu inseri uma junção T na linha de combustível, logo antes do filtro de combustível, e uma linha para o compartimento do passageiro do meu carro, onde foi conectada uma garrafa graduada de dois litros que eu montei em uma caixa onde o assento do passageiro normalmente é preso. Então, eu selei a linha de retorno de combustível, a qual em operação normal envia gasolina de uma bomba de volta para o tanque de combustível”.

Suponha que você chame de  $\mu$  as milhas médias que ele pode dirigir com pressão normal nos pneus. Um estimador não tendencioso para  $\mu$  é a média das corridas amostrais,  $\bar{x}$ . Mas Wayne tem uma ideia diferente. Ele decide usar o seguinte estimador: Ele vira uma moeda não viciada. Se der cara, ele adicionará mais cinco milhas a cada observação. Se der coroa, ele subtrairá cinco milhas de cada observação.

- a. Mostre que a estimativa de Wayne é, de fato, não tendenciosa.
- b. Compare o desvio-padrão da estimativa de Wayne com o desvio-padrão da média da amostra.
- c. Dada a sua resposta em (b), por que a estimativa de Wayne não faz sentido cientificamente?

## Exercícios para a Seção 7.3

7.3.1 Um *software* foi usado para calcular alguns sumários numéricos de uma amostra de dados. Os resultados são dispostos aqui:

Variável	$N$	Média	EP da Média	Desvio-Padrão	Variância	Soma	Soma dos Quadrados
$x$	?	?	2,05	10,25	?	3761,70	?

- a. Preencha os valores que faltam.
- b. Qual é a estimativa da média da população a partir da qual essa amostra foi retirada?

7.3.2 Suponha que tenhamos uma amostra aleatória de tamanho  $2n$ , proveniente de uma população denotada por  $X$ , e  $E(X) = \mu$  e  $V(X) = \sigma^2$ . Sejam

$$\bar{X}_1 = \frac{1}{2n} \sum_{i=1}^{2n} X_i \quad \text{e} \quad \bar{X}_2 = \frac{1}{n} \sum_{i=1}^n X_i$$

dois estimadores de  $\mu$ . Qual é o melhor estimador de  $\mu$ ? Explique sua escolha.

7.3.3 Suponha que tenhamos uma amostra aleatória  $X_1, X_2, \dots, X_n$  proveniente de uma população que seja  $N(\mu, \sigma^2)$ . Planejamos usar  $\hat{\theta} = \sum_{i=1}^n (X_i - \bar{X})^2 / c$  para estimar  $\sigma^2$ . Calcule a tendência em  $\hat{\theta}$ , como um estimador de  $\sigma^2$ , em função da constante  $c$ .

7.3.4 Suponha que  $\hat{\theta}_1$  e  $\hat{\theta}_2$  sejam estimadores não tendenciosos do parâmetro  $\theta$ . Sabemos que  $V(\hat{\theta}_1) = 10$  e  $V(\hat{\theta}_2) = 4$ . Qual é o melhor estimador e em que sentido ele é melhor? Calcule a eficiência relativa dos dois estimadores.

7.3.5 Sejam  $\hat{\theta}_1, \hat{\theta}_2$  e  $\hat{\theta}_3$  estimadores do parâmetro  $\theta$ . Sabemos que  $E(\hat{\theta}_1) = E(\hat{\theta}_2) = \theta$ ,  $E(\hat{\theta}_3) \neq \theta$ ,  $V(\hat{\theta}_1) = 12$ ,  $V(\hat{\theta}_2) = 10$  e  $E(\hat{\theta}_3 - \theta)^2 = 6$ . Compare esses três estimadores. Qual você prefere? Por quê?

7.3.6 Dados sobre a força (libras) de remoção de conectores usados em um motor de automóveis são os seguintes: 79,3; 75,1; 78,2; 74,1; 73,9; 75,0; 77,6; 77,3; 73,8; 74,6; 75,5; 74,0; 74,7; 75,9; 72,9; 73,8; 74,2; 78,1; 75,4; 76,3; 75,3; 76,2; 74,9; 78,0; 75,1; 76,8.

- a. Calcule a estimativa pontual da média da força de remoção de todos os conectores na população. Estabeleça que estimador você usou e por quê.
- b. Calcule a estimativa pontual da média da força de remoção que separa os 50 % dos conectores mais fracos da população dos 50 % dos conectores mais fortes.
- c. Calcule as estimativas pontuais da variância e do desvio-padrão da população.
- d. Calcule o erro-padrão da estimativa pontual encontrada no item (a). Interprete o erro-padrão.
- e. Calcule a estimativa pontual da proporção de todos os conectores na população cuja força de remoção é menor do que 73 libras.

7.3.7 Dados sobre a espessura de óxido de semicondutores são os seguintes: 425, 431, 416, 419, 421, 436, 418, 410, 431, 433, 423, 426, 410, 435, 436, 428, 411, 426, 409, 437, 422, 428, 413, 416.

- a. Calcule a estimativa pontual da média da espessura do óxido para todas as pastilhas na população.
- b. Calcule a estimativa pontual do desvio-padrão da espessura do óxido para todas as pastilhas na população.
- c. Calcule o erro-padrão da estimativa pontual do item (a).
- d. Calcule a estimativa pontual da mediana da espessura do óxido para todas as pastilhas na população.
- e. Calcule a estimativa pontual da proporção de pastilhas na população que tem uma espessura de óxido maior do que 430 angströms.

7.3.8 Suponha que  $X$  seja o número de “sucessos” observados em uma amostra de  $n$  observações, em que  $p$  é a probabilidade de sucesso em cada observação.

- a. Mostre que  $\hat{P} = X/n$  é um estimador não tendencioso de  $p$ .
- b. Mostre que o erro-padrão de  $\hat{P}$  é  $\sqrt{p(1-p)/n}$ . Como você estimaria o erro-padrão?

7.3.9  $\bar{X}_1$  e  $S_1^2$  são a média e a variância da amostra de uma população com média  $\mu_1$  e variância  $\sigma_1^2$ . Similarmente,  $\bar{X}_2$  e  $S_2^2$  são a média e a variância da amostra de uma segunda população independente com média  $\mu_2$  e variância  $\sigma_2^2$ . Os tamanhos das amostras são  $n_1$  e  $n_2$ , respectivamente.

- a. Mostre que  $\bar{X}_1 - \bar{X}_2$  é um estimador não tendencioso de  $\mu_1 - \mu_2$ .
- b. Encontre o erro-padrão de  $\bar{X}_1 - \bar{X}_2$ . Como você poderia estimar o erro-padrão?
- c. Suponha que ambas as populações tenham a mesma variância; isto é,  $\sigma_1^2 = \sigma_2^2 = \sigma^2$ . Mostre que

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

é um estimador não tendencioso de  $\sigma^2$ .

7.3.10 Considere uma variável aleatória normal padrão, com média 10 e desvio-padrão 4. Suponha que uma amostra aleatória de tamanho 16 seja retirada dessa distribuição e que a média da amostra seja calculada. Sabemos que o erro-padrão da média da amostra nesse caso é  $\sigma/\sqrt{n} = \sigma/\sqrt{16} = 1$ . Utilize o método *bootstrap* com  $n_b = 200$  amostras *bootstraps* para encontrar o erro-padrão da média da amostra. Compare o erro-padrão pela técnica *bootstrap* com o erro-padrão real.

7.3.11 Dos  $n_1$  estudantes de engenharia selecionados aleatoriamente da Arizona State University (ASU),  $X_1$  possuem um computador da Apple, e dos  $n_2$  estudantes de engenharia selecionados aleatoriamente do Instituto Tecnológico de Virgínia (Virginia Tech),  $X_2$  possuem um computador da Apple. Sejam  $p_1$  e  $p_2$  as probabilidades de que estudantes de engenharia, selecionados aleatoriamente, da ASU e do Virginia Tech, respectivamente, possuam computadores da Apple.

- a. Mostre que uma estimativa não tendenciosa para  $p_1 - p_2$  é  $(X_1/n_1) - (X_2/n_2)$ .
- b. Qual é o erro-padrão da estimativa pontual do item (a)?
- c. Como você calcularia uma estimativa pontual do erro-padrão calculado no item (b)?
- d. Suponha que  $n_1 = 200$ ,  $X_1 = 150$ ,  $n_2 = 250$  e  $X_2 = 185$ . Use os resultados do item (a) para calcular uma estimativa de  $p_1 - p_2$ .
- e. Use os resultados dos itens (b) a (d) para calcular uma estimativa do erro-padrão da estimativa.

7.3.12 Suponha que duas amostras aleatórias independentes (de tamanho  $n_1$  e  $n_2$ ), retiradas a partir de duas distribuições normais, estejam disponíveis. Usando o método *bootstrap*, explique como você estimaria o erro-padrão da diferença nas médias amostrais  $\bar{X}_1 - \bar{X}_2$ .

## Exercícios para a Seção 7.4

7.4.1 Seja  $X$  uma variável aleatória geométrica, com parâmetro  $p$ . Encontre o estimador de máxima verossimilhança de  $p$ , com base em uma amostra aleatória de tamanho  $n$ .

7.4.2 Considere a função densidade de probabilidade

$$f(x) = \frac{1}{\theta^2} x e^{-x/\theta}, \quad 0 \leq x < \infty, \quad 0 < \theta < \infty$$

Encontre o estimador de máxima verossimilhança de  $\theta$ .

7.4.3 Sejam  $X_1, X_2, \dots, X_n$  uniformemente distribuídos no intervalo 0 a  $a$ . Mostre que o estimador de momento de  $a$  é  $\hat{a} = 2\bar{X}$ . Ele é um estimador não tendencioso? Discuta quão razoável é esse estimador.

7.4.4 Considere a função densidade de probabilidade

$$f(x) = c(1 + \theta x), \quad -1 \leq x \leq 1$$

- a. Encontre o valor da constante  $c$ .
- b. Qual é o estimador de momento para  $\theta$ ?
- c. Mostre que  $\hat{\theta} = 3\bar{X}$  é um estimador não tendencioso para  $\theta$ .
- d. Encontre o estimador de máxima verossimilhança para  $\theta$ .

7.4.5 Considere a distribuição de Weibull

$$f(x) = \begin{cases} \frac{\beta}{\delta} \left(\frac{x}{\delta}\right)^{\beta-1} e^{-\left(\frac{x}{\delta}\right)^\beta}, & 0 < x \\ 0, & \text{caso contrário} \end{cases}$$

- a. Encontre a função verossimilhança, com base em uma amostra aleatória de tamanho  $n$ . Encontre o logaritmo da verossimilhança.
- b. Mostre que o logaritmo da verossimilhança é maximizado, resolvendo as equações

$$\beta = \left[ \frac{\sum_{i=1}^n x_i^\beta \ln(x_i)}{\sum_{i=1}^n x_i^\beta} - \frac{\sum_{i=1}^n \ln(x_i)}{n} \right]^{-1} \quad \delta = \left[ \frac{\sum_{i=1}^n x_i^\beta}{n} \right]^{1/\beta}$$

- c. Quais são as complicações envolvidas na resolução das duas equações do item (b)?

7.4.6 Sejam  $X_1, X_2, \dots, X_n$  uniformemente distribuídos no intervalo 0 a  $a$ . Lembre-se de que o estimador de máxima verossimilhança de  $a$  é  $\hat{a} = \max(X_i)$ .

- Argunte intuitivamente por que  $\hat{a}$  não pode ser um estimador não tendencioso para  $a$ .
- Suponha que  $E(\hat{a}) = na/(n+1)$ . É razoável que  $\hat{a}$  consistentemente subestime  $a$ ? Mostre que a tendência no estimador se aproxima de zero quando  $n$  aumenta.
- Proponha um estimador não tendencioso para  $a$ .
- Seja  $Y = \max(X_i)$ . Use o fato de que  $Y \leq y$  se, e somente se, cada  $X_i \leq y$ , para deduzir a função de distribuição cumulativa de  $Y$ . Mostre então que a função densidade de probabilidade de  $Y$  é

$$f(y) = \begin{cases} \frac{ny^{n-1}}{a^n}, & 0 \leq y \leq a \\ 0, & \text{caso contrário} \end{cases}$$

Use esse resultado para mostrar que o estimador de máxima verossimilhança para  $a$  é tendencioso.

e. Suponhamos dois estimadores não tendenciosos para  $a$ : o estimador de momento  $\hat{a}_1 = 2\bar{X}$  e  $\hat{a}_2 = [(n+1)/n] \max(X_i)$ , em que  $\max(X_i)$  é a maior observação em uma amostra aleatória de tamanho  $n$ . Podemos mostrar que  $V(\hat{a}_1) = a^2/(3n)$  e que  $V(\hat{a}_2) = a^2/[n(n+2)]$ . Mostre que, se  $n > 1$ ,  $\hat{a}_2$  é um estimador melhor do que  $\hat{a}$ . Em que sentido ele é um estimador melhor de  $a$ ?

7.4.7 Reconsidere os dados da espessura do óxido no Exercício 7.3.7 e suponha que seja razoável considerar que a espessura do óxido seja normalmente distribuída.

- Calcule as estimativas de máxima verossimilhança de  $\mu$  e de  $\sigma^2$ .
- Faça um gráfico da função verossimilhança na vizinhança de  $\hat{\mu}$  e  $\hat{\sigma}^2$ , as estimativas de máxima verossimilhança, e comente a sua forma.
- Suponha que o tamanho da amostra foi maior ( $n = 40$ ), porém as estimativas de máxima verossimilhança foram numericamente iguais aos valores obtidos no item (a). Faça um gráfico da função verossimilhança para  $n = 40$ , compare-o àquele do item (b), e comente o efeito do maior tamanho de amostra.

## Exercícios Suplementares para o Capítulo 7

7.S8 Suponha que uma variável aleatória seja normalmente distribuída, com média  $\mu$  e variância  $\sigma^2$ . Dessa distribuição, retiramos uma amostra aleatória com cinco observações. Qual será a função densidade de probabilidade conjunta da amostra?

7.S9 Transistores têm uma vida que é distribuída exponencialmente, com parâmetro  $\lambda$ . Uma amostra aleatória de  $n$  transistores é retirada. Qual será a função densidade de probabilidade conjunta da amostra?

7.S10 Um especialista comprou 25 resistores do vendedor 1 e 30 resistores do vendedor 2. Sejam  $X_{1,1}, X_{1,2}, \dots, X_{1,25}$  as resistências do vendedor 1, que são consideradas normal e independentemente distribuídas, com média igual a 100 V e desvio-padrão igual a 1,5 V. Similarmente, sejam  $X_{2,1}, X_{2,2}, \dots, X_{2,30}$  as resistências do vendedor 2, que são consideradas normal e independentemente distribuídas, com média igual a 105 V e desvio-padrão igual a 2,0 V. Qual é a distribuição amostral de  $\bar{X}_1 - \bar{X}_2$ ? Qual é o erro-padrão de  $\bar{X}_1 - \bar{X}_2$ ?

7.S11 Uma amostra aleatória de 36 observações foi retirada de uma distribuição normal, com média igual a 50 e desvio-padrão igual a 12. Encontre a probabilidade de a média da amostra estar no intervalo  $47 \leq X \leq 53$ . A suposição de normalidade é importante? Por quê?

7.S12 Um fabricante de dispositivos semicondutores retira uma amostra aleatória de 100 *chips* e os testa, classificando cada *chip* como defeituoso ou não defeituoso. Seja  $X_i = 0$ , se o *chip* for não defeituoso, e  $X_i = 1$ , se o *chip* for defeituoso. A fração da amostra com defeitos é

$$\hat{p} = \frac{X_1 + X_2 + \dots + X_{100}}{100}$$

Qual é a distribuição amostral da variável aleatória  $\hat{p}$ ?

7.S13 Uma variável aleatória  $x$  tem função densidade de probabilidade

$$f(x) = \frac{1}{2\theta^3} x^2 e^{-x/\theta}, \quad 0 < x < \infty, \quad 0 < \theta < \infty$$

Encontre o estimador de máxima verossimilhança para  $\theta$ .

7.S14 Seja  $f(x) = (1/\theta)x^{(1-\theta)/\theta}$ ,  $0 < x < 1$  e  $0 < \theta < \infty$ . Mostre que  $\hat{\theta} = -(1/n) \sum_{i=1}^n \ln(X_i)$  é um estimador de máxima verossimilhança para  $\theta$  e que  $\hat{\theta}$  é um estimador não tendencioso para  $\theta$ .

7.S15 Uma empresa de eletricidade instalou medidores especiais em dez casas de uma subdivisão que mede o consumo (demanda) de energia a cada hora do dia. A companhia está interessada na demanda de energia em uma hora específica – a hora em que o sistema experimenta o pico de consumo. Os dados (em kW) desses dez medidores são: 23,1; 15,6; 17,4; 20,1; 19,8; 26,4; 25,1; 20,5; 21,9 e 28,7. Se  $\mu$  for o verdadeiro pico médio da demanda para as dez casas nesse grupo tendo medidores especiais, estime  $\mu$ . Agora, suponha que a empresa queira estimar a demanda na hora do pico para todas as 5000 casas nessa subdivisão. Seja  $\theta$  essa grandeza. Estime  $\theta$  usando os dados fornecidos anteriormente. Estime a proporção de casas na subdivisão que demanda no mínimo 20 kW na hora de pico do sistema.

# Capítulo 8 Exercícios

## Exercícios para a Seção 8.1

8.1.1 Para uma população normal com variância conhecida  $\sigma^2$ , responda as seguintes questões:

- Qual é o nível de confiança para o intervalo  $\bar{x} - 2,14\sigma/\sqrt{n} \leq \mu \leq \bar{x} + 2,14\sigma/\sqrt{n}$ ?
- Qual é o nível de confiança para o intervalo  $\bar{x} - 2,49\sigma/\sqrt{n} \leq \mu \leq \bar{x} + 2,49\sigma/\sqrt{n}$ ?
- Qual é o nível de confiança para o intervalo  $\bar{x} - 1,85\sigma/\sqrt{n} \leq \mu \leq \bar{x} + 1,85\sigma/\sqrt{n}$ ?
- Qual é o nível de confiança para o intervalo  $\mu \leq \bar{x} + 2,00\sigma/\sqrt{n}$ ?
- Qual é o nível de confiança para o intervalo  $\bar{x} - 1,96\sigma/\sqrt{n} \leq \mu$ ?

8.1.2 Considere as expressões de intervalo unilateral de confiança para a média de uma população normal.

- Que valor de  $z_{\alpha}$  resultaria em um IC de 90 %?
- Que valor de  $z_{\alpha}$  resultaria em um IC de 95 %?
- Que valor de  $z_{\alpha}$  resultaria em um IC de 99 %?

8.1.3 Uma amostra aleatória foi retirada de uma distribuição normal e os seguintes intervalos de confiança construídos usando os mesmos dados: (38,02; 61,98) e (39,95; 60,05).

- Qual é o valor da média da amostra?
- Um desses intervalos é de 95 % de confiança e o outro é de 90 % de confiança. Qual deles é de 95 % de confiança e por quê?

8.1.4 Deseja-se obter uma estimativa de intervalo de confiança para o ganho em um circuito de um dispositivo semicondutor. Suponha que o ganho seja normalmente distribuído com desvio-padrão  $\sigma = 20$ .

- Quão grande  $n$  tem de ser, se o comprimento do IC de 95 % deve ser 40?
- Quão grande  $n$  tem de ser, se o comprimento do IC de 99 % deve ser 40?

8.1.5 Suponha que  $n = 100$  amostras aleatórias de água proveniente de um lago com água fresca foram retiradas, sendo medida a concentração (miligramas por litro) de cálcio. Um IC de 95 % para a concentração média de cálcio é  $0,49 \leq \mu \leq 0,82$ .

- Um IC de 99 %, calculado a partir dos dados da amostra, seria maior ou menor?
- Considere a seguinte afirmação: Há uma chance de 95 % de  $\mu$  estar entre 0,49 e 0,82. Essa afirmação é correta? Explique sua resposta.
- Considere a seguinte afirmação: Se  $n = 100$  amostras aleatórias de água proveniente do lago forem tomadas e o IC de 95 % para  $\mu$  for calculado e se esse processo for repetido 1000 vezes, 950 dos ICs conterão o valor verdadeiro de  $\mu$ . Essa afirmação está correta? Explique sua resposta.

8.1.6 O rendimento de um processo químico está sendo estudado. De experiências prévias com esse processo, sabe-se que o rendimento é normalmente distribuído e  $\sigma = 3$ . Os últimos cinco dias de operação da planta resultaram nos seguintes rendimentos percentuais: 91,6; 88,75; 90,8; 89,95 e 91,3. Encontre um intervalo bilateral com 95 % de confiança para o rendimento médio verdadeiro.

8.1.7 Um fabricante produz anéis para pistões de um motor de um carro. Sabe-se que o diâmetro do anel é distribuído normalmente com  $\sigma = 0,001$  milímetro. Uma amostra aleatória de 15 anéis tem um diâmetro médio de  $\bar{x} = 74,036$  milímetros.

- Construa um intervalo bilateral de confiança de 99 % para o diâmetro médio do anel do pistão.

- Construa um limite inferior de confiança de 99 % para o diâmetro médio do anel do pistão. Compare o limite inferior desse intervalo de confiança com o do item (a).

8.1.8 Um engenheiro civil está analisando a resistência à compressão do concreto. A resistência à compressão é distribuída normalmente com  $\sigma^2 = 1000$  (psi)<sup>2</sup>. Uma amostra aleatória de 12 corpos de prova tem uma resistência média à compressão de  $\bar{x} = 3250$  psi.

- Construa um intervalo bilateral de confiança de 95 % para a resistência média à compressão.
- Construa um intervalo bilateral de confiança de 99 % para a resistência média à compressão. Compare a largura desse intervalo de confiança com aquela calculada no item (a).

8.1.9 Suponha que no Exercício 8.1.8 desejássemos estimar a resistência à compressão, com um erro que fosse menor do que 15 psi, com 99 % de confiança. Qual o tamanho requerido da amostra?

8.1.10 Se o tamanho  $n$  da amostra for dobrado, de quanto o comprimento do IC para  $\mu$  na Equação 8.5 será reduzido? O que acontecerá ao comprimento do intervalo se o tamanho da amostra for aumentado quatro vezes?

8.1.11 De quanto o tamanho  $n$  da amostra deverá ser aumentado se o comprimento do IC para  $\mu$  na Equação 8.5 for reduzido à metade?

8.1.12 Ishikawa *et al.* ["Evaluation of Adhesiveness of *Acinetobacter* sp. Tol 5 to Abiotic Surfaces," *Journal of Bioscience and Bioengineering* (Vol. 113(6), pp. 719-725)] estudaram a adesão de vários filmes em superfícies sólidas para possível uso em tecnologias ambientais. Experimentos de adesão são conduzidos pela medição da absorvância em  $A_{590}$ . Suponha que, para a linhagem de bactérias *Acinetobacter*, cinco medidas tiveram as seguintes leituras: 2,69; 5,76; 2,67; 1,62 e 4,12 dinas-cm<sup>2</sup>. Admita que o desvio-padrão seja 0,66 dina-cm<sup>2</sup>.

- Encontre um intervalo de confiança de 95 % para a adesão média.
- Se os cientistas quiserem que o intervalo de confiança não seja maior do que 0,55 dina-cm<sup>2</sup>, quantas observações eles devem considerar?

8.1.13 Um artigo publicado no *Journal of Agricultural Science* ["The Use of Residual Maximum Likelihood to Model Grain Quality Characteristics of Wheat with Variety, Climatic and Nitrogen Fertilizer Effects" (1997, Vol. 128, pp. 135-142)] investigou médias do teor de proteína do grão cru de trigo (CP) e o número de queda de Hagberg (HFN) pesquisados no Reino Unido. A análise usou uma variedade de aplicações de fertilizante de nitrogênio (kg N/ha), temperatura (°C) e a quantidade mensal total de chuva (mm). Os dados mostrados a seguir descrevem temperaturas para o trigo crescido na Faculdade de Agricultura Harper Adams, entre 1982 e 1993. As temperaturas medidas em junho foram obtidas como segue:

15,2	14,2	14,0	12,2	14,4	12,5
14,3	14,2	13,5	11,8	15,2	

Considere que o desvio-padrão seja conhecido,  $\sigma = 0,5$ .

- Construa um intervalo bilateral de confiança de 99 % para a temperatura média.
- Construa um intervalo bilateral de confiança de 95 % para a temperatura média.
- Suponha que quiséssemos estar 95 % confiantes de que o erro na estimação da temperatura média fosse menor do que dois graus Celsius. Que tamanho da amostra deveria ser usado?
- Suponha que quiséssemos estar 95 % confiantes de que a largura total do intervalo bilateral de confiança para a temperatura média fosse 1,5 grau Celsius. Que tamanho da amostra deveria ser usado?

## Exercícios para a Seção 8.2

8.2.1 Uma amostra aleatória foi retirada a partir de uma distribuição normal. A saída de um *software* é dada a seguir:

Variável	$N$	Média	Erro-Padrão da Média	Desvio-Padrão	Variância	Soma
$x$	10	?	0,507	1,605	?	251,848

a. Preencha as grandezas que faltam.

b. Encontre um intervalo de confiança de 95 % para a média da população.

8.2.2 Encontre os valores dos seguintes percentis:  $t_{0,025;15}$ ,  $t_{0,05;10}$ ,  $t_{0,10;20}$ ,  $t_{0,005;25}$  e  $t_{0,001;30}$ .

8.2.3 Um engenheiro do setor de pesquisa de um fabricante de pneu está investigando a vida do pneu em relação a um novo componente da borracha. Ele fabricou 16 pneus e testou-os até o final da vida em um teste na estrada. A média e o desvio-padrão da amostra são 60.139,7 e 3645,94 quilômetros. Encontre um intervalo de confiança de 95 % para a vida média do pneu.

8.2.4 Determine o percentil  $t$  que é requerido para construir cada um dos seguintes intervalos unilaterais de confiança:

a. Nível de confiança = 95 %, graus de liberdade = 14.

b. Nível de confiança = 99 %, graus de liberdade = 19.

c. Nível de confiança = 99,9 %, graus de liberdade = 24.

8.2.5 Um artigo em *Obesity Research* ["Impaired Pressure Natriuresis in Obese Youths" (2003, Vol. 11, pp. 745-751)] descreveu um estudo em que todas as refeições foram fornecidas para 14 meninos magros durante três dias, seguidas por um estresse (com uma tarefa de videogame). A pressão sanguínea sistólica (PSS) média durante o teste foi 118,3 mm Hg, com um desvio-padrão de 9,9 mm Hg. Construa um intervalo unilateral superior de confiança de 99 % para a PSS média.

8.2.6 Um artigo em *Medicine and Science in Sports and Exercise* ["Maximal Leg-Strength Training Improves Cycling Economy in Previously Untrained Men" (2005, Vol. 37, pp. 131-136)] estudou o desempenho de pedalar antes e depois de oito semanas de treino de resistência da perna. Sete homens previamente não treinados fizeram o treino de resistência da perna três dias por semana durante oito semanas (com quatro séries de cinco repetições a 85 % de um máximo de repetição). O pico na potência durante o incremento de pedalada aumentou a média para 315 watts, com um desvio-padrão de 16 watts. Construa um intervalo de confiança de 95 % para a potência média de pico depois do treino.

8.2.7 O brilho de um tubo de imagem de televisão pode ser avaliado medindo-se a quantidade de corrente requerida para atingir determinado nível de brilho. Uma amostra de dez tubos resultou em  $\bar{x} = 317,2$  e  $s = 15,7$ . Encontre (em microampères) um intervalo de confiança de 99 % para a corrente média requerida. Estabeleça qualquer suposição necessária sobre a distribuição em foco dos dados.

8.2.8 Um artigo publicado no *Journal of Composite Materials* (dezembro de 1989, Vol. 23(12), pp. 1200-1215) descreve o efeito da delaminação na frequência natural de vigas feitas de laminados compostos. Cinco dessas vigas delaminadas foram submetidas a cargas, e as frequências (em hertz) resultantes foram:

230,66; 233,05; 232,58; 229,48; 232,58

Verifique a suposição de normalidade na população. Calcule um intervalo bilateral de confiança de 90 % para a frequência natural média.

8.2.9 A resistência do concreto à compressão está sendo testada por um engenheiro civil. Ele testa 12 corpos de prova e obtém os seguintes dados:

2216	2237	2249	2204
2225	2301	2281	2263
2318	2255	2275	2295

a. Verifique a suposição de que a resistência à compressão é normalmente distribuída. Inclua um gráfico na sua resposta.

b. Construa um intervalo bilateral de confiança de 95 % para a resistência média.

c. Construa um limite unilateral inferior de confiança de 95 % para a resistência média. Compare esse limite inferior com o limite inferior do intervalo bilateral de confiança e discuta por que eles são diferentes.

8.2.10 Um artigo em *Computers & Electrical Engineering* ["Parallel Simulation of Cellular Neural Networks" (1996, Vol. 22, pp. 61-84)] considerou a aceleração de redes neurais celulares (RNC) para uma arquitetura computacional paralela com finalidade geral, baseada em seis *transputers* em diferentes áreas. Os dados são mostrados a seguir:

3,775302	3,350679	4,217981	4,030324	4,639692
4,139665	4,395575	4,824257	4,268119	4,584193
4,930027	4,315973	4,600101		

a. Há evidência para suportar a suposição de que a aceleração de RNC é normalmente distribuída? Inclua um gráfico em sua resposta.

b. Construa um intervalo bilateral de confiança de 95 % para a aceleração média.

c. Construa um limite inferior de confiança de 95 % para a aceleração média.

8.2.11 Um artigo em *Nuclear Engineering International* (fevereiro de 1988, p. 33) descreve várias características de bastões combustíveis usados em um reator pertencente a uma utilidade elétrica na Noruega. Medidas da percentagem de enriquecimento de 12 bastões foram reportadas como segue:

2,94	3,00	2,90	2,75	3,00	2,95
2,90	2,75	2,95	2,82	2,81	3,05

a. Use um gráfico de probabilidade normal para verificar a suposição de normalidade.

b. Encontre um intervalo bilateral de confiança de 99 % para o percentual médio de enriquecimento. Você está confortável com a afirmação de que o percentual médio de enriquecimento é de 2,95 %? Por quê?

8.2.12 Um plano de saúde monitora o número de cateterismos feitos por mês em cada uma de suas clínicas. A seguir, são apresentados os dados do ano mais recente para uma clínica particular (a variável reportada é o número de cateterismos por mês expresso como o número de cateterismos por mil membros do plano de saúde):

2,31	2,09	2,36	1,95	1,98	2,25	2,16	2,07	1,88	1,94	1,97	2,02
------	------	------	------	------	------	------	------	------	------	------	------

a. Encontre um IC bilateral de 95 % para o número médio de cateterismos feitos por mês nessa clínica.

b. Historicamente, o número médio de exames feitos por todas as clínicas no sistema tem sido 1,95. Existe qualquer evidência de que essa clínica particular faça em média mais cateterismos do que a média de todo o sistema?

## Exercícios para a Seção 8.3

8.3.1 A percentagem de titânio em uma liga usada na fundição de aeronaves é medida em 51 peças selecionadas aleatoriamente. O desvio-padrão amostral é  $s = 0,37$ . Construa um intervalo bilateral de confiança de 95 % para  $\sigma$ .

8.3.2 Determine os valores dos seguintes percentis:

$$\chi^2_{0,05;10}, \chi^2_{0,025;15}, \chi^2_{0,01;12}, \chi^2_{0,95;20}, \chi^2_{0,99;18}, \chi^2_{0,995;16}, \text{ e } \chi^2_{0,005;25}$$

8.3.3 Um artigo em *Urban Ecosystems* ["Urbanization and Warming of Phoenix (Arizona, USA): Impacts, Feedbacks and Mitigation" (2002, Vol. 6, pp. 183-203)] menciona que Fênix é ideal para estudar os efeitos de uma ilha de aquecimento urbano porque ela cresceu de uma população de 300.000 para aproximadamente três milhões ao longo dos últimos 50 anos, sendo esse um período com um registro contínuo e detalhado do clima. As médias, nesses 50 anos, das temperaturas médias anuais em oito localidades de Fênix são mostradas a seguir. Com um gráfico de probabilidade, verifique a suposição de normalidade na população. Construa um intervalo de confiança de 95 % para o desvio-padrão das temperaturas médias anuais nas localidades.

Site	Temperatura Média (°C)
Sky Harbor Airport	23,3
Phoenix Greenway	21,7
Phoenix Encanto	21,6
Waddell	21,7
Litchfield	21,3
Laveen	20,7
Maricopa	20,9
Harlquahala	20,1

8.3.4 Um artigo em *Medicine and Science in Sports and Exercise* ["Electrostimulation Training Effects on the Physical Performance of Ice Hockey Players" (2005, Vol. 37, pp. 455-460)] considerou o uso de eletromioestimulação (EMS) como um método para treinar músculos saudáveis do esqueleto. Sessões de EMS consistiram em 30 contrações (duração de quatro segundos, 85 Hz) e foram executadas três vezes por semana durante três semanas em 17 jogadores de hóquei no gelo.

O teste de desempenho de patinar em dez metros mostrou um desvio-padrão de 0,09 segundo. Construa um intervalo de confiança de 95 % para o desvio-padrão do teste de desempenho de patinar.

8.3.5 Um artigo em *Technometrics* ["Two-Way Random Effects Analyses and Gauge R&R Studies" (1999, Vol. 41(3), pp. 202-211)] estudou a capacidade de um medidor medir o peso de papel. Os dados para medidas repetidas de uma folha de papel são mostrados na tabela a seguir. Construa um intervalo unilateral superior de confiança de 95 % para o desvio-padrão dessas medidas. Verifique a suposição de normalidade dos dados e comente sobre as suposições para o intervalo de confiança.

Observações				
3,481	3,448	3,485	3,475	3,472
3,477	3,472	3,464	3,472	3,470
3,470	3,470	3,477	3,473	3,474

8.3.6 Um artigo em *Cancer Research* ["Analyses of Litter-Matched Time-to-Response Data, with Modifications for Recovery of Interlitter Information" (1977, Vol. 37, pp. 3863-3868)] testou a capacidade de um medicamento na gênese de um tumor. Ratos foram selecionados aleatoriamente de ninhadas e receitados com o medicamento. Os tempos de aparecimento do tumor foram registrados conforme segue:

101, 104, 104, 77, 89, 88, 104, 96, 82, 70, 89, 91, 39, 103, 93, 85, 104, 104, 81, 67, 104, 104, 104, 87, 104, 89, 78, 104, 86, 76, 103, 102, 80, 45, 94, 104, 104, 76, 80, 72, 73

Calcule o intervalo de confiança de 95 % para o desvio-padrão do tempo até o aparecimento do tumor. Verifique a suposição de normalidade da população e comente sobre as suposições para o intervalo de confiança.

8.3.7 A partir dos dados de pH de chuva no Condado de Ingham, Michigan:

5,47 5,37 5,38 4,63 5,37 3,74 3,71 4,96 4,64 5,11 5,65 5,39 4,16 5,62 4,57 4,64 5,48 4,57 4,57 4,51 4,86 4,56 4,61 4,32 3,98 5,70 4,15 3,98 5,65 3,10 5,04 4,62 4,51 4,34 4,16 4,64 5,12 3,71 4,64

Encontre um intervalo de confiança bilateral de 95 % para o desvio-padrão de pH.

## Exercícios para a Seção 8.4

8.4.1 As pesquisas para a eleição presidencial de 2004 do crítico estado de Ohio forneceram os seguintes resultados. Havia 2020 pessoas consultadas nas pesquisas e 768 eram graduadas em faculdades. Das pessoas graduadas em faculdades, 412 votaram em George Bush.

- Calcule um intervalo de confiança de 95 % para a proporção de pessoas graduadas em Ohio que votaram em George Bush.
- Calcule o limite inferior de confiança de 95 % para a proporção de pessoas graduadas em Ohio que votaram em George Bush.

8.4.2 Um artigo em *Knee Surgery, Sports Traumatology, Arthroscopy* ["Arthroscopic Meniscal Repair with an Absorbable Screw: Results and Surgical Technique" (2005, Vol. 13, pp. 273-279)] mostrou que

somente 25 de 37 rompimentos (67,6 %), localizados entre 3 e 6 mm da borda do menisco, foram cicatrizados.

- Calcule um intervalo bilateral de confiança de 95 % para a proporção de tais rompimentos que cicatrizarão.
- Calcule um limite inferior de confiança de 95 % para a proporção de tais rompimentos que cicatrizarão.

8.4.3 Um artigo publicado no *Journal of the American Statistical Association* ["Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling" (1990, Vol. 85, pp. 972-985)] mediu o peso de 30 ratos sob controles experimentais. Suponha que haja 12 ratos abaixo do peso.

- a. Calcule um intervalo bilateral de confiança de 95 % para a verdadeira proporção de ratos que mostrariam subpeso a partir do experimento.
- b. Usando a estimativa pontual de  $p$ , obtida a partir da amostra preliminar, qual o tamanho necessário da amostra para estarmos 95 % confiantes de que o erro em estimar o valor verdadeiro de  $p$  seja menor do que 0,02?
- c. Quão grande deve ser a amostra se desejarmos estar no mínimo 95 % confiantes de que o erro em estimar  $p$  seja menor do que 0,02, independentemente do valor verdadeiro de  $p$ ?

**8.4.4** O Departamento de Transportes do Arizona deseja fazer uma pesquisa com residentes do estado para determinar que proporção da população gostaria de aumentar os limites de velocidade de autoestradas de 65 mph para 75 mph. Quantos residentes eles vão precisar examinar se eles quiserem estar no mínimo 99 % confiantes de que a proporção da amostra esteja dentro de 0,05 da proporção verdadeira?

**8.4.5** Os correios americanos usam o reconhecimento óptico de caracteres (OCR) desde a metade dos anos 1960. Em 1983, os correios começaram a implantar a tecnologia nas principais agências de todo o país ([www.britannica.com](http://www.britannica.com)). Suponha que em uma amostra aleatória de 500 dígitos dos CEPs (código de endereçamento postal) escritos a mão, 466 tenham sido lidos corretamente.

- a. Construa um intervalo de confiança de 95 % para a proporção verdadeira de dígitos corretos que podem ser lidos corretamente.
- b. Qual o tamanho da amostra necessário para reduzir a margem de erro para 1 %?
- c. Como a resposta para o item (b) mudaria se você tivesse considerado que a máquina de leitura lesse corretamente somente metade dos dígitos?

**8.4.6** Informações em um pacote de sementes indicam que 93 % delas germinarão. Das 200 sementes que você plantou, somente 180 germinaram.

- a. Com base nessa amostra, encontre um intervalo de confiança de 95 % para a proporção verdadeira de sementes que germinaram.
- b. Esse fato lhe fornece evidência de que a afirmação no pacote esteja errada?

**8.4.7** Use os dados do Exercício 8.4.5 para calcular o IC bilateral de Agresti-Coull para a proporção de dígitos lidos corretamente. Compare e discuta o IC em relação àquele calculado no Exercício 8.4.5.

**8.4.8** Use os dados do Exercício 8.4.2 para calcular o IC bilateral de Agresti-Coull para a proporção de lágrimas que vão curar. Compare e discuta o IC em relação àquele calculado no Exercício 8.4.2.

## Exercícios para a Seção 8.6

**8.6.1** Considere os dados dos testes de pneu, descritos no Exercício 8.2.3. Calcule um intervalo de previsão de 95 % para a vida do próximo pneu desse tipo, testado sob condições que são similares àquelas empregadas no teste original. Compare o comprimento do intervalo de previsão com o comprimento do IC de 95 % para a média da população.

**8.6.2** Considere a frequência natural das vigas descritas no Exercício 8.2.8. Calcule um intervalo de previsão de 90 % para o diâmetro da frequência natural da próxima viga desse tipo que será testada. Compare o comprimento do intervalo de previsão com o comprimento do IC de 90 % para a média da população.

**8.6.3** Considere o teste de brilho do tubo de televisão, descrito no Exercício 8.2.7. Calcule um intervalo de previsão de 99 % para o brilho do próximo tubo testado. Compare o comprimento do intervalo de previsão com o comprimento do IC de 99 % para a média da população.

**8.6.4** Considere o teste da resistência compressiva do concreto descrito no Exercício 8.2.9. Calcule um intervalo de previsão de 90 % para o próximo corpo de prova de concreto testado.

**8.6.5** Considere os dados de enriquecimento de bastões combustíveis, descritos no Exercício 8.2.11. Calcule um intervalo de previsão de 90 % para o enriquecimento do próximo bastão testado. Compare o comprimento do intervalo de previsão com o comprimento do IC de 99 % para a média da população.

**8.6.6** Como você obteria um limite unilateral de previsão para uma observação futura? Aplique esse procedimento para obter um limite unilateral de previsão de 95 % para a espessura da parede da próxima garrafa, para a situação descrita no Exercício 8.2.11.

**8.6.7** Considere os dados de testes de pneus do Exercício 8.2.3. Calcule um intervalo de tolerância de 95 % para a vida dos pneus que tem nível de confiança de 95 %. Compare o comprimento do intervalo de tolerância com o comprimento do IC de 95 % para a média da população. Qual é o menor intervalo? Discuta a diferença na interpretação desses dois intervalos.

**8.6.8** Considere os dados de brilho da televisão do Exercício 8.2.7. Calcule um intervalo de tolerância de 99 % para o brilho dos tubos de televisão que tem um nível de confiança de 95 %. Compare o comprimento do intervalo de previsão com o comprimento do IC de 99 % para a média da população. Qual é o menor intervalo? Discuta a diferença na interpretação desses dois intervalos.

**8.6.9** Considere os dados de enriquecimento de bastões de combustível do Exercício 8.2.11. Calcule um intervalo de tolerância de 99 % para o enriquecimento dos bastões que tem um nível de confiança de 95 %. Compare o comprimento do intervalo de tolerância com o comprimento do IC de 95 % para a média da população.

**8.6.10** Considere os dados de resistência do concreto do Exercício 8.2.9. Calcule um intervalo de tolerância de 90 % para a resistência compressiva do concreto que tem um nível de confiança de 90 %.

## Exercícios Suplementares para o Capítulo 8

8.S7 Considere o intervalo de confiança para  $\mu$ , com desvio-padrão conhecido  $\sigma$ :

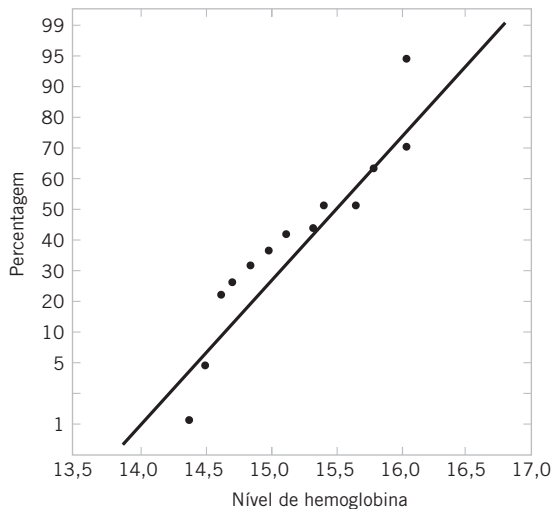
$$\bar{x} - z_{\alpha_1}\sigma/\sqrt{n} \leq \mu \leq \bar{x} + z_{\alpha_2}\sigma/\sqrt{n}$$

em que  $\alpha_1 + \alpha_2 = \alpha$ . Seja  $\alpha = 0,05$ . Encontre o intervalo para  $\alpha_1 = \alpha_2 = \alpha/2 = 0,025$ . Agora, encontre o intervalo para o caso de  $\alpha_1 = 0,01$  e  $\alpha_2 = 0,04$ . Qual é o menor intervalo? Há alguma vantagem em ter um intervalo de confiança “simétrico”?

8.S8 Um artigo publicado no *Journal of Sports Science* [“Iron Status in Winter Olympic Sports” (1987, Vol. 5(3), pp. 261-271)] apresentou os resultados de uma investigação do nível de hemoglobina de jogadores de hóquei no gelo nas Olimpíadas do Canadá. Os dados reportados (em g/dL) são mostrados a seguir:

15,3	16,0	14,4	16,2	16,2
14,9	15,7	15,3	14,6	15,7
16,0	15,0	15,7	16,2	14,7
14,8	14,6	15,6	14,5	15,2

a. Fornecido o seguinte gráfico de probabilidade dos dados, qual é a suposição lógica acerca da distribuição básica dos dados?



b. Explique por que essa verificação da distribuição baseada nos dados amostrais será importante, se quisermos construir um intervalo de confiança para a média.

c. Com base nesses dados amostrais, um intervalo de confiança de 95 % para a média é (15,04; 15,62). É razoável inferir que a média verdadeira poderia ser 14,5? Explique sua resposta.

d. Explique por que essa verificação da distribuição baseada nos dados amostrais será importante, se quisermos construir um intervalo de confiança para a variância.

e. Com base nesses dados amostrais, um intervalo de confiança de 95 % para a variância é [0,22; 0,82]. É razoável inferir que a variância verdadeira poderia ser 0,35? Explique sua resposta.

f. É razoável usar esses intervalos de confiança para inferir sobre a média e a variância dos níveis de hemoglobina

(i) de doutores canadenses? Explique sua resposta.

(ii) de crianças canadenses de 6 a 12 anos? Explique sua resposta.

8.S9 Uma população normal tem média  $\mu = 50$  e variância  $\sigma^2 = 5$  conhecidas. Mostre que a probabilidade aproximada de que a variância da amostra seja maior do que ou igual a 7,44, menor do que ou igual a 2,56, para uma amostra aleatória de tamanho

a.  $n = 16$ .      b.  $n = 30$ .      c.  $n = 71$ .

d. Compare suas respostas dos itens (a) a (c) com a probabilidade aproximada de que a variância amostral é maior do que ou igual a 7,44. Explique por que essa probabilidade da extremidade está crescendo ou diminuindo com o aumento do tamanho da amostra.

e. Compare suas respostas dos itens (a) a (c) com a probabilidade aproximada de que a variância amostral é menor do que ou igual a 2,56. Explique por que essa probabilidade da extremidade está crescendo ou diminuindo com o aumento do tamanho da amostra.

8.S10 Durante as temporadas de 1999 e 2000 de beisebol, houve muita especulação de que o grande número, pouco comum, de corridas à base (*home runs*) executadas foi em razão, no mínimo, em parte, de uma bola com mais energia. Uma maneira de testar a “disposição” de uma bola de beisebol é lançá-la em direção a uma superfície vertical, com uma velocidade conhecida  $V_L$ , e medir a razão da velocidade de saída  $V_O$  da bola em relação a  $V_L$ . A razão  $R = V_O/V_L$  é chamada de *coeficiente de restituição*. A seguir, são apresentadas as medidas de coeficiente de restituição para 40 bolas de beisebol selecionadas aleatoriamente. As bolas foram arremessadas em direção a uma superfície de carvalho, por meio de uma máquina de lançamento.

0,6248	0,6237	0,6118	0,6159	0,6298	0,6192
0,6520	0,6368	0,6220	0,6151	0,6121	0,6548
0,6226	0,6280	0,6096	0,6300	0,6107	0,6392
0,6230	0,6131	0,6223	0,6297	0,6435	0,5978
0,6351	0,6275	0,6261	0,6262	0,6262	0,6314
0,6128	0,6403	0,6521	0,6049	0,6170	
0,6134	0,6310	0,6065	0,6214	0,6141	

a. Há evidência que suporte a suposição de que o coeficiente de restituição seja normalmente distribuído?

b. Encontre um IC de 99 % para o coeficiente médio de restituição.

c. Encontre um intervalo de previsão de 99 % para o coeficiente de restituição para a próxima bola de beisebol que será testada.

d. Encontre um intervalo que conterá 99 % dos valores do coeficiente de restituição com 95 % de confiança.

e. Explique a diferença entre os três intervalos calculados nos itens (b), (c) e (d).

8.S11 Considere os dados do coeficiente de restituição da bola de beisebol do Exercício 8.S10. Suponha que qualquer bola que tenha um coeficiente de restituição que exceda 0,635 seja considerada muito ativa. Com base nos dados disponíveis, que proporção das bolas de beisebol na população amostrada corresponde a bolas muito ativas? Encontre um limite inferior de 95 % para essa proporção.

8.S12 Um artigo publicado no *ASCE Journal of Energy Engineering* [“Overview of Reservoir Release Improvements at 20 TVA Dams” (Vol. 125, abril 1999, pp. 1-17)] apresenta, a seguir, dados para concentrações de oxigênio dissolvido em correntes em 20 barragens do sistema do Vale do Tennessee. As observações são (em miligramas por litro): 5,0; 3,4; 3,9; 1,3; 0,2; 0,9; 2,7; 3,7; 3,8; 4,1; 1,0; 1,0; 0,8; 0,4; 3,8; 4,5; 5,3; 6,1; 6,9 e 6,5.

a. Há evidência que suporte a suposição de que a concentração de oxigênio dissolvido seja normalmente distribuída?

- b. Encontre um IC de 95 % para a concentração média de oxigênio dissolvido.
- c. Encontre um intervalo de previsão de 95 % para a concentração média de oxigênio dissolvido para a próxima corrente do sistema que será testada.
- d. Encontre um intervalo que conterá 95 % dos valores da concentração média de oxigênio dissolvido com 99 % de confiança.
- e. Explique a diferença entre os três intervalos calculados nos itens (b), (c) e (d).

8.S13 Um artigo em *The Engineer* (“Redesign for Suspect Wiring”, junho de 1990) reportou os resultados de uma investigação sobre erros relativos ao cabeamento em aeronaves comerciais de passageiros, que podem produzir informações defeituosas à tripulação. Esse erro relativo ao cabeamento pode ter sido responsável pelo desastre de um avião da *British Midland Airways*, em janeiro de 1989, fazendo com que o piloto desligasse o motor errado. De 1600 aviões selecionados aleatoriamente, oito tinham erros relativos ao cabeamento que poderiam mostrar informação incorreta à tripulação.

- a. Encontre um intervalo de confiança de 99 % para a proporção de aeronaves que tenham tais erros relativos ao cabeamento.
- b. Suponha que usemos a informação nesse exemplo, de modo a fornecer uma estimativa preliminar de  $p$ . Quão grande a amostra seria para produzir uma estimativa de  $p$  diferente do valor

verdadeiro por no máximo 0,008, de modo a nos deixar 99 % confiantes?

c. Suponha que não tivéssemos uma estimativa preliminar de  $p$ . Quão grande a amostra seria, se quiséssemos estar no mínimo 99 % confiantes de que a proporção amostral seria diferente do valor verdadeiro por no máximo 0,008, independentemente do valor verdadeiro de  $p$ ?

d. Comente a utilidade da informação preliminar para o cálculo do tamanho necessário da amostra.

8.S14 Um artigo publicado no *Journal of Human Nutrition and Dietetics* [“The Validation of Energy and Protein Intakes by Doubly Labeled Water and 24-Hour Urinary Nitrogen Excretion in Post-Obese Subjects” (1995, Vol. 8, pp. 51-64)] mostrou a energia admitida expressa como uma taxa metabólica basal (TMB), dada em MJ.

5,40	5,67	5,79	6,85	6,92
5,70	6,08	5,48	5,44	5,51

a. Use um gráfico de probabilidade normal para verificar a suposição de normalidade.

b. Encontre um intervalo bilateral de confiança de 99 % para a média de TMB.

## Capítulo 9 Exercícios

### Exercícios para a Seção 9.1

9.1.1 Em cada uma das seguintes situações, estabeleça se esse é um problema corretamente posto de teste de hipóteses e por quê.

- a.  $H_0: \mu = 25, H_1: \mu \neq 25$     b.  $H_0: \sigma > 10, H_1: \sigma = 10$   
 c.  $H_0: \bar{x} = 50, H_1: \bar{x} \neq 50$     d.  $H_0: p = 0,1, H_1: p = 0,5$   
 e.  $H_0: s = 30, H_1: s > 30$

9.1.2 Um fabricante de semicondutores coleta dados provenientes de uma nova ferramenta e conduz um teste de hipóteses, com a hipótese nula sendo a largura média de uma dimensão crítica igual a 100 nm. A conclusão é não rejeitar a hipótese nula. Esse resultado fornece uma evidência forte de que a média da dimensão crítica é igual a 100 nm? Explique.

9.1.3 Um fabricante de fibra têxtil está investigando um novo fio, que a companhia afirma ter um alongamento médio de 12 quilogramas, com desvio-padrão de 0,5 quilograma. A companhia deseja testar a hipótese  $H_0: \mu = 12$  contra  $H_1: \mu < 12$ , usando uma amostra aleatória de quatro espécimes.

- a. Qual será a probabilidade do erro tipo I, se a região crítica for definida como  $\bar{x} < 11,5$  quilogramas?
- b. Encontre  $\beta$  para o caso em que o alongamento médio verdadeiro seja de 11,25 quilogramas.
- c. Encontre  $\beta$  para o caso em que a média verdadeira seja de 11,5 quilogramas.

9.1.4 Repita o Exercício 9.1.3, usando um tamanho de amostra de  $n = 16$  e a mesma região crítica.

9.1.5 No Exercício 9.1.3, calcule o valor  $P$ , se a estatística observada for

a.  $\bar{x} = 11,25$     b.  $\bar{x} = 11,0$     c.  $\bar{x} = 11,75$

9.1.6 No Exercício 9.1.3, calcule a probabilidade de um erro tipo II, se a alongação média verdadeira for de 11,5 quilogramas e

- a.  $\alpha = 0,05$  e  $n = 4$     b.  $\alpha = 0,05$  e  $n = 16$   
 c. Compare os valores de  $\beta$  calculados nos itens anteriores. Que conclusão você pode tirar?

9.1.7 Uma companhia de produtos para consumidor está formulando um xampu novo e está interessada na altura (em milímetros) da espuma. A altura da espuma tem distribuição aproximadamente normal, com um desvio-padrão de 20 milímetros. A companhia deseja testar  $H_0: \mu = 175$  milímetros versus  $H_1: \mu > 175$  milímetros, usando os resultados de  $n = 10$  amostras.

- a. Encontre a probabilidade  $\alpha$  do erro tipo I, se a região crítica for  $\bar{x} > 185$ .
- b. Qual será a probabilidade do erro tipo II, se a altura média verdadeira da espuma for 185 milímetros?
- c. Encontre  $\beta$  para a média verdadeira de 195 milímetros.

9.1.8 No Exercício 9.1.7, calcule a probabilidade de um erro tipo II, se a verdadeira altura média de espuma for 185 milímetros e

- a.  $\alpha = 0,05$  e  $n = 10$     b.  $\alpha = 0,05$  e  $n = 16$   
 c. Compare os valores de  $\beta$  calculados nos itens anteriores. Que conclusão você pode tirar?

9.1.9 No Exercício 9.1.7, encontre o limite da região crítica se a probabilidade do erro tipo I for

- a.  $\alpha = 0,01$  e  $n = 10$     b.  $\alpha = 0,05$  e  $n = 10$   
 c.  $\alpha = 0,01$  e  $n = 16$     d.  $\alpha = 0,05$  e  $n = 16$

9.1.10 Um fabricante está interessado na voltagem de saída de um fornecimento de energia usado em um computador pessoal. A voltagem

de saída é considerada normalmente distribuída, com desvio-padrão igual a 0,25 volt. O fabricante deseja testar  $H_0: \mu = 5$  volts contra  $H_1: \mu \neq 5$  volts, usando  $n = 8$  unidades.

a. A região de aceitação é  $4,85 \leq \bar{x} \leq 5,15$ . Encontre o valor de  $\alpha$ .

b. Encontre a potência do teste para detectar uma voltagem de saída média verdadeira de 5,1 volts.

**9.1.11** A proporção de adultos com terceiro grau vivendo em Tempe, Arizona, é estimada como  $p = 0,4$ . Para testar essa hipótese, seleciona-se uma amostra aleatória de 15 adultos. Se o número de graduados estiver entre quatro e oito, a hipótese será aceita; do contrário, concluiremos que  $p \neq 0,4$ .

a. Encontre a probabilidade do erro tipo I para esse procedimento, considerando  $p = 0,4$ .

b. Encontre a probabilidade de cometer um erro tipo II se a proporção verdadeira for realmente  $p = 0,2$ .

**9.1.12** Acredita-se que  $p = 0,3$  seja a proporção, em Fênix, de moradores favoráveis à construção de estradas com pedágios, de modo a completar o sistema de rodovias. Se uma amostra aleatória de dez

moradores mostrar que um ou menos é favorável a essa proposta, concluiremos que  $p < 0,3$ .

a. Encontre a probabilidade do erro tipo I se a proporção verdadeira for  $p = 0,3$ .

b. Encontre a probabilidade de cometer um erro tipo II com esse procedimento se  $p = 0,2$ .

c. Qual será a potência desse procedimento se a proporção verdadeira for  $p = 0,2$ ?

**9.1.13** Se plotarmos a probabilidade de aceitar  $H_0: \mu = \mu_0$  versus vários valores de  $\mu$  e conectarmos os pontos com uma curva suave, obteremos a **curva característica operacional** (ou a **curva CO**) do procedimento de teste. Essas curvas são usadas extensivamente em aplicações industriais de teste de hipóteses para mostrar a sensibilidade e o desempenho relativo do teste. Quando a média verdadeira é realmente igual a  $\mu_0$ , a probabilidade de aceitar  $H_0$  é  $1 - \alpha$ .

a. Construa uma curva CO para o Exercício 9.1.7, usando valores da média verdadeira  $\mu$  de 178, 181, 184, 187, 190, 193, 196 e 199.

b. Converta a curva CO em um gráfico de **função de potência** do teste.

## Exercícios para a Seção 9.2

**9.2.1** Estabeleça as hipóteses nula e alternativa em cada caso.

a. Um teste de hipóteses será usado para potencialmente fornecer evidência de que a média da população é maior do que dez.

b. Um teste de hipóteses será usado para potencialmente fornecer evidência de que a média da população não é igual a sete.

c. Um teste de hipóteses será usado para potencialmente fornecer evidência de que a média da população é menor do que cinco.

**9.2.2** Uma hipótese será usada para testar que uma média da população é igual a cinco contra a alternativa de que a média da população não é igual a cinco, com variância conhecida  $\sigma$ . Qual é o valor crítico para a estatística de teste  $Z_0$ , considerando os seguintes níveis de significância?

a. 0,01    b. 0,05    c. 0,10

**9.2.3** A saída de um *software* é dada a seguir:

### Z para Uma Amostra:

Teste de  $\mu = 20$  versus  $> 20$

Desvio-padrão considerado = 0,75

Variável	N	Média	DP	EP da Média	Z	P
x	10	19,889	?	0,237	?	?

a. Complete as grandezas que faltam. Que conclusões você poderia tirar?

b. Esse é um teste unilateral ou bilateral?

c. Use a tabela da normal e os dados anteriores para construir um IC bilateral de 95 % para a média.

d. Qual seria o valor  $P$  se a hipótese alternativa fosse  $H_1: \mu \neq 20$ ?

**9.2.4** Para o teste de hipóteses  $H_0: \mu = 10$  contra  $H_1: \mu > 10$  e variância conhecida, calcule o valor  $P$  para cada uma das seguintes estatísticas de teste.

a.  $z_0 = 2,05$     b.  $z_0 = -1,84$     c.  $z_0 = 0,4$

**9.2.5** A saída de um *software* é dada a seguir.

### Z para Uma Amostra:

Teste de  $\mu = 99$  versus  $> 99$

Desvio-padrão considerado = 2,5

Variável	N	Média	DP	EP da Média	Z	P
x	12	100,039	2,365	?	1,44	0,075

a. Complete as grandezas que faltam. Que conclusões você poderia tirar?

b. Esse é um teste unilateral ou bilateral?

c. Se a hipótese tivesse sido  $H_0: \mu = 98$  versus  $H_0: \mu > 98$ , você rejeitaria a hipótese nula com nível de significância igual a 0,05? Você pode responder isso sem se referir à tabela da normal?

d. Use a tabela da normal e os dados anteriores para construir um limite inferior de 95 % para a média.

e. Qual seria o valor  $P$  se a hipótese alternativa fosse  $H_1: \mu \neq 99$ ?

**9.2.6** Um produtor fabrica eixos para um motor de automóvel. O desgaste (0,0001 polegada) dos eixos depois de 100.000 milhas é de interesse, visto que é provável ter um impacto nas reivindicações de garantia. Uma amostra aleatória de  $n = 15$  eixos é testada e  $\bar{x} = 2,78$ . Sabe-se que  $\sigma = 0,9$  e que o desgaste é normalmente distribuído.

a. Teste  $H_0: \mu = 3$  versus  $H_1: \mu \neq 3$ , usando  $\alpha = 0,05$ .

b. Qual é a potência desse teste se  $\mu = 3,25$ ?

c. Que tamanho da amostra seria requerido para detectar uma média verdadeira de 3,75, se quiséssemos que a potência fosse, no mínimo, 0,9?

**9.2.7** Sabe-se que a vida, em horas, de uma bateria é aproximadamente distribuída normalmente, com desvio-padrão  $\sigma = 1,25$  hora. Uma amostra aleatória de dez baterias tem uma vida média de  $\bar{x} = 40,5$  horas.

a. Há evidência que suporte a alegação de que a vida da bateria excede 40 horas? Use  $\alpha = 0,5$ .

b. Qual é o valor  $P$  para o teste do item (a)?

c. Qual será o erro  $\beta$  para o teste do item (a), se a média verdadeira for de 42 horas?

d. Que tamanho da amostra seria requerido para assegurar que  $\beta$  não excede 0,10, se a vida média verdadeira fosse de 44 horas?

e. Explique como você poderia responder à questão no item (a), calculando um limite apropriado de confiança para a vida da bateria.

9.2.8 *Supercavitação* é uma tecnologia de propulsão para veículos submarinos, que pode aumentar grandemente sua velocidade. Isso ocorre acima de aproximadamente 50 metros por segundo, quando a pressão cai o suficiente para permitir que a água se dissocie em vapor de água, formando uma bolha de gás atrás do veículo. Quando a bolha de gás envolve completamente o veículo, a supercavitação ocorre. Oito testes foram conduzidos em um protótipo de um veículo submarino em um tanque, com uma velocidade média observada de  $\bar{x} = 102,2$  metros por segundo. Considere que a velocidade seja normalmente distribuída, com desvio-padrão conhecido  $\sigma = 4$  metros por segundo.

- Teste as hipóteses  $H_0: \mu = 100$  versus  $H_1: \mu < 100$ , usando  $\alpha = 0,05$ .
- Qual é o valor  $P$  para o teste do item (a)?
- Calcule a potência do teste se a velocidade média verdadeira for tão baixa quanto 95 metros por segundo.
- Que tamanho da amostra seria requerido para detectar uma velocidade média verdadeira tão baixa quanto 95 metros por segundo, se quiséssemos que a potência do teste fosse no mínimo 0,85?
- Explique como a questão no item (a) poderia ser respondida, construindo um limite unilateral de confiança para a velocidade média.

9.2.9 Pesquisadores médicos desenvolveram um novo coração artificial, construído principalmente de titânio e plástico. O coração durará e operará quase indefinidamente, uma vez implantado no

corpo do paciente; porém, o conjunto de baterias necessita ser recarregado a cada quatro horas, aproximadamente. Uma amostra aleatória de 50 conjuntos de baterias é selecionada e submetida a um teste de vida. A vida média dessas baterias é de 4,05 horas. Considere que a vida das baterias seja normalmente distribuída, com desvio-padrão de  $\sigma = 0,2$  hora.

- Há evidência para suportar a alegação de que a vida média da bateria excede quatro horas? Use  $\alpha = 0,05$ .
- Qual é o valor  $P$  para o teste do item (a)?
- Calcule a potência do teste se a vida média verdadeira da bateria for de 4,5 horas.
- Que tamanho da amostra seria requerido para detectar a vida média verdadeira da bateria de 4,5 horas, se quiséssemos que a potência do teste fosse, no mínimo, 0,9?
- Explique como a questão no item (a) poderia ser respondida, construindo um limite unilateral de confiança para a vida média.

9.2.10 A estirpe bacteriana *Acinetobacter* tem sido testada por causa das propriedades de adesão. Uma amostra de cinco medidas apresentou as leituras de 2,69; 5,76; 2,67; 1,62 e 4,12 dinas-cm<sup>2</sup>. Considere que o desvio-padrão seja conhecido, sendo igual a 0,66 dina-cm<sup>2</sup>, e que os cientistas estejam interessados em alta adesão (no mínimo 2,5 dinas-cm<sup>2</sup>).

- A hipótese alternativa deve ser unilateral ou bilateral?
- Teste a hipótese em que a adesão média é 2,5 dinas-cm<sup>2</sup>.
- Qual é o valor  $P$  da estatística de teste?

## Exercícios para a Seção 9.3

9.3.1 Uma hipótese será usada para testar se a média de uma população é igual a dez contra a alternativa de que a média da população é maior do que dez, com variância desconhecida. Qual é o valor crítico para a estatística de teste  $T_0$  para os seguintes níveis de significância?

- $\alpha = 0,01$  e  $n = 20$
- $\alpha = 0,05$  e  $n = 12$
- $\alpha = 0,10$  e  $n = 15$

9.3.2 Considere a saída computacional a seguir.

### T para Uma Amostra:

Teste de  $\mu = 91$  versus  $\mu > 91$

Variável	N	Média	DP	Limite		
				EP da Média	Inferior de 95 %	T P
x	20	92,379	0,717	?	?	?

- Preencha os valores que faltam. Você pode calcular limites para o valor  $P$ . Que conclusões você pode tirar?
- Esse é um teste unilateral ou bilateral?
- Se a hipótese tivesse sido  $H_0: \mu = 90$  versus  $H_1: \mu > 90$ , suas conclusões mudariam?

9.3.3 Para o teste de hipóteses  $H_0: \mu = 7$  contra  $H_1: \mu \neq 7$ , com variância desconhecida e  $n = 20$ , aproxime o valor  $P$  para cada uma das seguintes estatísticas de teste.

- $t_0 = 2,05$
- $t_0 = -1,84$
- $t_0 = 0,4$

9.3.4 Considere a saída computacional a seguir.

### T para Uma Amostra:

Teste de  $\mu = 34$  versus  $\mu > 34$

Variável	N	Média	DP	IC de 95 %		T P
				EP da Média	IC de 95 %	
x	16	35,274	1,783	?	(34,324, 36,224)	? 0,012

- Quantos graus de liberdade existem para a estatística de teste  $t$ ?
- Preencha as grandezas que faltam.
- Em que nível de significância a hipótese nula pode ser rejeitada?
- Se a hipótese tivesse sido  $H_0: \mu = 34$  versus  $H_1: \mu > 34$ , o valor  $P$  seria maior ou menor?
- Se a hipótese tivesse sido  $H_0: \mu = 34,5$  versus  $H_1: \mu \neq 34,5$ , você teria rejeitado a hipótese nula no nível de 0,05?

9.3.5 Para o teste de hipóteses  $H_0: \mu = 5$  contra  $H_1: \mu < 5$ , com variância desconhecida e  $n = 12$ , aproxime o valor  $P$  para cada uma das seguintes estatísticas de teste.

- $t_0 = 2,05$
- $t_0 = -1,84$
- $t_0 = 0,4$

9.3.6 Um artigo publicado no *ASCE Journal of Energy Engineering* (1999, Vol. 125, pp. 59-75) descreveu um estudo das propriedades da inércia térmica de concreto aerado em autoclave, usado como material de construção. Cinco amostras do material foram testadas em uma estrutura, e a temperatura (°C) média no interior reportada foi: 23,01; 22,22; 22,04; 22,62 e 22,59.

- Teste as hipóteses  $H_0: \mu = 22,5$  versus  $H_1: \mu \neq 22,5$ , usando  $\alpha = 0,05$ . Encontre o valor  $P$ .
- Verifique a suposição de que a temperatura no interior é normalmente distribuída.
- Calcule a potência do teste se a temperatura média verdadeira no interior é tão alta quanto 22,75.
- Que tamanho da amostra seria requerido para detectar uma temperatura média verdadeira, no interior, tão alta quanto 22,75, se quiséssemos que a potência do teste fosse no mínimo 0,9?
- Explique como a questão no item (a) poderia ser respondida, construindo um intervalo bilateral para a temperatura média no interior.

9.3.7 Um artigo publicado em *Growth: A Journal Devoted to Problems of Normal and Abnormal Growth* ["Comparison of Measured and Estimated Fat-Free Weight, Fat, Potassium and Nitrogen of Growing Guinea Pigs" (1982, Vol. 46(4), pp. 306-321)] mediu o peso (em gramas) do corpo das cobaias em seu nascimento.

421,0	452,6	456,1	494,6	373,8
90,5	110,7	96,4	81,7	102,4
241,0	296,0	317,0	290,9	256,5
447,8	687,6	705,7	879,0	88,8
296,0	273,0	268,0	227,5	279,3
258,5	296,0			

- Teste a hipótese de que o peso médio do corpo é de 300 gramas. Use  $\alpha = 0,05$ .
- Qual é o menor nível de significância ao qual você estará sujeito para rejeitar a hipótese nula?
- Explique como você responderia à pergunta do item (a) com um intervalo bilateral de confiança para o peso médio do corpo.

9.3.8 Semear nuvens tem sido estudado durante muitas décadas como um procedimento de mudança do tempo (para um estudo interessante desse assunto, veja o artigo em *Technometrics*, "A Bayesian Analysis of a Multiplicative Treatment Effect in Weather Modification", 1975, Vol. 17, pp. 161-166). A precipitação pluviométrica, em acre-pé,<sup>1</sup> proveniente de 20 nuvens que foram selecionadas aleatoriamente e semeadas com nitrato de prata, segue: 18,0; 30,7; 19,8; 27,1; 22,3; 18,8; 31,8; 23,4; 21,2; 27,9; 31,9; 27,1; 25,0; 24,7; 26,9; 21,8; 29,2; 34,8; 26,7 e 31,6.

- Você pode sustentar a afirmação de que a precipitação média das nuvens semeadas excede 25 acres-pés? Use  $\alpha = 0,01$ . Encontre o valor  $P$ .
- Verifique se a precipitação é normalmente distribuída.
- Calcule a potência do teste, se a precipitação média verdadeira for de 27 acres-pés.
- Que tamanho da amostra seria requerido para detectar uma precipitação média verdadeira de 27,5 acres-pés, se quiséssemos que a potência do teste fosse no mínimo 0,9?
- Explique como a questão no item (a) poderia ser respondida, construindo um limite unilateral para a precipitação média.

9.3.9 Um artigo de 1992 publicado no *Journal of the American Medical Association* ("A Critical Appraisal of 98,6 Degrees F, the Upper Limit of the Normal Body Temperature, and Other Legacies of Carl Reinhold August Wunderlich") reportou temperatura do corpo, gênero e taxa do coração para um número de pessoas. As temperaturas do corpo para 25 mulheres foram: 97,8; 97,2; 97,4; 97,6; 97,8; 97,9; 98,0; 98,0; 98,0; 98,1; 98,2; 98,3; 98,3; 98,4; 98,4; 98,4; 98,5; 98,6; 98,6; 98,7; 98,8; 98,8; 98,9; 98,9 e 99,0.

- Teste a hipótese  $H_0: \mu = 98,6$  versus  $H_1: \mu \neq 98,6$ , usando  $\alpha = 0,05$ . Encontre o valor  $P$ .

<sup>1</sup> Acre-pé é uma unidade de volume, comumente usada nos Estados Unidos, para grandes volumes de água. É definida como o volume de água necessário para cobrir um acre (66 por 660 pés) de área superficial e um pé de profundidade. Esse volume corresponde a 1233,5 metros cúbicos. (N. de T.)

- Verifique a suposição de que a temperatura do corpo feminino é normalmente distribuída.
- Calcule a potência do teste se a temperatura média verdadeira do corpo feminino é tão alta quanto 98,0.
- Que tamanho da amostra seria requerido para detectar uma temperatura média verdadeira do corpo feminino tão baixa quanto 98,2, se quiséssemos que a potência do teste fosse no mínimo 0,9?
- Explique como a questão no item (a) poderia ser respondida, construindo um intervalo bilateral para a temperatura média do corpo feminino.

9.3.10 Reconsidere os dados de *Medicine and Science in Sports and Exercise*, descritos no Exercício 8.2.6. O tamanho da amostra foi igual a sete e a média e o desvio-padrão da amostra foram de 315 watts e 16 watts, respectivamente.

- Há evidência de que a resistência da perna excede 300 watts, com um nível de significância de 0,05? Encontre o valor  $P$ .
- Calcule a potência do teste se a resistência verdadeira for de 305 watts.
- Que tamanho da amostra seria requerido para detectar uma média verdadeira de 305 watts se quiséssemos que a potência do teste fosse no mínimo 0,9?
- Explique como a questão no item (a) poderia ser respondida com um intervalo de confiança.

9.3.11 O Exercício 6.2.8 forneceu dados sobre as alturas de estudantes femininas de engenharia em ASU.

- Você pode sustentar a afirmação de que a altura média de estudantes femininas de engenharia em ASU é, no mínimo, de 65 polegadas? Use  $\alpha = 0,05$ . Encontre o valor  $P$ .
- Verifique a suposição de normalidade.
- Calcule a potência do teste se a altura média verdadeira for de 68 polegadas.
- Que tamanho da amostra seria requerido para detectar uma altura média verdadeira de 66 polegadas, se quiséssemos que a potência do teste fosse no mínimo 0,8?

9.3.12 O Exercício 6.2.10 apresentou dados da concentração de sólidos suspensos na água de um lago.

- Teste as hipóteses  $H_0: \mu = 55$  versus  $H_1: \mu \neq 55$ . Use  $\alpha = 0,05$ . Encontre o valor  $P$ .
- Verifique a suposição de normalidade.
- Calcule a potência do teste se a concentração média verdadeira for tão baixa quanto 50.
- Que tamanho da amostra seria requerido para detectar uma concentração média verdadeira tão baixa quanto 50, se quiséssemos que a potência do teste fosse no mínimo 0,9?

9.3.13 Em pouco mais de um mês, de 5 de junho de 1879 a 2 de julho de 1879, Albert Michelson mediu a velocidade da luz no ar 100 vezes (Stigler, *Annals of Statistics*, 1977). Hoje, sabemos que o valor verdadeiro é 299.734,5 km/s. Os dados de Michelson têm uma média de 299.852,4 km/s, com um desvio-padrão de 79,01.

- Encontre um intervalo de confiança bilateral de 95 % para a média verdadeira (o valor verdadeiro da velocidade da luz).
- O que o intervalo de confiança diz acerca da acurácia das medidas de Michelson?

## Exercícios para a Seção 9.4

9.4.1 Considere o teste de  $H_0: \sigma^2 = 7$  contra  $H_1: \sigma^2 \neq 7$ . Quais são os valores críticos para a estatística de teste  $\chi_0^2$  para os seguintes níveis de significância e tamanhos de amostra?

- a.  $\alpha = 0,01$  e  $n = 20$       b.  $\alpha = 0,05$  e  $n = 12$   
 c.  $\alpha = 0,10$  e  $n = 15$

9.4.2 Considere o teste de  $H_0: \sigma^2 = 5$  contra  $H_1: \sigma^2 < 5$ . Aproxime o valor  $P$  para cada uma das seguintes estatísticas de teste.

- a.  $x_0^2 = 25,2$  e  $n = 20$       b.  $x_0^2 = 15,2$  e  $n = 12$   
 c.  $x_0^2 = 4,2$  e  $n = 15$

9.4.3 Os dados de *Medicine and Science in Sports and Exercise*, descritos no Exercício 8.3.4, consideraram o desempenho de um jogador de hóquei no gelo depois de um treino com eletroestimulação. Em resumo, havia 17 jogadores e o desvio-padrão do desempenho da amostra foi 0,09 segundo.

- a. Existe forte evidência para concluir que o desvio-padrão do tempo do desempenho excede o valor histórico de 0,75 segundo? Use  $\alpha = 0,05$ . Encontre o valor  $P$  para esse teste.

- b. Discuta como o item (a) poderia ser respondido, construindo um intervalo unilateral de confiança de 95 % para  $\sigma$ .

9.4.4 Reconsidere a percentagem de titânio em uma liga usada na fundição de aeronaves do Exercício 8.3.1. Lembre-se de que  $s = 0,37$  e  $n = 51$ .

- a. Teste a hipótese  $H_0: \sigma = 0,25$  versus  $H_1: \sigma \neq 0,25$ , usando  $\alpha = 0,05$ . Estabeleça quaisquer suposições necessárias acerca da distribuição em foco dos dados. Encontre o valor  $P$ .  
 b. Explique como você poderia responder à pergunta do item (a), construindo um intervalo bilateral de confiança de 95 % para  $\sigma$ .

9.4.5 Dados para a vida de pneus foram descritos no Exercício 8.2.3. O desvio-padrão da amostra foi de 3645,94 quilômetros e  $n = 16$ .

- a. Você pode concluir, usando  $\alpha = 0,05$ , que o desvio-padrão da vida de um pneu é menor do que 4000 quilômetros? Estabeleça quaisquer suposições necessárias acerca da distribuição em foco dos dados. Encontre o valor  $P$  para esse teste.  
 b. Explique como você responderia à pergunta do item (a) construindo um intervalo unilateral de confiança de 95 % para  $\sigma$ .

## Exercícios para a Seção 9.5

9.5.1 Considere a saída computacional a seguir.

**Teste e IC para Uma Proporção:**

Teste de  $p = 0,4$  vs  $p \neq 0,4$

X	N	p para amostra	IC de 95 %	Valor Z	Valor P
98	275	?	(0,299759, 0,412968)	?	?

Use a aproximação normal.

- a. Esse é um teste unilateral ou bilateral?  
 b. Complete os termos que faltam.  
 c. A aproximação normal foi usada no problema. Ela foi apropriada?

9.5.2 Suponha que 1000 consumidores sejam entrevistados e 850 declarem estar satisfeitos ou muito satisfeitos com os produtos e serviços de certa corporação.

- a. Teste a hipótese  $H_0: p = 0,9$  contra  $H_1: p \neq 0,9$ , com  $\alpha = 0,05$ . Encontre o valor  $P$ .  
 b. Explique como a questão do item (a) poderia ser respondida pela construção de um intervalo bilateral de confiança de 95 % para  $p$ .

9.5.3 Um artigo no *British Medical Journal* ["Comparison of Treatment of Renal Calculi by Operative Surgery, Percutaneous Nephrolithotomy, and Extracorporeal Shock Wave Lithotripsy" (1986, Vol. 292, pp. 879-882)] encontrou que a remoção de pedras nos rins pela nefrolitotomia percutânea (NP) tem uma taxa de sucesso igual a 289 de 350 pacientes. O método tradicional foi 78 % efetivo.

- a. Há evidência de que a taxa de sucesso para NP seja maior do que a taxa histórica de sucesso? Encontre o valor  $P$ .  
 b. Explique como a questão do item (a) poderia ser respondida com um intervalo de confiança.

9.5.4 Um artigo em *Fortune* (21 de setembro de 1992) afirmou que aproximadamente metade de todos os engenheiros continua seus estudos acadêmicos além do grau de bacharelado, recebendo no final o grau de mestre ou doutor. Dados de um artigo em *Engineering Horizons* (primavera de 1990) indicaram que 117 de 484 novos engenheiros graduados estavam planejando fazer pós-graduação.

- a. Os dados da *Engineering Horizons* são consistentes com a afirmação reportada pela *Fortune*? Use  $\alpha = 0,05$  para alcançar as suas conclusões. Encontre o valor  $P$  para o teste.  
 b. Discuta como a questão do item (a) poderia ser respondida pela construção de um intervalo bilateral de confiança para  $p$ .

9.5.5 Uma amostra aleatória foi composta de 500 eleitores registrados em Fênix. A esses eleitores foi perguntado se seriam favoráveis ao uso de combustíveis oxigenados para reduzir a poluição do ar ao longo do ano. Se mais de 315 eleitores responderem positivamente, concluiremos que no mínimo 60 % dos eleitores são favoráveis ao uso desses combustíveis.

- a. Encontre a probabilidade do erro tipo I, se exatamente 60 % dos eleitores forem favoráveis ao uso desses combustíveis.  
 b. Qual será a probabilidade do erro tipo II,  $\beta$ , se 75 % dos eleitores forem favoráveis a essa ação?

9.5.6 Um fabricante de computadores envia *laptops* com as baterias completamente carregadas, de modo que os consumidores possam começar a usar suas compras tão logo as tirem da caixa. No seu último modelo, 85 % dos consumidores receberam as baterias completamente carregadas. Para simular chegadas, a companhia enviou 100 novos modelos de *laptops* para vários locais ao redor do país. Dos 100 *laptops* enviados, 96 chegaram indicando 100 % de carga. Os dados fornecem evidência de que essa taxa do modelo é no mínimo tão alta quanto a do modelo anterior? Teste a hipótese com  $\alpha = 0,05$ .

## Exercícios para a Seção 9.7

9.7.1 Considere a seguinte tabela de frequências de observações para a variável aleatória  $X$ :

Valores	0	1	2	3	4
Frequência	4	21	10	13	2

a. Com base nessas 50 observações, a distribuição binomial, com  $n = 6$  e  $p = 0,25$ , é um modelo apropriado? Faça um procedimento de adequação de ajuste com  $\alpha = 0,05$ .

b. Calcule o valor  $P$  para esse teste.

9.7.2 Considere a seguinte tabela de frequência de observações para a variável aleatória  $X$ .

Valores	0	1	2	3	4
Frequência observada	24	30	31	11	4

a. Com base nessas 100 observações, a distribuição de Poisson, com uma média de 1,2, é um modelo apropriado? Faça um procedimento de adequação de ajuste com  $\alpha = 0,05$ .

b. Calcule o valor  $P$  para esse teste.

9.7.3 O número de carros passando pela interseção da Avenida Cruzeiro com a Avenida Central foi tabelado por um grupo de estudantes de engenharia civil. Eles obtiveram os dados da seguinte tabela:

Veículos por Minuto	Frequência Observada	Veículos por Minuto	Frequência Observada
40	14	53	102
41	24	54	96
42	57	55	90
43	111	56	81
44	194	57	73

Veículos por Minuto	Frequência Observada	Veículos por Minuto	Frequência Observada
45	256	58	64
46	296	59	61
47	378	60	59
48	250	61	50
49	185	62	42
50	171	63	29
51	150	64	18
52	110	65	15

a. A suposição de distribuição de Poisson parece apropriada como um modelo de probabilidade para esse processo? Use  $\alpha = 0,05$ .

b. Calcule o valor  $P$  para esse teste.

9.7.4 O número de chamadas que chegam a uma mesa telefônica de meio-dia a uma hora da tarde, durante os dias de semana, de segunda a sexta-feira, é monitorado por seis semanas (ou seja, 30 dias). Seja  $X$  o número de chamadas durante aquele período de uma hora. A frequência relativa de chamadas foi registrada e reportada como

Valor	5	6	8	9	10
Frequência relativa	0,067	0,067	0,100	0,133	0,200
Valor	11	12	13	14	15
Frequência relativa	0,133	0,133	0,067	0,033	0,067

a. A suposição de distribuição de Poisson parece apropriada como um modelo de probabilidade para esses dados? Use  $\alpha = 0,05$ .

b. Calcule o valor  $P$  para esse teste.

## Exercícios para a Seção 9.8

9.8.1 A Floresta Hopkins é uma reserva florestal de 2600 acres, localizada na interseção entre três estados americanos: Nova York, Vermont e Massachusetts. Pesquisadores monitoram as reservas florestais para estudar as variações ecológicas de longo prazo. Eles conduziram pesquisas de árvores, de arbustos e de ervas existentes em vários locais da floresta, durante aproximadamente 100 anos. A seguir, são apresentados alguns dados da pesquisa de três espécies de carvalho na mesma localização ao longo de três períodos de tempo muito diferentes.

	Espécies (Acre)			Total
	Pensylvanicum	Rubrum	Saccharum	
Ano 1936	12	27	94	133
1972	22	40	52	114
2011	97	25	18	140
Total	131	92	164	387

A distribuição de espécies parece ser independente do ano? Teste a hipótese com  $\alpha = 0,05$ . Encontre o valor  $P$  da estatística de teste.

9.8.2 A taxa de sobrevivência para passageiros do *Titanic* realmente dependeu do tipo de bilhete que eles tinham? A tabela a seguir

apresenta os dados para as 2201 pessoas a bordo listadas pelo tipo de bilhete que tinham e se sobreviveram ou não. A sobrevivência parece ser independente da classe do bilhete? (Teste a hipótese com  $\alpha = 0,05$ .) Qual é o valor  $P$  da estatística de teste?

	Tripulação	Primeira	Segunda	Terceira	Total
Vivos	212	202	118	178	710
Mortos	673	123	167	528	1491
Total	885	325	285	706	2201

9.8.3 Uma companhia opera quatro máquinas em três turnos todo dia. Dos registros de produção, são coletados os seguintes dados do número de interrupções:

Turno	Máquinas			
	A	B	C	D
1	41	20	12	16
2	31	11	9	14
3	15	17	16	10

Teste a hipótese (usando  $\alpha = 0,05$ ) de que as interrupções são independentes do turno. Encontre o valor  $P$  para esse teste.

9.8.4 Pacientes em um hospital são classificados como cirúrgicos ou médicos. Mantém-se um registro do número de vezes que os pacientes necessitam de serviços da enfermaria durante a noite e se esses pacientes estão ou não sob cuidados médicos. Os dados são apresentados aqui:

Cuidados Médicos	Categoria do Paciente	
	Cirúrgico	Médico
Sim	46	52
Não	36	43

Teste a hipótese (usando  $\alpha = 0,01$ ) de que as chamadas pelos pacientes médico-cirúrgicos são independentes de se os pacientes estejam recebendo cuidados médicos. Encontre o valor  $P$  para esse teste.

9.8.5 As notas em um curso de estatística e em um curso de pesquisa operacional, tomadas simultaneamente, são dadas a seguir para um grupo de estudantes.

Nota de Estatística	Nota de Pesquisa Operacional			
	A	B	C	Outro
A	25	6	17	13
B	17	16	15	6
C	18	4	18	10
Outro	10	8	11	20

As notas em estatística e em pesquisa operacional estão relacionadas? Use  $\alpha = 0,01$  para concluir algo. Qual é o valor  $P$  para esse teste?

9.8.6 Um experimento com cartuchos de artilharia resulta nos seguintes dados para as características das deflexões e amplitudes laterais. Você concluiria que deflexão e amplitude são independentes? Use  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?

Faixa (jardas)	Deflexão Lateral		
	Esquerda	Normal	Direita
0–1999	6	14	8
2000–5999	9	11	4
6000–11.999	8	17	6

9.8.7 Um estudo está sendo feito sobre as falhas de um componente eletrônico. Há quatro tipos possíveis de falhas e duas posições de montagem do componente. Os seguintes dados foram obtidos:

Posição de Montagem	Tipo de Falha			
	A	B	C	D
1	22	46	18	9
2	4	17	6	12

Você concluiria que o tipo de falha é independente da posição de montagem? Use  $\alpha = 0,01$ . Qual é o valor  $P$  para esse teste?

9.8.8 Perguntamos a estudantes, provenientes de uma amostra aleatória, suas opiniões a respeito de mudanças propostas no núcleo do currículo de seu curso. Os resultados são dados a seguir.

Classe	Opinião	
	A Favor	Contra
Primeiro Ano	120	80
Segundo Ano	70	130
Terceiro Ano	60	70
Quarto Ano	40	60

Teste a hipótese de que a opinião sobre a mudança é independente da classe. Use  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?

9.8.9 Um artigo no *British Medical Journal* [“Comparison of Treatment of Renal Calculi by Operative Surgery, Percutaneous Nephrolithotomy, and Extracorporeal Shock Wave Lithotripsy” (1986, Vol. 292, pp. 879-882)] encontrou que a remoção de pedras nos rins pela nefrolitotomia percutânea (NP) tem uma taxa de sucesso de 289 de 350 (83 %) pacientes. No entanto, quando o diâmetro das pedras foi considerado, os resultados pareceram diferentes. Para pedras menores que 2 cm, 87 % (234/270) dos casos tiveram sucesso. Para pedras maiores ou iguais a 2 cm, uma taxa de sucesso de 69 % (55/80) foi observada para NP.

- Os sucessos e o tamanho das pedras são independentes? Use  $\alpha = 0,05$ .
- Encontre o valor  $P$  para esse teste.

## Exercícios para a Seção 9.9

9.9.1 O teor de titânio em uma liga de um avião é um determinante importante da resistência. Uma amostra de 20 corpos de prova revela o seguinte teor (em percentagem) de titânio:

8,32; 8,05; 8,93; 8,65; 8,25; 8,46; 8,52; 8,35; 8,36; 8,41; 8,42; 8,30; 8,71; 8,75; 8,60; 8,83; 8,50; 8,38; 8,29; 8,46.

A mediana do teor de titânio deve ser 8,5 %.

- Use o teste dos sinais com  $\alpha = 0,05$  para investigar essa hipótese. Encontre o valor  $P$  para esse teste.
- Use a aproximação normal para o teste dos sinais para testar  $H_0: \mu = 8,5$  versus  $H_1: \mu \neq 8,5$ , com  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?

9.9.2 Dez amostras foram retiradas de um banho de recobrimento usado em um processo de fabricação de eletrônicos, e o pH do banho foi determinado. Os valores de pH da amostra são 7,91; 7,85; 6,82;

8,01; 7,46; 6,95; 7,05; 7,35; 7,25 e 7,42. A engenharia de fabricação acredita que o pH tenha uma mediana igual a 7,0.

- Os dados da amostra indicam que essa afirmação está correta? Use o teste dos sinais, com  $\alpha = 0,05$ , para investigar essa hipótese. Encontre o valor  $P$  para esse teste.
- Use a aproximação normal para o teste dos sinais para testar  $H_0: \mu = 7,0$  versus  $H_1: \mu \neq 7,0$ . Qual é o valor  $P$  para esse teste?

9.9.3 O diâmetro de um mancal de esferas foi medido por um inspetor usando um novo tipo de paquímetro. Os resultados (em mm) foram: 0,265; 0,263; 0,266; 0,267; 0,267; 0,265; 0,267; 0,267; 0,265; 0,268; 0,268 e 0,263.

- Use o teste de Wilcoxon do posto sinalizado para avaliar a afirmação de que o diâmetro médio da esfera é 0,265 mm. Use  $\alpha = 0,05$ .
- Use a aproximação normal para o teste dos sinais. Com  $\alpha = 0,05$ , que conclusões você pode tirar?

**9.9.4** O nível (em ppm) de impureza é rotineiramente medido em um produto químico intermediário. Os seguintes dados foram observados em um teste recente:

2,4; 2,5; 1,7; 1,6; 1,9; 2,6; 1,3; 1,9; 2,0; 2,5; 2,6; 2,3; 2,0; 1,8; 1,3; 1,7; 2,0; 1,9; 2,3; 1,9; 2,4; 1,6.

Você pode afirmar que a mediana do nível de impureza seja menor do que 2,5 ppm?

- Estabeleça e teste a hipótese apropriada, usando o teste dos sinais com  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?
- Use a aproximação normal para o teste dos sinais para testar  $H_0: \mu = 2,5$  versus  $H_1: \mu < 2,5$ , com  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?

**9.9.5** Zarcão pode ser usado em painéis de alumínio. O tempo de secagem do zarcão é uma consideração importante no processo de fabricação. Foram selecionados 20 painéis, e os tempos de secagem foram: 1,6; 1,3; 1,5; 1,6; 1,7; 1,9; 1,8; 1,6; 1,4; 1,8; 1,9; 1,8; 1,7; 1,5; 1,6; 1,4; 1,3; 1,6; 1,5 e 1,8. Existe alguma evidência de que o tempo médio de secagem do zarcão seja maior do que 1,5 hora?

**9.9.6** Um novo tipo de ponteira pode ser usado em uma máquina de teste de dureza Rockwell. Selecionam-se oito corpos de prova provenientes de lingotes de teste de uma liga de níquel. Cada corpo de prova é testado usando-se a nova ponteira. As leituras de dureza na escala C Rockwell são 63, 65, 58, 60, 55, 57, 53 e 59. Os resultados suportam a hipótese de que a dureza média é maior do que 60 em um nível de 0,05?

## Exercícios para a Seção 9.10

**9.10.1** Um fabricante de produtos químicos tem de identificar um novo fornecedor de matéria-prima, que é um componente essencial para um produto particular. O fornecedor anterior foi capaz de entregar o material com uma massa molar média de 3500. O novo fornecedor tem de mostrar equivalência com esse valor de massa molar. Se o novo fornecedor puder entregar o material cuja massa molar esteja dentro de 50 unidades desse valor, ele será considerado equivalente. Uma amostra aleatória de dez lotes de produto está disponível, sendo a média e o desvio-padrão da massa molar iguais a 3550 e 25, respectivamente.

- Estabeleça as hipóteses apropriadas que têm de ser testadas para demonstrar a equivalência.
- Quais são as suas conclusões usando  $\alpha = 0,05$ ?

**9.10.2** No desenvolvimento de uma droga genérica, um fabricante de biofármacos necessita mostrar equivalência com o produto corrente. A variável de interesse é a taxa de absorção do produto. O produto atual tem uma taxa de absorção de 18 mg/h. Se o novo produto genérico tiver uma taxa de absorção dentro de 0,50 mg/h desse valor, ele será considerado equivalente. Uma amostra aleatória de 20 unidades do produto está disponível e a média e o desvio-padrão da amostra para a taxa de absorção são iguais a 18,22 mg/h e 0,92 mg/h, respectivamente.

- Estabeleça as hipóteses apropriadas que têm de ser testadas para demonstrar a equivalência.
- Quais são as suas conclusões usando  $\alpha = 0,05$ ?

**9.10.3** A resistência média de adesão de um cimento tem de ser no mínimo igual a 10.000 psi. O processo pelo qual esse material é fabricado tem de mostrar equivalência a esse padrão. Se o processo puder fabricar cimentos com uma resistência média de adesão de no mínimo 9750 psi, o material será considerado equivalente ao padrão. Uma amostra aleatória de seis observações encontra-se disponível, e a média e o desvio-padrão da resistência de adesão são iguais a 9360 psi e 42,6 psi, respectivamente.

- Estabeleça as hipóteses apropriadas que têm de ser testadas para demonstrar a equivalência.
- Quais são as suas conclusões usando  $\alpha = 0,05$ ?

**9.10.4** A resistência média de ruptura de um isolante de cerâmica tem de ser no mínimo igual a 10 psi. O processo pelo qual esse isolante é fabricado tem de mostrar equivalência a esse padrão. Se o processo puder fabricar isolantes com uma resistência média de ruptura de no mínimo 9,5 psi, o material será considerado equivalente ao padrão. Uma amostra aleatória de 50 isolantes está disponível, e a média e o desvio-padrão da resistência de ruptura são iguais a 9,31 psi e 0,22 psi, respectivamente.

- Estabeleça as hipóteses apropriadas que têm de ser testadas para demonstrar a equivalência.
- Quais são as suas conclusões usando  $\alpha = 0,05$ ?

## Exercícios para a Seção 9.11

**9.11.1** Suponha que oito conjuntos de hipóteses sobre uma proporção de população da forma

$$H_0: p = 0,3 \quad H_1: p > 0,3$$

tenham sido testados e que os valores  $P$  para esses testes sejam 0,15; 0,83; 0,103; 0,024; 0,03; 0,07; 0,09 e 0,13. Use o procedimento de Fisher para combinar todos esses valores  $P$ . Existe evidência suficiente para concluir que a proporção de população exceda 0,30?

**9.11.2** Suponha que dez conjuntos de hipóteses da forma

$$H_0: \mu = \mu_0 \quad H_1: \mu \neq \mu_0$$

tenham sido testados e que os valores  $P$  para esses testes sejam 0,12; 0,08; 0,93; 0,02; 0,01; 0,05; 0,88; 0,15; 0,13 e 0,06. Use o procedimento de Fisher para combinar todos esses valores  $P$ . Que conclusões você pode tirar dessas hipóteses?

**9.11.3** A massa média de um pacote de peixe congelado tem de ser igual a 22 onças. Cinco amostras independentes foram selecionadas e as hipóteses estatísticas a respeito da massa média foram testadas. Os valores  $P$  que resultaram desses testes foram 0,065; 0,0924; 0,073; 0,025 e 0,021. Existe evidência suficiente para concluir que a massa média do pacote não seja igual a 22 onças?

9.11.4 O desvio-padrão do volume de enchimento de um recipiente de um produto farmacêutico tem de ser menor do que 0,2 onça, de modo a assegurar que o recipiente seja cheio, de forma acurada. Seis amostras independentes foram selecionadas e as hipóteses estatísticas

a respeito do desvio-padrão foram testadas. Os valores  $P$  que resultaram foram 0,15; 0,091; 0,075; 0,02; 0,04 e 0,06. Existe evidência suficiente para concluir que o desvio-padrão do volume de enchimento seja menor do que 0,2 onça?

## Exercícios Suplementares para o Capítulo 9

9.S5 Considere a saída computacional a seguir.

### T para Uma Amostra:

Teste de  $\mu = 100$  versus  $\mu \neq 100$

Variável	N	Média	DP	EP da Média	IC de 95 %	T	P
X	16	98,33	4,61	?	(?, ?)	?	?

- Quantos graus de liberdade existem para a estatística de teste  $t$ ?
- Preencha as grandezas que faltam. Você pode usar limites para o valor  $P$ .
- Quais são suas conclusões se  $\alpha = 0,05$ ?
- Quais serão suas conclusões se a hipótese for  $H_0: \mu = 100$  versus  $H_a: \mu > 100$ ?

9.S6 Considere a saída computacional a seguir.

### Z para Uma Amostra:

Teste de  $\mu = 26$  versus  $\mu > 26$

Desvio-padrão considerado = 1,5

Variável	N	Média	DP	EP da Média	Z	P
X	?	26,541	2,032	0,401	?	?

- Preencha as informações que faltam.
- Esse é um teste unilateral ou bilateral?
- Quais são suas conclusões se  $\alpha = 0,05$ ?
- Encontre um IC bilateral de 95 % para a média.

9.S7 Considere a saída computacional a seguir.

### T para Uma Amostra:

Teste de  $\mu = 85$  versus  $\mu < 85$

Variável	N	Média	DP	EP da Média	T	P
X	25	84,331	?	0,631	?	?

- Quantos graus de liberdade existem para a estatística de teste  $t$ ?
- Preencha as grandezas que faltam. Você pode usar limites para o valor  $P$ .
- Quais são suas conclusões se  $\alpha = 0,05$ ?
- Encontre um limite superior de 95 % para a média.
- Quais serão suas conclusões se a hipótese for  $H_0: \mu = 100$  versus  $H_a: \mu > 100$ ?

9.S8 Um artigo em *Fire Technology* ["An Experimental Examination of Dead Air Space for Smoke Alarms" (2009, Vol. 45, pp. 97-115)] estudou o desempenho de detectores de fumaça instalados, no mínimo, a 100 mm da parede adjacente, se instalados em um teto plano, e no mínimo a 100 mm e no máximo a 300 mm da superfície adjacente do teto, se instalados nas paredes. A finalidade dessa regra é evitar a instalação de alarmes de fumaça no "espaço morto de ar", em que se considera ser difícil a fumaça alcançar. Um número de experimentos interessantes foi descrito no trabalho. Resultados para o tempo (em

segundos) para sinalizar tal experimento com combustível de gravetos de pinho em um quarto aberto, usando alarmes fotoelétricos de fumaça, são: 220, 225, 297, 315, 282 e 313.

- Existe evidência suficiente para suportar a hipótese de que o tempo médio para sinalizar seja menor do que 300 segundos?
- Existe alguma preocupação prática acerca da suposição de uma distribuição normal como um modelo para os dados tempo-sinalização?
- Encontre um IC bilateral de 95 % para o tempo médio para sinalizar.

9.S9 Considere a descrição do Exercício 9.S10. Depois de coletar uma amostra, estamos interessados em testar  $H_0: p = 0,10$  versus  $H_1: p \neq 0,10$ , com  $\alpha = 0,05$ . Para cada uma das situações seguintes, calcule o valor  $p$  para esse teste:

- $n = 50$  e  $\hat{p} = 0,095$
- $n = 100$  e  $\hat{p} = 0,095$
- $n = 500$  e  $\hat{p} = 0,095$
- $n = 1000$  e  $\hat{p} = 0,095$

e. Comente o efeito do tamanho da amostra sobre o valor  $P$  verdadeiro observado do teste.

9.S10 Um fabricante de semicondutores retira uma amostra aleatória, de tamanho  $n$ , de *chips*, testando-os e classificando-os como defeituosos ou não defeituosos. Seja  $X_i = 0$ , se o *chip* for não defeituoso, e  $X_i = 1$ , se o *chip* for defeituoso. A fração defeituosa da amostra é

$$\hat{p} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Quais são a distribuição amostral e as estimativas da média amostral e da variância amostral de  $\hat{p}$ , quando

- O tamanho da amostra for  $n = 50$ ?
- O tamanho da amostra for  $n = 80$ ?
- O tamanho da amostra for  $n = 100$ ?
- Compare suas respostas para os itens (a) a (c) e comente o efeito do tamanho da amostra sobre a variância da distribuição amostral.

9.S11 Suponha que estejamos testando  $H_0: p = 0,5$  versus  $H_a: p \neq 0,5$ . Suponha que  $p$  seja o valor verdadeiro da proporção da população.

- Usando  $\alpha = 0,05$ , encontre a potência do teste para  $n = 100$ , 150 e 300, considerando  $p = 0,6$ . Comente o efeito do tamanho da amostra sobre a potência do teste.
- Usando  $\alpha = 0,01$ , encontre a potência do teste para  $n = 100$ , 150 e 300, considerando  $p = 0,6$ . Compare suas respostas com aquelas obtidas no item (a) e comente o efeito de  $\alpha$  sobre a potência do teste para diferentes tamanhos de amostras.
- Usando  $\alpha = 0,05$ , encontre a potência do teste para  $n = 100$ , considerando  $p = 0,08$ .

Compare sua resposta com aquela obtida no item (a) e comente o efeito do valor verdadeiro de  $p$  sobre a potência do teste para o mesmo tamanho de amostra e nível  $\alpha$ .

- Usando  $\alpha = 0,01$ , qual o tamanho requerido da amostra, se  $p = 0,6$  e se quisermos  $\beta = 0,05$ ? Qual o tamanho requerido da amostra se  $p = 0,8$  e se quisermos  $\beta = 0,05$ ? Compare os dois tamanhos de amostra e comente o efeito do valor verdadeiro de  $p$  sobre o tamanho requerido da amostra, quando  $\beta$  for mantido aproximadamente constante.

9.S12 O sistema de resfriamento em um submarino nuclear consiste em um arranjo de tubos soldados, pelos quais circula um líquido refrigerante. As especificações requerem que a resistência da solda tenha de encontrar ou exceder 150 psi.

a. Suponha que os engenheiros de projeto decidam testar a hipótese  $H_0: \mu = 150$  versus  $H_1: \mu > 150$ . Explique por que essa escolha de hipótese alternativa é melhor que  $H_1: \mu < 150$ .

b. Uma amostra aleatória de 20 soldas resulta em  $\bar{x} = 153,7$  psi e  $s = 11,3$  psi. Que conclusões você pode tirar em relação à hipótese no item (a)? Estabeleça quaisquer suposições necessárias sobre a distribuição em foco dos dados.

9.S13 Uma companhia de biotecnologia produz uma droga terapêutica, cuja concentração tem um desvio-padrão de quatro gramas por litro. Um novo método de produzir essa droga foi proposto, embora esteja envolvido algum custo adicional. O gerente autorizará uma mudança na técnica de produção, somente se o desvio-padrão da concentração no novo processo for menor do que quatro gramas por litro. Os pesquisadores escolheram  $n = 10$  e obtiveram os seguintes dados em gramas por litro. Faça a análise necessária para determinar se uma mudança na técnica de produção deveria ser implementada.

16,628	16,630
16,622	16,631
16,627	16,624
16,623	16,622
16,618	16,626

9.S14 Um fabricante de instrumentos de medidas com precisão afirma que o desvio-padrão no uso dos instrumentos é, no máximo, igual a 0,00002 milímetro. Um analista, que não está ciente da afirmação, usa o instrumento oito vezes e obtém um desvio-padrão da amostra de 0,00001 milímetro.

a. Confirme, usando um procedimento de teste e um nível  $\alpha$  de 0,01, que não há evidência suficiente para confirmar a afirmação de que o desvio-padrão dos instrumentos seja de, no máximo, 0,00002. Estabeleça quaisquer suposições necessárias sobre a distribuição em foco dos dados.

b. Explique por que o desvio-padrão da amostra,  $s = 0,00001$ , é menor do que 0,00002, embora os resultados do procedimento de teste estatístico não confirmem a afirmação.

9.S15 Considere os dados de qualidade de água do Exercício 6.2.10.

a. Esses dados suportam a afirmação de que a concentração média de sólidos suspensos não excede 50 partes por milhão? Use  $\alpha = 0,05$ .

b. Qual é o valor  $P$  para o teste do item (a)?

c. A distribuição normal parece ser um modelo razoável para esses dados? Faça um teste apropriado de adequação de ajuste para responder a essa questão.

9.S16 Considere os dados de resistência ao cisalhamento de soldas de ponto do Exercício 6.2.9. A distribuição normal parece ser um modelo razoável para esses dados? Faça um teste apropriado de adequação de ajuste para responder a essa questão.

9.S17 Um artigo em *Food Testing and Analysis* ["Improving Reproducibility of Refractometry Measurements of Fruit Juices" (1999, Vol. 4(4), pp. 13-17)] mediu a concentração (Brix) de açúcar em suco claro de maçã. Todas as leituras foram tomadas a 20 °C:

11,48	11,45	11,48	11,47	11,48
11,50	11,42	11,49	11,45	11,44
11,45	11,47	11,46	11,47	11,43
11,50	11,49	11,45	11,46	11,47

a. Teste a hipótese  $H_0: \mu = 11,5$  versus  $H_1: \mu \neq 11,5$ , usando  $\alpha = 0,05$ . Encontre o valor  $P$ .

b. Calcule a potência do teste, se a média verdadeira for 11,4.

c. Que tamanho de amostra seria requerido para detectar uma concentração média verdadeira de 11,45, se quiséssemos que a potência do teste fosse no mínimo 0,9?

d. Explique como a pergunta do item (a) poderia ser respondida construindo um intervalo de confiança bilateral para a concentração média de açúcar.

e. Há evidência para confirmar a suposição de que a concentração de açúcar seja normalmente distribuída?

9.S18 Considere a saída computacional a seguir.

#### Teste e IC para Uma Proporção:

Teste de  $P = 0,25$  versus  $P < 0,25$

X	N	Amostra P	Limite	Valor Z	Valor P
53	225	0,235556	0,282088	?	?

Usando a aproximação pela normal:

a. Preencha as informações que faltam.

b. Quais são suas conclusões, se  $\alpha = 0,05$ ?

c. A aproximação da binomial pela normal foi usada aqui. Ela foi apropriada?

d. Encontre um limite superior de confiança de 95 % para a proporção verdadeira.

e. Qual seria o valor  $P$  e quais seriam suas conclusões, se a hipótese alternativa fosse  $H_1: p \neq 0,25$ ?

9.S19 Um artigo em *Experimental Brain Research* ["Synapses in the Granule Cell Layer of the Rat Dentate Gyrus: Serial-Sectionin Study" (1996, Vol. 112(2), pp. 237-243)] mostrou a razão entre os números de sinapses simétricas e totais em células e segmentos iniciais do axônio de células granulares reconstruídas no giro denteado de um rato com 12 semanas de vida:

0,65	0,90	0,78	0,94	0,40	0,94
0,91	0,86	0,53	0,84	0,42	0,50
0,50	0,68	1,00	0,57	1,00	1,00
0,84	0,9	0,91	0,92	0,96	
0,96	0,56	0,67	0,96	0,52	
0,89	0,60	0,54			

a. Use os dados para testar  $H_0: \sigma^2 = 0,02$  versus  $H_1: \sigma^2 \neq 0,02$ , usando  $\alpha = 0,05$ .

b. Encontre o valor  $P$  para o teste.

9.S20 Um artigo em *Biological Trace Element Research* ["Interaction of Dietary Calcium, Manganese, and Manganese Source (Mn Oxide or Mn Methionine Complex) on Chick Performance and Manganese Utilization" (1991, Vol. 29(3), pp. 217-228)] mostrou os seguintes resultados de ensaio de manganês em tecidos do fígado de pintos alimentados com dietas ricas em Ca.

6,02	6,08	7,11	5,73	5,32	7,10
5,29	5,84	6,03	5,99	4,53	6,81

a. Teste a hipótese  $H_0: \sigma^2 = 0,6$  versus  $H_1: \sigma^2 \neq 0,6$ , usando  $\alpha = 0,01$ .

b. Qual é o valor  $P$  para esse teste?

c. Discuta como o item (a) poderia ser respondido pela elaboração de um intervalo de confiança bilateral de 99 % para  $\sigma$ .

9.S21 Considere o experimento do brilho do tubo de imagem de uma televisão, descrito no Exercício 8.2.7.

a. Para esse tamanho de amostra  $n = 10$ , os dados confirmam a afirmação de que o desvio-padrão de corrente é menor do que 20 microampères?

b. Em vez de  $n = 10$ , suponha que o tamanho da amostra tenha sido 51. Repita a análise feita no item (a), usando  $n = 51$ .

c. Compare suas respostas e comente como o tamanho da amostra afeta suas conclusões obtidas nos itens (a) e (b).

## Capítulo 10 Exercícios

### Exercícios para a Seção 10.1

10.1.1 Considere o teste de hipóteses  $H_0: \mu_1 = \mu_2$  contra  $H_1: \mu_1 \neq \mu_2$ , com variâncias conhecidas  $\sigma_1 = 10$  e  $\sigma_2 = 5$ . Suponha que os tamanhos de amostra sejam  $n_1 = 10$  e  $n_2 = 15$  e que  $\bar{x}_1 = 4,7$  e  $\bar{x}_2 = 7,8$ . Use  $\alpha = 0,05$ .

- Teste a hipótese e encontre o valor  $P$ .
- Explique como o teste poderia ser conduzido com um intervalo de confiança.
- Qual é a potência do teste, no item (a), para uma diferença verdadeira nas médias de 3?
- Considerando tamanhos iguais de amostras, que tamanho de amostra deveria ser usado para se obter  $\beta = 0,05$ , se a diferença verdadeira nas médias fosse 3? Suponha  $\alpha = 0,05$ .

10.1.2 Duas máquinas são usadas para encher garrafas de plástico que têm um volume líquido de 16,0 onças. O volume de enchimento pode ser considerado normal, com um desvio-padrão  $\sigma_1 = 0,020$  e  $\sigma_2 = 0,025$  onça. Um membro do grupo de engenheiros da qualidade suspeita de que ambas as máquinas encham até o mesmo volume líquido médio, independente de esse volume ser ou não 16,0 onças. Uma amostra aleatória de dez garrafas é retirada na saída de cada máquina.

Máquina 1		Máquina 2	
16,03	16,01	16,02	16,03
16,04	15,96	15,97	16,04
16,05	15,98	15,96	16,02
16,05	16,02	16,01	16,01
16,02	15,99	15,99	16,00

- Você acha que o engenheiro está correto? Considere  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?
- Calcule um intervalo de confiança de 95 % para a diferença de médias. Dê uma interpretação prática para esse intervalo.
- Qual é a potência do teste no item (a) para uma diferença verdadeira de médias igual a 0,04?
- Supondo tamanhos iguais de amostra, que tamanho de amostra deveria ser usado para assegurar que  $\beta = 0,05$ , se a diferença verdadeira de médias for 0,04? Considere  $\alpha = 0,05$ .

10.1.3 Estão sendo estudadas as taxas de queima de dois diferentes propelentes sólidos usados no sistema de escape da tripulação de aeronaves. Sabe-se que ambos os propelentes têm aproximadamente o mesmo desvio-padrão de taxa de queima; ou seja,  $\sigma_1 = \sigma_2 = 3$  centímetros por segundo. Duas amostras aleatórias de  $n_1 = 20$  e  $n_2 = 20$  espécimes são testadas; as taxas médias de queima das amostras são  $\bar{x}_1 = 18$  centímetros por segundo e  $\bar{x}_2 = 24$  centímetros por segundo.

- Teste a hipótese de que ambos os propelentes têm a mesma taxa média de queima. Use  $\alpha = 0,05$ . Qual é o valor  $P$ ?

b. Construa um intervalo de confiança de 95 % para a diferença de médias  $\mu_1 - \mu_2$ . Qual é o significado prático desse intervalo?

c. Qual é o erro  $\beta$  do teste no item (a), se a diferença verdadeira na taxa média de queima for 2,5 centímetros por segundo?

d. Supondo tamanhos iguais de amostra, que tamanho de amostra é necessário para se obter uma potência de 0,9, se a diferença verdadeira de médias for de 14 cm/s?

10.1.4 Um polímero é fabricado em uma batelada de um processo químico. Medidas de viscosidade são normalmente feitas em cada batelada, e a longa experiência com o processo tem indicado que sua variabilidade é razoavelmente estável, com  $\sigma = 20$ . Quinze bateladas de medidas de viscosidade são dadas a seguir:

724, 718, 776, 760, 745, 759, 795, 756, 742, 740, 761, 749, 739, 747, 742

Faz-se uma mudança no processo, que consiste em alterar o tipo de catalisador empregado. Depois da mudança no processo, oito medidas de viscosidade são feitas:

735, 775, 729, 755, 783, 760, 738, 780

Suponha que a variabilidade do processo não seja afetada pela alteração de catalisador. Se a diferença na viscosidade média da batelada fosse dez ou menos, o fabricante gostaria de detectá-la com uma alta probabilidade.

- Formule e teste uma hipótese apropriada, considerando  $\alpha = 0,10$ . Quais são suas conclusões? Encontre o valor  $P$ ?
- Encontre um intervalo de confiança de 90 % para a diferença de viscosidades médias em cada batelada, resultante do processo de mudança.
- Compare os resultados dos itens (a) e (b) e discuta sobre o que você achou.

10.1.5 Um artigo em *Industrial Engineer* (setembro de 2012) reportou um estudo de fontes potenciais de ferimento em veterinários de cavalos, conduzido em um hospital veterinário universitário. Forças na mão foram medidas para várias atividades comuns em que os veterinários estão envolvidos quando fazem exames ou tratam cavalos. Vamos considerar as forças nas mãos para as tarefas de levantar e usar o ultrassom. Admita que ambos os tamanhos de amostra sejam iguais a 6, que a força média da amostra para levantar seja 6,0 libras, com um desvio-padrão de 1,5 libra, e que a força média da amostra para usar o ultrassom seja de 6,2 libras, com um desvio-padrão de 0,3 libra (dados lidos de gráficos no artigo). Considere que os desvios-padrão sejam conhecidos. Há evidência para concluir que as duas atividades resultem em diferenças significativamente diferentes para as mãos?

## Exercícios para a Seção 10.2

10.2.1 Considere o resultado a seguir.

**Teste T para Duas Amostras e IC:**

Amostra	N	Média	DP	EP da Média
1	12	10,94	1,26	0,36
2	16	12,15	1,99	0,50

$$\text{Diferença} = \mu_1 - \mu_2$$

Estimativa para a diferença:  $-1,210$

IC de 95 % para a diferença:  $(-2,560; 0,140)$

Teste  $t$  da diferença =  $0$  (vs  $\neq 0$ )

Valor  $t$  = ? Valor  $P$  = ? GL = ?

Ambos usam o DP combinado = ?

- Preencha os valores que faltam. Esse é um teste unilateral ou bilateral? Use os limites inferior e superior para o valor  $P$ .
- Quais as suas conclusões se  $\alpha = 0,05$ ? E se  $\alpha = 0,01$ ?
- Esse teste foi feito considerando que as variâncias das duas populações são iguais. Isso parece razoável?
- Suponha que a hipótese tivesse sido  $H_0: \mu_1 = \mu_2$  versus  $H_a: \mu_1 < \mu_2$ . Quais seriam suas conclusões se  $\alpha = 0,05$ ?

10.2.2 Considere o teste de hipóteses  $H_0: \mu_1 = \mu_2$  contra  $H_1: \mu_1 \neq \mu_2$ . Suponha que os tamanhos de amostra sejam  $n_1 = 15$  e  $n_2 = 15$ , que  $\bar{x}_1 = 4,7$  e  $\bar{x}_2 = 7,8$  e que  $s_1^2 = 4$  e  $s_2^2 = 6,25$ . Considere que  $\sigma_1^2 = \sigma_2^2$  e que os dados sejam retirados de distribuições normais. Use  $\alpha = 0,05$ .

- Teste a hipótese e encontre o valor  $P$ .
- Explique como o teste poderia ser conduzido com um intervalo de confiança.
- Qual é a potência do teste no item (a) para uma diferença verdadeira de médias igual a 3?
- Considerando tamanhos iguais de amostras, que tamanho de amostra deveria ser utilizado para se obter  $\beta = 0,05$ , se a diferença verdadeira de médias fosse  $-2$ ? Suponha  $\alpha = 0,05$ .

10.2.3 O diâmetro de bastões de aço fabricados em duas máquinas extrusoras diferentes está sendo investigado. Duas amostras aleatórias com tamanhos  $n_1 = 15$  e  $n_2 = 17$  são selecionadas, e as médias e variâncias das amostras são  $\bar{x}_1 = 8,73$ ,  $s_1^2 = 0,35$ ,  $\bar{x}_2 = 8,68$  e  $s_2^2 = 0,40$ , respectivamente. Suponha que  $\sigma_1^2 = \sigma_2^2$  e que os dados sejam retirados de uma população normal.

- Há evidência que confirme a afirmação de que as duas máquinas produzem bastões com diferentes diâmetros médios? Use  $\alpha = 0,05$  para chegar a essa conclusão. Encontre o valor  $P$ .
- Construa um intervalo de confiança de 95 % para a diferença no diâmetro médio dos bastões. Interprete esse intervalo.

10.2.4 Considere o teste de hipóteses  $H_0: \mu_1 = \mu_2$  contra  $H_1: \mu_1 \neq \mu_2$ . Suponha que os tamanhos de amostra sejam  $n_1 = 10$  e  $n_2 = 10$ , que  $\bar{x}_1 = 7,8$  e  $\bar{x}_2 = 5,6$  e que  $s_1^2 = 4$  e  $s_2^2 = 9$ . Considere que  $\sigma_1^2 = \sigma_2^2$  e que os dados sejam retirados de distribuições normais. Use  $\alpha = 0,05$ .

- Teste a hipótese e encontre o valor  $P$ .
- Explique como o teste poderia ser conduzido com um intervalo de confiança.
- Qual será a potência do teste no item (a), se  $\mu_1$  for três unidades maior que  $\mu_2$ ?
- Considerando tamanhos iguais de amostras, que tamanho de amostra deveria ser usado para se obter  $\beta = 0,05$ , se  $\mu_1$  fosse três unidades maior que  $\mu_2$ ? Admita  $\alpha = 0,05$ .

10.2.5 Dois catalisadores podem ser empregados em um processo químico em batelada. Doze bateladas foram preparadas usando catalisador 1, resultando em um rendimento médio de 86 e um desvio-padrão da amostra de 3. Quinze bateladas foram preparadas usando catalisador 2, resultando em um rendimento médio de 89, com um desvio-padrão de 2. Considere que as medidas de rendimento sejam distribuídas aproximadamente de forma normal, com o mesmo desvio-padrão.

- Há evidência que confirme a afirmação de que o catalisador 2 produz um rendimento maior do que o catalisador 1? Use  $\alpha = 0,01$ .
- Encontre um intervalo de confiança de 99 %, para a diferença entre os rendimentos médios, que possa ser usado para testar a afirmação do item (a).

10.2.6 Dois fornecedores fabricam uma engrenagem de plástico usada em uma impressora a laser. A resistência ao impacto (medida em lbf-ft) dessas engrenagens é uma característica importante. Uma amostra aleatória de dez engrenagens do fornecedor 1 resulta em  $\bar{x}_1 = 290$  e  $s_1 = 12$ , enquanto outra amostra aleatória de 16 engrenagens do segundo fornecedor resulta em  $\bar{x}_2 = 321$  e  $s_2 = 22$ .

- Há evidência confirmando a afirmação de que o fornecedor 2 fornece engrenagens com maiores resistências médias ao impacto? Use  $\alpha = 0,05$  e considere que ambas as populações sejam normalmente distribuídas e que as variâncias não sejam iguais. Qual é o valor  $P$  para esse teste?
- Os dados confirmam a afirmação de que a resistência média ao impacto de engrenagens provenientes do fornecedor 2 é no mínimo 25 lbf-ft maior que aquela do fornecedor 1? Faça as mesmas suposições adotadas no item (a).
- Construa uma estimativa do intervalo de confiança para a diferença na resistência média ao impacto e explique como esse intervalo poderia ser empregado para responder à questão colocada referente às diferenças entre os fornecedores.

10.2.7 Um artigo em *Electronic Components and Technology Conference* (2001, Vol. 52, pp. 1167-1171) comparou os processos de serragem com um único eixo e com duplo eixo para pastilhas metalizadas de cobre. A largura dos cavacos posteriores foi medida em um total de 15 dispositivos de cada tipo, encontrando-se  $\bar{x}_{\text{único}} = 66,385$ ,  $s_{\text{único}} = 7,895$  e  $\bar{x}_{\text{duplo}} = 45,278$ ,  $s_{\text{duplo}} = 8,612$ .

- Os dados confirmam a afirmação de que ambos os processos produzem o mesmo tipo de cavaco? Use  $\alpha = 0,05$  e considere que ambas as populações sejam normalmente distribuídas e tenham a mesma variância. Encontre o valor  $P$  para esse teste.
- Construa um intervalo bilateral de confiança de 95 % para a diferença de médias dos processos de serragem. Compare esse intervalo aos resultados do item (a).
- Se o erro  $\beta$  do teste não exceder 0,1 quando a diferença verdadeira nas lascas for 15, que tamanhos de amostra terão de ser usados? Use  $\alpha = 0,05$ .

10.2.8 Um filme fotocondutor é fabricado com uma espessura nominal de  $25 \times 10^{-3}$  polegada. O engenheiro da produção deseja aumentar a velocidade média do filme e acredita que isso possa ser atingido pela redução da espessura para  $20 \times 10^{-3}$  polegada. Oito amostras de cada espessura do filme são fabricadas em um processo piloto de produção, e a velocidade (em microjoules por polegada quadrada) do filme é medida. Para o filme de  $25 \times 10^{-3}$  polegada, os resultados dos dados das amostras são  $\bar{x}_1 = 1,15$  e  $s_1 = 0,11$ , enquanto para o filme de  $20 \times 10^{-3}$  polegada, os dados resultam em  $\bar{x}_2 = 1,06$  e  $s_2 = 0,09$ . Note que um aumento na velocidade do filme baixaria o valor da observação em microjoules por polegada quadrada.

a. Os dados confirmam a afirmação de que a redução da espessura do filme aumenta a sua velocidade média? Use  $\alpha = 0,10$  e considere que as variâncias das duas populações sejam iguais e que a população em foco da velocidade do filme seja normalmente distribuída. Qual é o valor  $P$  para esse teste?

b. Encontre um intervalo de confiança de 95 %, para a diferença entre as duas médias, que pode ser usado para testar a afirmação do item (a).

**10.2.9** Um artigo em *Radio Engineering and Electronic Physics* [1984, Vol. 29(3), pp. 63-66] investigou o comportamento de um gerador estocástico na presença de um ruído externo. O número de períodos foi medido em uma amostra de 100 trens para cada um dos dois níveis diferentes de voltagem do ruído, 100 e 150 mV. Para 100 mV, o número médio de períodos em um trem foi de 7,9 com  $s = 2,6$ . Para 150 mV, a média foi de 6,9 com  $s = 2,4$ .

a. Originalmente, suspeitou-se de que, aumentando a voltagem do ruído, reduziria o número médio de períodos. Os dados suportam essa afirmação? Use  $\alpha = 0,01$  e considere que cada população seja normalmente distribuída e que as variâncias das duas populações sejam iguais. Qual é o valor  $P$  para esse teste?

b. Calcule um intervalo de confiança para responder à questão do item (a).

**10.2.10** Um artigo em *IEEE International Symposium on Electromagnetic Compatibility* ["EM Effects of Different Mobile Handsets on Rats Brain" (2002, Vol. 2, pp. 667-670)] quantificou a absorção de energia eletromagnética e o efeito térmico resultante proveniente de telefones celulares. Os resultados experimentais foram obtidos a partir de experimentos *in vivo*, conduzidos em ratos. Os valores da pressão sanguínea arterial (mmHg) para o grupo de controle (oito ratos) durante o experimento são  $\bar{x}_1 = 90$ ,  $s_1 = 5$  e para o grupo de teste (nove ratos) são  $\bar{x}_2 = 115$  e  $s_2 = 10$ .

a. Há evidência que suporte a afirmação de que o grupo de teste tem uma pressão sanguínea maior? Use  $\alpha = 0,05$  e considere que ambas as populações são distribuídas normalmente, porém as variâncias não são iguais. Qual é o valor  $P$  para esse teste?

b. Calcule um intervalo de confiança para responder à questão do item (a).

c. Os dados confirmam a afirmação de que a pressão sanguínea média do grupo de teste é, no mínimo, 15 mmHg maior que a do grupo de controle? Faça as mesmas suposições feitas no item (a).

d. Explique como a questão do item (c) poderia ser respondida com um intervalo de confiança.

**10.2.11** A distância global viajada por uma bola de golfe é testada, atirando a bola com o *Iron Byron*, um jogador mecânico, com um balanço que é dito emular o legendário campeão, Byron Nelson. Dez bolas, selecionadas aleatoriamente a partir de duas marcas diferentes, são testadas, sendo a distância global medida. Os dados são:

Marca 1: 275, 286, 287, 271, 283, 271, 279, 275, 263, 267

Marca 2: 258, 244, 260, 265, 273, 281, 271, 270, 263, 268

a. Há evidência de que a distância global seja distribuída aproximadamente de forma normal? Uma suposição de variâncias iguais é justificada?

b. Teste a hipótese de que ambas as marcas de bola têm iguais distâncias médias globais. Use  $\alpha = 0,05$ . Qual é o valor  $P$ ?

c. Construa um intervalo de confiança de 95 % para a diferença média entre as duas marcas de bolas de golfe.

d. Qual é a potência do teste estatístico do item (b) para detectar uma diferença verdadeira na distância média global de cinco jardas?

e. Que tamanho de amostra seria requerido para detectar uma diferença verdadeira na distância média global de três jardas com potência de aproximadamente 0,75?

**10.2.12** O "efeito mola" em um taco de golfe poderia ser determinado medindo o coeficiente de restituição (a razão entre a velocidade de saída da bola e a velocidade de chegada de uma bola de golfe atirada na cabeça do taco). Doze tacos direcionadores, selecionados aleatoriamente e produzidos por dois fabricantes de taco, são testados, sendo o coeficiente de restituição medido. Os dados são:

Taco 1: 0,8406, 0,8104, 0,8234, 0,8198, 0,8235, 0,8562,  
0,8123, 0,7976, 0,8184, 0,8265, 0,7773, 0,7871

Taco 2: 0,8305, 0,7905, 0,8352, 0,8380, 0,8145, 0,8465,  
0,8244, 0,8014, 0,8309, 0,8405, 0,8256, 0,8476

a. Há evidência de que o coeficiente de restituição seja distribuído aproximadamente de forma normal? Uma suposição de variâncias iguais é justificada?

b. Teste a hipótese de que ambas as marcas de taco têm iguais coeficientes médios de restituição. Use  $\alpha = 0,05$ . Qual é o valor  $P$  do teste?

c. Construa um intervalo de confiança de 95 % para a diferença média entre as duas marcas de tacos de golfe.

d. Qual é a potência do teste estatístico do item (b) para detectar uma diferença verdadeira no coeficiente de restituição de 0,2?

e. Que tamanho de amostra seria requerido para detectar uma diferença verdadeira no coeficiente médio de restituição de 0,1 com potência de aproximadamente 0,8?

**10.2.13** Um artigo em *Quality Engineering* [2013, Vol. 25(1)] apresentou dados para ciclos de falha de juntas soldadas em diferentes temperaturas para tipos diferentes de placas de circuito impresso (PCI). São mostrados, a seguir, dados de falhas para duas temperaturas (20 e 60 °C) para PCI de cobre-níquel-ouro.

20°C 218, 265, 279, 282, 336, 469, 496, 507, 685, 685

60°C 185, 242, 254, 280, 305, 353, 381, 504, 556, 697

a. Teste a hipótese nula, com  $\alpha = 0,05$ , de que os ciclos de falha são os mesmos em ambas as temperaturas. A hipótese alternativa é unilateral ou bilateral?

b. Encontre um intervalo de confiança de 95 % para os ciclos médios de falha para as duas temperaturas.

c. O intervalo de confiança de 95 % contém o valor zero? Explique a conexão com a conclusão que você alcançou no item (a).

d. Os gráficos de probabilidade normal dos ciclos de falha indicam qualquer violação das suposições para os testes e para o intervalo de confiança que você fez?

**10.2.14** Um artigo em *Quality Engineering* [2012, Vol. 24(1)] descreveu um experimento sobre rebolo abrasivo. Os seguintes dados são referentes à força (em N) de moagem desse experimento em dois níveis diferentes de vibração.

Baixa 242, 249, 235, 250, 254, 244, 258, 311, 237, 261, 314, 252

Alta 302, 421, 419, 399, 317, 311, 350, 363, 392, 367, 301, 302

a. Há evidência para confirmar a suposição de que a força média de moagem aumenta com o nível de vibração?

b. Encontre um intervalo de confiança de 95 % para a diferença na força média de moagem para os dois níveis de vibração.

c. O intervalo de confiança de 95 % contém o valor zero? Explique a conexão com a conclusão que você alcançou no item (a).

d. Os gráficos de probabilidade normal da força de moagem indicam qualquer violação das suposições para os testes e para o intervalo de confiança que você fez?

## Exercícios para a Seção 10.3

**10.3.1** Um engenheiro elétrico tem de projetar um circuito para operar com uma quantidade máxima de corrente em um tubo de imagem, de modo a encontrar uma imagem com brilho suficiente. Dentro de suas restrições de projeto, ele desenvolveu dois circuitos candidatos e testou protótipos de cada um. Os dados resultantes (em microampères) são apresentados a seguir:

**Circuito 1:** 251, 255, 258, 257, 250, 251, 254, 250, 248

**Circuito 2:** 250, 253, 249, 256, 259, 252, 260, 251

- Use o teste de Wilcoxon da soma dos postos para testar  $H_0: \mu_1 = \mu_2$  contra a alternativa  $H_1: \mu_1 > \mu_2$ . Use  $\alpha = 0,025$ .
- Aplique a aproximação normal para o teste de Wilcoxon da soma dos postos. Considere que  $\alpha = 0,05$ . Encontre o valor  $P$  aproximado para essa estatística de teste.

**10.3.2** Um dos autores deste livro viaja regularmente para Seattle, Washington. Ele usa uma das duas companhias aéreas, Delta ou Alasca. Atrasos de voos são, por vezes, inevitáveis, porém ele estaria disposto a viajar mais pela companhia aérea que tivesse um melhor registro de chegadas na hora marcada. O número de minutos que seu voo atrasou nas últimas seis viagens em cada linha aérea é mostrado a seguir. Há qualquer evidência de qual companhia aérea tem desempenho superior com relação ao número de chegadas na hora? Use  $\alpha = 0,01$  e o teste de Wilcoxon da soma dos postos.

**Delta:** 13, 10, 1, 4, 0, 9 (minutos de atraso)

**Alasca:** 15, 8, 3, -1, -2, 4 (minutos de atraso)

**10.3.3** O fabricante de uma banheira está interessado em testar dois aquecedores diferentes para seu produto. O aquecedor que produzir o máximo ganho de calor depois de 15 minutos será o preferível. O fabricante obtém dez amostras de cada unidade aquecedora e testa cada uma. O ganho de calor (aumento de temperatura em °F) depois de 15 minutos é mostrado a seguir:

**Unidade 1:** 25, 27, 29, 31, 30, 26, 24, 32, 33, 38

**Unidade 2:** 31, 33, 32, 35, 34, 29, 38, 35, 37, 30

- Existe alguma razão para suspeitar de que uma unidade seja superior à outra? Use  $\alpha = 0,05$  e o teste de Wilcoxon da soma dos postos.
- Use a aproximação normal para o teste de Wilcoxon da soma dos postos. Considere que  $\alpha = 0,05$ . Encontre o valor  $P$  aproximado para essa estatística de teste.

**10.3.4** Considere a distância viajada por uma bola de golfe, descrita no Exercício 10.2.11.

- Aplique o teste de Wilcoxon da soma dos postos para investigar se as médias diferem. Use  $\alpha = 0,05$ .
- Considere a aproximação normal para o teste de Wilcoxon da soma dos postos com  $\alpha = 0,05$ . Encontre o valor  $P$  aproximado para esse teste.

## Exercícios para a Seção 10.4

**10.4.1** O gerente de uma frota de carros está testando duas marcas de pneus radiais. Ele coloca, ao acaso, um pneu de cada marca nas duas rodas traseiras de oito carros e anda com os carros até que os pneus se desgastem. Os dados (em quilômetros) são mostrados a seguir. Encontre um intervalo de confiança de 99 % para a diferença na vida média. Com base em seus cálculos, qual a marca que você prefere?

Carro	Marca 1	Marca 2
1	36.925	34.318
2	45.300	42.280
3	36.240	35.500
4	32.100	31.950
5	37.210	38.015
6	48.360	47.800
7	38.200	37.810
8	33.500	33.215

**10.4.2** Um cientista de computação está investigando a utilidade de duas diferentes linguagens de programação na melhoria das tarefas computacionais. Doze programadores experientes, familiarizados com ambas as linguagens, codificaram uma função-padrão nas duas linguagens. O tempo (em minutos) foi registrado, e os dados são mostrados a seguir.

Programador	Tempo	
	Linguagem de Programação 1	Linguagem de Programação 2
1	17	18
2	16	14
3	21	19
4	14	11
5	18	23
6	24	21
7	16	10
8	14	13
9	21	19
10	23	24
11	13	15
12	18	20

- A suposição de a diferença no tempo de codificação ser normalmente distribuída é razoável?
- Encontre um intervalo de confiança de 95 % para a diferença nos tempos médios de codificação. Existe alguma indicação de que uma linguagem de programação seja preferível?

10.4.3 Um artigo publicado no *Journal of Aircraft* [“Equivalent Plate Analysis of Aircraft Wing Box Structures with General Planform Geometry” (Vol. 23, 1986, pp. 859-864)] descreveu uma nova formulação do método de análise de placa equivalente, que é capaz de modelar estruturas de aviões, tais como vigas-caixão nas asas de aviões, e que produz resultados similares ao método de análise por elementos finitos, que é mais laborioso computacionalmente. Frequências naturais de vibração para a estrutura das vigas-caixão nas asas de aviões são calculadas usando ambos os métodos, sendo os resultados mostrados a seguir para as sete primeiras frequências naturais.

Freq.	Elemento Finito, Ciclos/s	Placas Equivalentes, Ciclos/s
1	14,58	14,76
2	48,52	49,10
3	97,22	99,99
4	113,99	117,53
5	174,73	181,22
6	212,72	220,14
7	277,38	294,80

- a. Os dados sugerem que os dois métodos fornecem o mesmo valor médio para a frequência natural de vibração? Use  $\alpha = 0,05$ . Encontre o valor  $P$ .
- b. Encontre um intervalo de confiança de 95 % para a diferença média entre os dois métodos.

10.4.4 Dois diferentes testes analíticos podem ser aplicados para determinar o nível de impurezas em ligas de aço. Oito espécimes são testados usando ambos os procedimentos, sendo os resultados mostrados na tabela a seguir.

Espécime	Teste 1	Teste 2
1	1,2	1,4
2	1,3	1,7
3	1,5	1,5
4	1,4	1,3
5	1,7	2,0
6	1,8	2,1
7	1,4	1,7
8	1,3	1,6

- a. Há evidência suficiente para concluir que os testes diferem no nível médio de impureza? Use  $\alpha = 0,01$ .
- b. Há evidência para confirmar a afirmação de que o teste 1 gera uma diferença média 0,1 unidade menor que o teste 2? Use  $\alpha = 0,05$ .
- c. Se a média do Teste 1 for 0,1 menor que a média do teste 2, é importante detectar isso com uma probabilidade de no mínimo 0,90. O emprego de oito ligas foi adequado? Se não, quantas ligas deveriam ter sido usadas?

10.4.5 Neurocientistas conduziram uma pesquisa em uma prisão canadense para ver se o confinamento solitário afetaria a atividade das ondas cerebrais [“Change in EEG Alpha Frequency and Evoked Response Latency During Solitary Confinement”, *Journal of Abnormal Psychology* (1972, Vol. 7, pp. 54-59)]. Eles selecionaram

aleatoriamente 20 ocupantes, distribuindo-os em dois grupos, indo metade deles para o confinamento solitário e a outra metade para o confinamento regular. Eis os dados:

Não Confinado	Confinado
10,7	9,6
10,7	10,4
10,4	9,7
10,9	10,3
10,5	9,2
10,3	9,3
9,6	9,9
11,1	9,5
11,2	9,0
10,4	10,9

- a. O teste  $t$  pareado é apropriado para testar se as frequências médias da onda alfa são as mesmas nos dois grupos? Explique.
- b. Faça um teste apropriado.

10.4.6 Em *Biometrics* (1990, Vol. 46, pp. 673-687), os autores analisaram a circunferência de cinco laranjeiras (marcadas como A-E), que foram medidas em sete ocasiões ( $x_i$ ).

Árvore	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
A	30	58	87	115	120	142	145
B	33	69	111	156	172	203	203
C	30	51	75	108	115	139	140
D	32	62	112	167	179	209	214
E	30	49	81	125	142	174	177

- a. Compare o aumento médio na circunferência nos períodos 1 a 2 com o aumento médio para os períodos 2 a 3. O aumento é a diferença na circunferência nos dois períodos. Essas médias são significativamente diferentes com  $\alpha = 0,10$ ?
- b. Há evidência de que o aumento médio no período 1 a 2 seja maior que o aumento médio no período 6 a 7 com  $\alpha = 0,05$ ?
- c. As suposições do teste do item (a) são violadas porque os mesmos dados (circunferência do período 2) são usados para calcular ambos os aumentos médios?

10.4.7 Em uma série de testes para estudar a eficácia de *ginkgo biloba* na memória, Solomon *et al.* olharam primeiro diferenças nos testes de memória de pessoas em seis semanas, antes e depois de aderirem ao estudo [“Ginkgo for Memory Enhancement: A Randomized Controlled Trial”, *Journal of the American Medical Association* (2002, Vol. 288, pp. 835-840)]. Para 99 pacientes que não receberam medicamento, o aumento médio na categoria fluência (número de palavras geradas em um minuto) foi 1,07 palavra, com um desvio-padrão de 3,195 palavras. Pesquisadores queriam saber se o número médio de palavras lembradas foi positivo.

- a. Esse é um teste unilateral ou bilateral?
- b. Faça um teste de hipóteses para determinar se o aumento da média é zero.
- c. Por que esse pode ser visto como um teste  $t$  emparelhado?
- d. O que a conclusão diz acerca da importância de incluir placebos em tais testes?

## Exercícios para a Seção 10.5

10.5.1 Para uma distribuição  $F$ , encontre o seguinte:

- a.  $f_{0,25;5;10}$
- b.  $f_{0,10;24;9}$
- c.  $f_{0,05;8;15}$
- d.  $f_{0,75;5;10}$
- e.  $f_{0,90;24;9}$
- f.  $f_{0,95;8;15}$

10.5.2 Duas companhias químicas podem fornecer uma matéria-prima, cuja concentração de determinado elemento é importante. A concentração média para ambos os fornecedores é a mesma, porém suspeitamos de que a variabilidade na concentração pode diferir entre as

duas companhias. O desvio-padrão da concentração em uma amostra aleatória de  $n_1 = 10$  bateladas produzidas pela companhia 1 é  $s_1 = 4,7$  gramas por litro, enquanto, para a companhia 2, uma amostra aleatória, de  $n_2 = 16$  bateladas, resulta em  $s_2 = 5,8$  gramas por litro. Há evidência suficiente para concluir que as variâncias das duas populações difiram? Use  $\alpha = 0,05$ .

**10.5.3** Considere o teste de hipóteses  $H_0: \sigma_1^2 = \sigma_2^2$  contra  $H_1: \sigma_1^2 < \sigma_2^2$ . Suponha que os tamanhos de amostra sejam  $n_1 = 5$  e  $n_2 = 10$  e que  $s_1^2 = 23,2$  e  $s_2^2 = 28,8$ . Use  $\alpha = 0,05$ . Teste a hipótese e explique como o teste poderia ser conduzido usando um intervalo de confiança para  $\sigma_1/\sigma_2$ .

**10.5.4** Considere os dados sobre a resistência ao impacto da engrenagem do Exercício 10.2.6. Há evidência suficiente para concluir que a

variância da resistência ao impacto seja diferente para os dois fornecedores? Use  $\alpha = 0,05$ .

**10.5.5** Reconsidere os dados da distância global para as bolas de golfe do Exercício 10.2.11. Existe evidência que confirme a afirmação de que o desvio-padrão da distância global é o mesmo para ambas as marcas de bolas (use  $\alpha = 0,05$ )? Explique como responder a essa questão com um intervalo de confiança de 95 % para  $\sigma_1/\sigma_2$ .

**10.5.6** Reconsidere os dados do coeficiente de restituição do Exercício 10.2.12. Os dados sugerem que o desvio-padrão da distância global seja o mesmo para ambas as marcas de tacos direcionadores (use  $\alpha = 0,05$ )? Explique como responder a essa questão com um intervalo de confiança para  $\sigma_1/\sigma_2$ .

## Exercícios para a Seção 10.6

**10.6.1** Considere o seguinte resultado:

### Teste e IC para Duas Proporções:

Amostra	X	N	Amostra p
1	54	250	0,216000
2	60	290	0,206897

Diferença =  $p_1 - p_2$

Estimativa para a diferença: 0,00910345

IC de 95 % para a diferença: (-0,0600031; 0,0782100)

Teste para a diferença

= 0 (vs  $\neq 0$ ): Z = ? Valor P = ?

- Esse é um teste unilateral ou bilateral?
- Preencha os valores que faltam.
- A hipótese nula pode ser rejeitada?
- Construa um IC aproximado de 90 % para a diferença nas duas proporções.

**10.6.2** Um artigo em *Knee Surgery, Sports Traumatology, Arthroscopy* (2005, Vol. 13(4), pp. 273-279) considerou o reparo artroscópico no menisco, feito com um parafuso absorvível. Resultados mostraram que, para rompimentos maiores que 25 milímetros, 14 de 18 (78 %) reparos foram bem-sucedidos, e, para rompimentos mais curtos, 22 de 30 (73 %) dos reparos tiveram sucesso.

- Há evidência de que a taxa de sucesso seja maior para rompimentos mais longos? Use  $\alpha = 0,05$ . Qual é o valor  $P$ ?
- Calcule um limite unilateral de confiança de 95 % para a diferença de proporções que pode ser usado para responder à questão do item (a).

**10.6.3** Dois tipos diferentes de máquinas de moldagem por injeção são usados para formar peças de plásticos. Uma peça é considerada defeituosa se tiver excesso de encolhimento ou se for descolorida. Duas amostras aleatórias, cada uma de tamanho 300, são selecionadas e 15 peças defeituosas são encontradas na amostra da máquina 1, enquanto oito peças defeituosas são encontradas na amostra da máquina 2.

- É razoável concluir que ambas as máquinas produzam a mesma fração de peças defeituosas, usando  $\alpha = 0,05$ ? Encontre o valor  $P$  para esse teste.

b. Construa um intervalo de confiança de 95 % para a diferença das duas frações defeituosas.

c. Suponha que  $p_1 = 0,05$  e  $p_2 = 0,01$ . Com os tamanhos das amostras dados aqui, qual é a potência do teste para essa alternativa bilateral?

d. Suponha que  $p_1 = 0,05$  e  $p_2 = 0,01$ . Determine o tamanho necessário da amostra para detectar essa diferença, com uma probabilidade de no mínimo 0,9.

e. Suponha que  $p_1 = 0,05$  e  $p_2 = 0,02$ . Com os tamanhos das amostras dados aqui, qual é a potência do teste para essa alternativa bilateral?

f. Suponha que  $p_1 = 0,05$  e  $p_2 = 0,02$ . Determine o tamanho necessário da amostra para detectar essa diferença, com uma probabilidade de no mínimo 0,9.

**10.6.4** De uma amostra aleatória de 500 adultos residentes em uma vila, 385 foram favoráveis ao aumento do limite de velocidade para 75 mph em uma autoestrada, enquanto em outra amostra de 400 adultos, residentes em outra vila, 267 foram favoráveis a esse aumento do limite de velocidade.

a. Esses dados indicam que há uma diferença entre os residentes das duas vilas no tocante ao apoio para aumentar o limite de velocidade? Use  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?

b. Construa um intervalo de confiança de 95 % para a diferença das duas proporções. Forneça uma interpretação prática desse intervalo.

**10.6.5** A revista *New England Journal of Medicine* reportou um experimento para julgar a eficiência de cirurgia em homens diagnosticados com câncer de próstata. A metade aleatoriamente atribuída de 695 (347) homens no estudo teve cirurgias e 18 deles depois morreram de câncer de próstata comparados com 31 dos 348 que não tiveram cirurgia. Existe alguma evidência para sugerir que a cirurgia diminuiu a proporção daqueles que morreram de câncer de próstata?

**10.6.6** Considere os dados de velocidade limite em autoestradas introduzidos no Exercício 10.6.4. Encontre um IC de 99 % para a diferença das duas proporções, usando o procedimento alternativo de IC descrito nesta seção. Compare os comprimentos do IC do Exercício 10.6.4 com este. Discuta as causas possíveis de qualquer diferença que você observou.

## Exercícios Suplementares para o Capítulo 10

10.S7 Considere a saída computacional a seguir.

**Teste T para Duas Amostras e IC:**

Amostra	N	Média	DP	EP da Média
1	20	11,87	2,23	?
2	20	12,73	3,19	

0,71 Diferença =  $\mu_1 - \mu_2$   
 Estimativa para a diferença: -0,860  
 IC de 95 % para a diferença: (?, ?)  
 Teste T da diferença = 0 (vs  $\neq 0$ ):  
 Valor T = ? Valor P = ? GL = ?  
 Ambos usam o DP Combinado = ?

a. Preencha os valores que faltam. Você pode usar limites para o valor  $P$ .

b. Esse é um teste unilateral ou bilateral?

c. Quais serão suas conclusões, se  $\alpha = 0,05$ ? E se  $\alpha = 0,10$ ?

10.S8 Um especialista comprou 25 resistores de um vendedor 1 e 35 resistores de um vendedor 2. Cada resistência do resistor é medida com os seguintes resultados:

**Vendedor 1**

96,8	100,0	100,3	98,5	98,3	98,2
99,6	99,4	99,9	101,1	103,7	97,7
99,7	101,1	97,7	98,6	101,9	101,0
99,4	99,8	99,1	99,6	101,2	98,2
98,6					

**Vendedor 2**

106,8	106,8	104,7	104,7	108,0	102,2
103,2	103,7	106,8	105,1	104,0	106,2
102,6	100,3	104,0	107,0	104,3	105,8
104,0	106,3	102,2	102,8	104,2	103,4
104,6	103,5	106,3	109,2	107,2	105,4
106,4	106,8	104,1	107,1	107,7	

a. Qual é a suposição necessária à distribuição de modo a testar a afirmação de que a variância da resistência do produto do vendedor 1 não é significativamente diferente da variância da resistência do produto do vendedor 2? Faça um procedimento gráfico para verificar essa suposição.

b. Faça um procedimento estatístico apropriado de teste de hipóteses de modo a determinar se o especialista pode afirmar que a variância da resistência do produto do vendedor 1 é significativamente diferente daquela do produto do vendedor 2.

10.S9 O anúncio de um produto líquido usado em uma dieta afirma que o seu uso por um mês resulta em uma perda média de massa de no mínimo três libras. Oito indivíduos usaram o produto por um mês, e os dados referentes à perda média resultante são reportados a seguir. Use os procedimentos de teste de hipóteses para responder às questões a seguir.

a. Os dados confirmam a afirmação do fabricante do produto de dieta, com a probabilidade de um erro tipo I estabelecida em 0,05?

b. Os dados confirmam a afirmação do fabricante do produto de dieta, com a probabilidade de um erro tipo I estabelecida em 0,01?

c. Por meio de um esforço para aumentar as vendas, o produtor está considerando a possibilidade de mudar suas afirmações de “no mínimo três libras” para “no mínimo cinco libras”. Repita os itens (a) e (b) para testar essa nova afirmação.

Indivíduo	Massa Inicial (lb)	Massa Final (lb)
1	165	161
2	201	195
3	195	192
4	198	193
5	155	150
6	143	141
7	150	146
8	187	183

10.S10 O experimento, em 1954, da vacina Salk contra a poliomielite focou na sua eficiência de combate à paralisia. Notou-se que, sem um grupo de controle de crianças, não haveria uma base para avaliar a eficácia da vacina Salk. Então a vacina foi administrada a um grupo, e um placebo (visualmente idêntico à vacina, porém sem efeito algum) foi administrado a um segundo grupo. Por motivos éticos e por se suspeitar de que o conhecimento da administração da vacina afetaria os diagnósticos subsequentes, o experimento foi conduzido de maneira tal que a identidade das crianças não fosse revelada. Ou seja, nem os indivíduos e nem os administradores sabiam quem havia recebido a vacina e quem havia recebido o placebo. Os dados reais para esse experimento são apresentados a seguir.

Grupo do placebo:  $n = 201.299$ : 110 casos observados de pólio

Grupo da vacina:  $n = 200.745$ : 33 casos observados de pólio

a. Considere o procedimento de teste de hipóteses para determinar se a proporção de crianças nos dois grupos que contraíram paralisia é estatisticamente diferente. Use uma probabilidade de erro tipo I igual a 0,05.

b. Repita o item (a), usando uma probabilidade de erro tipo I igual a 0,01.

c. Compare suas conclusões dos itens (a) e (b) e explique por que elas são as mesmas ou por que são diferentes.

10.S11 Em 1990, uma amostra aleatória de 1500 telefones residenciais em Fênix mostrou que 387 dos números não tinham sido listados. No mesmo ano, uma amostra aleatória de 1200 telefones em Scottsdale mostrou que 310 dos números não tinham sido listados.

a. Encontre o intervalo de confiança de 95 % para a diferença das duas proporções e use esse intervalo de confiança para determinar se há, entre as duas cidades, uma diferença estatisticamente significativa nas proporções de números não listados.

b. Encontre o intervalo de confiança de 90 % para a diferença das duas proporções e use esse intervalo de confiança para determinar se há, entre as duas cidades, uma diferença estatisticamente significativa nas proporções de números não listados.

c. Suponha que todos os números na descrição do problema tenham sido dobrados. Isto é, 774 residentes de 3000 amostras de Fênix e 620 residentes de 2400 de Scottsdale tiveram números não listados. Repita os itens (a) e (b) e comente o efeito nos seus resultados, se o tamanho da amostra for aumentado sem variar as proporções.

**10.S12** Um fabricante de uma nova drágea para aliviar dor gostaria de demonstrar que seu produto trabalha duas vezes mais rápido que o produto do concorrente. Especificamente, ele gostaria de testar

$$H_0: \mu_1 = 2\mu_2$$

$$H_1: \mu_1 > 2\mu_2$$

em que  $\mu_1$  é o tempo médio de absorção do produto adversário e  $\mu_2$  é o tempo médio de absorção do produto novo. Considerando que as variâncias  $\sigma_1^2$  e  $\sigma_2^2$  são conhecidas, desenvolva um procedimento para testar essa hipótese.

**10.S13** Um estudo sobre economia de combustível foi conduzido em dois automóveis alemães: Mercedes e Volkswagen. Um veículo de cada marca foi selecionado, registrando-se a quantidade de milhas percorridas com dez tanques de combustível em cada carro. Esses dados (em milhas por galão) são mostrados a seguir:

Mercedes		Volkswagen	
24,7	24,9	41,7	42,8
24,8	24,6	42,3	42,4
24,9	23,9	41,6	39,9
24,7	24,9	39,5	40,8
24,5	24,8	41,9	29,6

a. Construa um gráfico de probabilidade normal para cada um dos conjuntos de dados. Com base nesses gráficos, é razoável considerar que eles sejam retirados de uma população normal?

b. Suponha ter sido determinado que o menor valor observado dos dados do Mercedes foi registrado erroneamente, sendo o valor correto igual a 24,6. Além disso, o menor valor observado dos dados do Volkswagen foi também registrado erroneamente, sendo o valor correto igual a 39,6. Construa novamente gráficos de probabilidade normal para cada um dos conjuntos de dados com os valores corretos. Com base nesses novos gráficos, é razoável considerar que eles sejam retirados de uma população normal?

c. Compare suas respostas dos itens (a) e (b) e comente o efeito dessas observações erradas na suposição de normalidade.

d. Usando os dados corretos do item (b) e um intervalo de confiança de 95 %, há evidência que confirme a afirmação de que a variabilidade no consumo de combustível por milha seja maior para Volkswagen do que para Mercedes?

e. Refaça o item (d) deste problema, usando um procedimento adequado de teste de hipóteses. Você obteve a mesma resposta que obteve originalmente? Por quê?

**10.S14** Suponha que estejamos testando  $H_0: \mu_1 = \mu_2$  versus  $H_1: \mu_1 \neq \mu_2$  e que planejemos usar amostras de mesmo tamanho, provenientes de duas populações. Ambas as populações são consideradas normais, com variâncias desconhecidas, porém iguais. Se usarmos  $\alpha = 0,05$  e se a média verdadeira for  $\mu_1 = \mu_2 + \sigma$ , qual será o tamanho da amostra que tem de ser usado para que a potência desse teste seja, no mínimo, 0,90?

**10.S15** Uma máquina de teste de dureza Rockwell pressiona uma ponteira em um corpo de prova, usando a profundidade da depressão

resultante para indicar a dureza. Duas ponteiras diferentes estão sendo comparadas para determinar se fornecem as mesmas leituras de dureza Rockwell na escala C. Nove corpos de prova são testados, com ambas as ponteiras sendo testadas em cada corpo de prova. Os dados são mostrados na tabela seguinte.

Corpo de Prova	Ponteira		Corpo de Prova	Ponteira	
	1	2		1	2
1	47	46	6	41	41
2	42	40	7	45	46
3	43	45	8	45	46
4	40	41	9	49	48
5	42	43			

a. Estabeleça quaisquer suposições necessárias para testar a afirmação de que ambas as ponteiras produzem as mesmas leituras de dureza Rockwell na escala C. Verifique essas suposições nos dados que você tem.

b. Aplique um método estatístico apropriado para determinar se os dados confirmam a afirmação de que a diferença nas leituras de dureza Rockwell na escala C das duas ponteiras é significativamente diferente de zero.

c. Suponha que, se as duas ponteiras diferirem nas leituras da dureza média por 1,0, a potência do teste deverá ser no mínimo 0,9. Para  $\alpha = 0,01$ , quantos corpos de prova devem ser usados no teste?

**10.S16** Um artigo publicado no *Journal of the Environmental Engineering Division* ["Distribution of Toxic Substances in Rivers" (1982, Vol. 108, pp. 639-649)] investigou a concentração de várias substâncias orgânicas hidrofóbicas no Rio Wolf, no Tennessee. Medidas de hexaclorobenzeno (HCB) em nanograma por litro são feitas em diferentes profundidades a jusante de um lixão abandonado. Dados para duas profundidades são mostrados a seguir.

Superfície: 3,74; 4,61; 4,00; 4,67; 4,87; 5,12; 4,52; 5,29; 5,74; 5,48

Fundo: 5,44; 6,88; 5,37; 5,44; 5,03; 6,48; 3,89; 5,85; 6,85; 7,16

a. Quais são as suposições requeridas para testar a afirmação de que a concentração média de HCB é a mesma em ambas as profundidades? Verifique essas suposições para as quais você tem a informação.

b. Aplique um procedimento apropriado para determinar se os dados confirmam a afirmação do item (a).

c. Suponha que a diferença verdadeira nas concentrações médias seja de 2,0 nanogramas por litro. Para  $\alpha = 0,05$ , qual é a potência de um teste estatístico para  $H_0: \mu_1 = \mu_2$  versus  $H_1: \mu_1 \neq \mu_2$ ?

d. Que tamanho de amostra seria requerido para detectar uma diferença de 1,0 nanograma por litro com  $\alpha = 0,05$ , se a potência tem de ser no mínimo 0,9?

**10.S17** Considere os dados dos números de telefone não listados do Exercício 10.S11. Encontre ICs de 95 % para a diferença nas proporções dos números de telefone não listados para as residências de Fênix e de Scottsdale, usando ambos os procedimentos descritos neste capítulo. Compare os comprimentos desses dois intervalos e comente qualquer diferença que você possa observar.

# Capítulo 11 Exercícios

## Exercícios para a Seção 11.2

**11.2.1** Diabetes e obesidade são preocupações sérias de saúde nos Estados Unidos e na maioria do mundo desenvolvido. A medida do teor de gordura de uma pessoa é uma maneira de monitorar o progresso de controle de peso; porém, medi-lo acuradamente exige equipamento caro de raios X ou uma piscina para submergir o indivíduo. Em vez disso, usa-se com frequência o índice de massa corporal (IMC) como uma aproximação da gordura corporal porque é mais fácil de medir:  $IMC = \text{massa (kg)}/[\text{altura(m)}]^2 = 703 \text{ massa (lb)}/[\text{altura(polégada)}]^2$ . Em um estudo com 250 homens da Universidade de Brigham Young, o IMC e a gordura corporal foram medidos. Pesquisadores encontram as seguintes estatísticas:

$$\begin{aligned} \sum_{i=1}^n x_i &= 6322,28 & \sum_{i=1}^n x_i^2 &= 162674,18 \\ \sum_{i=1}^n y_i &= 4757,90 & \sum_{i=1}^n y_i^2 &= 107679,27 \\ \sum_{i=1}^n x_i y_i &= 125471,10 \end{aligned}$$

- Calcule as estimativas de mínimos quadrados dos coeficientes angular e linear. Faça um gráfico da linha de regressão.
- Use a equação da linha ajustada para prever que gordura corporal seria observada, em média, para um homem com um IMC de 30.
- Suponha que a gordura corporal de um homem com um IMC de 25 seja 25 %. Encontre o resíduo para essa observação.
- O IMC igual a 25 no item (c) foi superestimado ou subestimado? Explique brevemente.

**11.2.2** Um artigo em *Technometrics*, de S. C. Narula e J. F. Wellington [“Prediction, Linear Regression, and a Minimum Sum of Relative Errors” (1977, Vol. 19(2), pp. 185-190)], apresenta dados de preços de vendas e taxas anuais para 24 casas. Esses dados são mostrados na seguinte tabela:

Preço de Venda/1000	Taxas (local, escola, município)/1000	Preço de Venda/1000	Taxas (local, escola, município)/1000
25,9	4,9176	30,0	5,0500
29,5	5,0208	36,9	8,2464
27,9	4,5429	41,9	6,6969
25,9	4,5573	40,5	7,7841
29,9	5,0597	43,9	9,0384
29,9	3,8910	37,5	5,9894
30,9	5,8980	37,9	7,5422
28,9	5,6039	44,5	8,7951
35,9	5,8282	37,9	6,0831
31,5	5,3003	38,9	8,3607
31,0	6,2712	36,9	8,1400
30,9	5,9592	45,8	9,1416

- Considerando que um modelo de regressão linear simples seja apropriado, obtenha um ajuste por mínimos quadrados, relacionando o preço de venda com as taxas pagas. Qual é a estimativa de  $\sigma^2$ ?
- Encontre o preço médio de venda, dado que as taxas pagas são  $x = 7,50$ .
- Calcule o valor ajustado de  $y$  correspondendo a  $x = 5,8980$ . Encontre o resíduo correspondente.
- Calcule o valor ajustado  $\hat{y}_i$  para cada valor de  $x_i$  usado para ajustar o modelo. Construa então um gráfico de  $\hat{y}_i$  contra o valor observado correspondente,  $y_i$ , e comente o que esse gráfico aparentaria se a relação entre  $y$  e  $x$  fosse uma linha reta determinística (sem erro aleatório). O gráfico obtido indica realmente que as taxas pagas representam uma variável efetiva de regressão na previsão do preço de venda?

**11.2.3** Um artigo em *Concrete Research* [“Near Surface Characteristics of Concrete: Intrinsic Permeability” (1989, Vol. 41, pp. 87-97)] apresentou dados sobre resistência à compressão,  $x$ , e permeabilidade intrínseca,  $y$ , de várias misturas e curas de concreto. Um sumário das grandezas é:  $n = 14$ ,  $\sum y_i = 572$ ,  $\sum y_i^2 = 23.530$ ,  $\sum x_i = 43$ ,  $\sum x_i^2 = 157,42$  e  $\sum x_i y_i = 1697,80$ . Considere que as duas variáveis estejam relacionadas por um modelo de regressão linear simples.

- Calcule as estimativas de mínimos quadrados dos coeficientes angular e linear. Estime  $\sigma^2$ . Faça um gráfico da linha de regressão.
- Use a equação da linha ajustada para prever que valor de permeabilidade seria observado, se a resistência à compressão fosse  $x = 4,3$ .
- Dê uma estimativa da permeabilidade média, quando a resistência à compressão for  $x = 3,7$ .
- Suponha que o valor observado da permeabilidade em  $x = 3,7$  seja  $y = 46,1$ . Calcule o valor do resíduo correspondente.

**11.2.4** Um artigo publicado no *Journal of Sound and Vibration* [“Measurement of Noise-Evoked Blood Pressure by Means of Averaging Method: Relation between Blood Pressure Rise and SPL” (1991, Vol. 151(3), pp. 383-394)] descreveu um estudo investigativo da relação entre exposição ao barulho e hipertensão. Os seguintes dados são representativos daqueles reportados no artigo.

$y$	1	0	1	2	5	1	4	6	2	3
$x$	60	63	65	70	70	70	80	90	80	80
$y$	5	4	6	8	4	5	7	9	7	6
$x$	85	89	90	90	90	90	94	100	100	100

- Desenhe um diagrama de dispersão de  $y$  (aumento da pressão sanguínea em milímetros de mercúrio) versus  $x$  (nível da pressão sonora em decibéis). Parece razoável propor um modelo de regressão linear simples para essa situação?
- Ajuste o modelo de regressão linear simples usando o método dos mínimos quadrados. Encontre uma estimativa de  $\sigma^2$ .
- Encontre o aumento médio previsto para a pressão sanguínea, associado a um nível de pressão sonora de 85 decibéis.

**11.2.5** Um artigo publicado no *Tappi Journal* (março de 1986) apresentou dados sobre a concentração (em gramas por litro) do licor verde de  $Na_2S$  e da produção (toneladas por dia) de uma máquina de papel. Os dados (lidos a partir de um gráfico) são mostrados na seguinte tabela:

y	40	42	49	46	44	48
x	825	830	890	895	890	910
y	46	43	53	52	54	57
x	915	960	990	1010	1012	1030

a. Ajuste um modelo de regressão linear simples, relacionando a concentração do licor verde de  $\text{Na}_2\text{S}$ ,  $y$ , com a produção,  $x$ . Encontre uma estimativa de  $\sigma^2$ . Desenhe um diagrama de dispersão dos dados e do modelo resultante do ajuste pelo método dos mínimos quadrados.

b. Encontre o valor ajustado  $y$ , correspondente a  $x = 910$ , e o resíduo associado.

c. Encontre a concentração média do licor verde de  $\text{Na}_2\text{S}$ , quando a taxa de produção for de 950 toneladas por dia.

**11.2.6** Um artigo publicado no *Journal of Environmental Engineering* (1989, Vol. 115(3), pp. 608-619) relatou os resultados de um estudo a respeito da ocorrência de sódio e cloreto nas correntes superficiais de um rio na parte central de Rhode Island. Os dados a seguir se referem à concentração de cloreto (em miligramas por litro),  $y$ , e à área (em %) das encostas exploradas para análise,  $x$ .

y	4,4	6,6	9,7	10,6	10,8	10,9
x	0,19	0,15	0,57	0,70	0,67	0,63
y	11,8	12,1	14,3	14,7	15,0	17,3
x	0,47	0,70	0,60	0,78	0,81	0,78
y	19,2	23,1	27,4	27,7	31,8	39,5
x	0,69	1,30	1,05	1,06	1,74	1,62

a. Desenhe um diagrama de dispersão dos dados. Um modelo de regressão linear simples parece apropriado aqui?

b. Ajuste um modelo de regressão linear simples usando o método dos mínimos quadrados. Encontre uma estimativa de  $\sigma^2$ .

c. Estime a concentração média de cloreto para 1 % de área das encostas exploradas.

d. Encontre o valor ajustado correspondente a  $x = 0,47$  e o resíduo associado.

**11.2.7** Um artigo publicado no *Journal of the American Ceramic Society* ["Rapid Hot-Pressing of Ultrafine PSZ Powders" (1991, Vol. 74, pp. 1547-1553)] considerou a microestrutura de pó ultrafino de zircônia parcialmente estabilizada como uma função da temperatura. Os dados são mostrados a seguir:

x = Temperatura (°C):	1100	1200	1300	1100	1500
		1200	1300		
y = Porosidade (%):	30,8	19,2	6,0	13,5	11,4
		7,7	3,6		

a. Ajuste um modelo de regressão linear simples, usando o método dos mínimos quadrados. Encontre uma estimativa para  $\sigma^2$ .

b. Estime a porosidade média para uma temperatura de 1400 °C.

c. Encontre o valor ajustado correspondente a  $y = 11,4$  e o resíduo associado.

d. Desenhe um diagrama de dispersão dos dados. Um modelo de regressão linear simples parece apropriado aqui? Explique.

**11.2.8** Um artigo em *Wood Science and Technology* ["Creep in Chipboard, Part 3: Initial Assessment of the Influence of Moisture Content and Level of Stressing on Rate of Creep and Time to Failure" (1981, Vol. 15, pp. 125-144)] estudou a deflexão (mm) de papelão a partir de níveis de umidade relativa. Considere que as duas variáveis estejam relacionadas de acordo com o modelo de regressão linear simples. Os dados são mostrados a seguir:

x = Nível de umidade (%):	54	54	61	61	68
y = Deflexão (mm):	16.473	18.693	14.305	15.121	13.505
x = Nível de umidade (%):	68	75	75	75	
y = Deflexão (mm):	11.640	11.168	12.534	11.224	

a. Calcule as estimativas de mínimos quadrados dos coeficientes angular e linear. Qual é a estimativa de  $\sigma^2$ ? Faça um gráfico do modelo de regressão com os dados.

b. Encontre a estimativa da deflexão média, se a umidade relativa puder ser limitada para 65 %.

c. Estime a mudança na deflexão média associada a um aumento de 5 % na umidade relativa.

d. De modo a diminuir a deflexão média em 1 milímetro, qual o aumento que tem de ser gerado na umidade relativa?

e. Dado que a umidade relativa é de 68 %, encontre o valor ajustado da deflexão e o resíduo correspondente.

**11.2.9** Um artigo publicado no *Journal of Environmental Engineering Division* ["Least Squares Estimates of BOD Parameters" (1980, Vol. 106, pp. 1197-1202)] tomou uma amostra do Rio Holston abaixo de Kingport, Tennessee, durante o mês de agosto de 1977. O teste de demanda bioquímica de oxigênio (DBO) foi conduzido durante um período de tempo dado em dias. Os dados resultantes são mostrados a seguir:

Tempo (dias):	1	2	4	6	8	10	12	14	16
		18	20						
DBO (mg/litro):	0,6	0,7	1,5	1,9	2,1	2,6	2,9	3,7	3,5
		3,7	3,8						

a. Considerando que um modelo de regressão linear seja apropriado, ajuste o modelo de regressão relacionando DBO ( $y$ ) com o tempo ( $x$ ). Qual é a estimativa de  $\sigma^2$ ?

b. Qual é a estimativa do nível esperado de DBO para um tempo de 15 dias?

c. Que variação no DBO médio é esperada quando o tempo varia por 3 dias?

d. Suponha que o tempo usado seja de 6 dias. Calcule o valor ajustado de  $y$  e o resíduo correspondente.

e. Calcule  $\hat{y}_i$  ajustado para cada valor de  $x_i$  usado para ajustar o modelo. Construa então um gráfico de  $\hat{y}_i$  versus os valores observados correspondentes  $y_i$  e comente o que esse gráfico pareceria se a relação entre  $y$  e  $x$  fosse uma linha reta determinística (nenhum erro aleatório). O gráfico obtido realmente indica que o tempo é uma variável regressora efetiva na previsão de DBO?

**11.2.10** Mostre que, em um modelo de regressão linear simples, o ponto  $(\bar{x}, \bar{y})$  está exatamente sobre a linha de regressão dos mínimos quadrados.

Exercícios para a Seção 11.4

11.4.1 Lembre-se da regressão do percentual de gordura corporal sobre o IMC, proveniente do Exercício 11.2.1.

- a. Estime o desvio-padrão do erro.
- b. Estime o desvio-padrão do coeficiente angular.
- c. Qual é o valor da estatística  $t$  para o coeficiente angular?
- d. Teste a hipótese que  $\beta_1 = 0$ , com  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?

11.4.2 Considere o modelo de regressão linear simples  $y = 10 + 25x + \epsilon$ , em que o termo do erro aleatório é normal e independentemente distribuído com média zero e desvio-padrão 2. Use um *software* para gerar uma amostra de oito observações, uma em cada nível  $x = 10, 12, 14, 16, 18, 20, 22$  e  $24$ .

- a. Ajuste o modelo de regressão linear por mínimos quadrados e encontre as estimativas dos coeficientes angular e linear.
- b. Encontre a estimativa de  $\sigma^2$ .
- c. Encontre os erros-padrão dos coeficientes angular e linear.
- d. Agora, use o *software* para gerar uma amostra de 16 observações, sendo cada duas nos mesmos níveis de  $x$  usados previamente. Ajuste o modelo usando os mínimos quadrados.
- e. Encontre a estimativa de  $\sigma^2$  para o novo modelo do item (d). Compare-o com a estimativa obtida no item (b). Qual o impacto que o aumento no tamanho da amostra teve sobre a estimativa?
- f. Encontre os erros-padrão dos coeficientes angular e linear, usando o novo modelo do item (d). Compare esses erros-padrão com aqueles encontrados no item (c). Qual o impacto que o aumento no tamanho da amostra teve sobre os erros-padrão estimados?

11.4.3 Considere a saída computacional a seguir.

A equação de regressão é

$$Y = 12,9 + 2,34 x$$

Preditor	Coef.	EP do Coef.	T	P
Constante	12,857	1,032	?	?
X	2,3445	0,1150	?	?
$S = 1,48111 \quad R^2 = 98,1\% \quad R^2_{\text{ajustado}} = 97,9\%$				

Análise de Variância

Fonte	GL	SQ	MQ	F	P
Regressão	1	912,43	912,43	?	?
Erro residual	8	17,55	?		
Total	9	929,98			

- a. Preencha a informação que falta. Você pode usar limites para os valores  $P$ .
- b. Você pode concluir que o modelo define uma relação linear útil?
- c. Qual é a sua estimativa de  $\sigma^2$ ?

11.4.4 Considere os dados do Exercício 11.2.2 sobre  $y =$  preço de venda e  $x =$  taxas pagas.

- a. Teste  $H_0: \beta_1 = 0$ , usando o teste  $t$ ; use  $\alpha = 0,05$ .
- b. Teste  $H_0: \beta_1 = 0$ , usando a análise de variância com  $\alpha = 0,05$ . Discuta a relação desse teste com o teste do item (a).
- c. Estime os erros-padrão dos coeficientes angular e linear.
- d. Teste a hipótese de  $\beta_0 = 0$ .

11.4.5 Considere os dados do Exercício 11.2.3 sobre  $x =$  resistência à compressão e  $y =$  permeabilidade intrínseca de concreto.

- a. Teste a significância da regressão, usando  $\alpha = 0,05$ . Encontre o valor  $P$  para esse teste. Você pode concluir que o modelo específica uma relação linear útil entre essas duas variáveis?
- b. Estime  $\sigma^2$  e o desvio-padrão de  $\hat{\beta}_1$ .
- c. Qual é o erro-padrão do coeficiente linear desse modelo?

11.4.6 Considere os dados do Exercício 11.2.4 sobre  $y =$  aumento da pressão sanguínea e  $x =$  nível de pressão sonora.

- a. Teste a significância da regressão, usando  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?
- b. Estime os erros-padrão dos coeficientes angular e linear.
- c. Teste  $H_0: \beta_0 = 0$  versus  $H_1: \beta_0 \neq 0$ , usando  $\alpha = 0,05$ . Encontre o valor  $P$  para esse teste.

11.4.7 Considere os dados do Exercício 11.2.5 sobre  $y =$  concentração do licor verde de  $\text{Na}_2\text{S}$  e  $x =$  produção em uma máquina de papel.

- a. Teste a significância da regressão, usando  $\alpha = 0,05$ . Encontre o valor  $P$  para esse teste.
- b. Estime os erros-padrão dos coeficientes angular e linear.
- c. Teste  $H_0: \beta_0 = 0$  versus  $H_1: \beta_0 \neq 0$ , usando  $\alpha = 0,05$ . Encontre o valor  $P$  para esse teste.

11.4.8 Considere os dados do Exercício 11.2.6 sobre  $y =$  concentração de cloreto e  $x =$  área das encostas.

- a. Teste a hipótese  $H_0: \beta_1 = 0$  versus  $H_1: \beta_1 \neq 0$ , usando o procedimento da análise de variância, com  $\alpha = 0,01$ .
- b. Encontre o valor  $P$  para o teste no item (a).
- c. Estime os erros-padrão de  $\hat{\beta}_0$  e  $\hat{\beta}_1$ ?
- d. Teste  $H_0: \beta_1 = 0$  versus  $H_1: \beta_1 \neq 0$ , usando  $\alpha = 0,01$ . Que conclusões você pode tirar? Você acha que o modelo poderia ter melhor ajuste se a interseção fosse removida?

11.4.9 Considere os dados do Exercício 11.2.9 sobre  $y =$  demanda de oxigênio e  $x =$  tempo.

- a. Teste a significância da regressão com  $\alpha = 0,01$ . Encontre o valor  $P$  para esse teste. Que conclusões você pode tirar?
- b. Estime os erros-padrão dos coeficientes angular e linear.
- c. Teste a hipótese de que  $\beta_0 = 0$ .

11.4.10 Considere os dados do Exercício 11.2.8 sobre  $y =$  deflexão e  $x =$  nível de umidade relativa.

- a. Teste a significância da regressão com  $\alpha = 0,01$ . Qual é o valor  $P$  para esse teste? Estabeleça as conclusões que resultam desse teste.
- b. Esse modelo parece adequado?
- c. Estime os erros-padrão dos coeficientes angular e linear.

11.4.11 Um artigo publicado no *The Journal of Clinical Endocrinology and Metabolism* ["Simultaneous and Continuous 24-Hour Plasma and Cerebrospinal Fluid Leptin Measurements: Dissociation of Concentrations in Central and Peripheral Compartments" (2004, Vol. 89, pp. 258-265)] estudou a demografia de plasma simultâneo e contínuo 24 horas e medidas de leptina no fluido cerebrospinal. Os dados são mostrados a seguir:

$y =$ IMC ( $\text{kg}/\text{m}^2$ ):	19,92	20,59	29,02	20,78	25,97
	20,39	23,29	17,27	35,24	
$x =$ Idade (ano):	45,5	34,6	40,6	32,9	28,2
	52,1	33,3	47,0		

a. Teste a significância da regressão com  $\alpha = 0,05$ . Encontre o valor  $P$  para esse teste. Você pode concluir que o modelo especifica uma relação linear útil entre essas duas variáveis?

b. Estime  $\sigma^2$  e o desvio-padrão de  $\hat{\beta}_1$ .  
c. Qual é o erro-padrão do coeficiente linear desse modelo?

## Exercícios para as Seções 11.5 e 11.6

**11.6.1** Usando a regressão do Exercício 11.2.1,

- Encontre um intervalo de confiança de 95 % para o coeficiente angular.
- Encontre um intervalo de confiança de 95 % para o percentual médio de gordura corporal para um homem com um IMC de 25.
- Encontre um intervalo de previsão de 95 % para o percentual médio de gordura corporal para um homem com um IMC de 25.
- Qual intervalo é maior: o intervalo de confiança ou o intervalo de previsão? Explique brevemente.

**11.6.2** Considere os dados do Exercício 11.2.2 sobre  $y =$  preço de venda de casas e  $x =$  taxas pagas. Encontre um intervalo de confiança de 95 % para cada um dos itens:

- $\beta_1$
- $\beta_0$
- Preço médio de venda quando as taxas pagas forem  $x = 7,50$ .
- Calcule o intervalo de previsão de 95 % para o preço de venda quando as taxas pagas forem  $x = 7,50$ .

**11.6.3** Considere os dados no Exercício 11.2.3 sobre  $y =$  permeabilidade intrínseca do concreto e  $x =$  resistência à compressão. Encontre um intervalo de confiança de 95 % para cada um dos itens:

- Coefficiente angular.
- Coefficiente linear.
- Permeabilidade média quando  $x = 2,5$ .
- Encontre um intervalo de previsão de 95 % para a permeabilidade quando  $x = 2,5$ . Explique por que esse intervalo é mais largo do que o intervalo no item (c).

**11.6.4** O Exercício 11.2.4 apresentou os dados sobre  $y =$  aumento da pressão sanguínea e  $x =$  o nível de pressão sonora. Encontre um intervalo de confiança de 95 % para cada um dos itens:

- $\beta_1$
- $\beta_0$
- Aumento médio da pressão sanguínea quando o nível de pressão sonora for 85 decibéis.
- Encontre o intervalo de previsão de 95 % para o aumento médio da pressão sanguínea quando a pressão sonora for 85 decibéis.

**11.6.5** Considere os dados do Exercício 11.2.5 sobre  $y =$  concentração do licor verde de  $\text{Na}_2\text{S}$  e  $x =$  produção em uma máquina de papel. Encontre um intervalo de confiança de 99 % para cada um dos itens:

- $\beta_1$
- $\beta_0$
- Concentração média de  $\text{Na}_2\text{S}$  quando a produção for  $x = 910$  t/dia.
- Encontre o intervalo de previsão de 99 % para a concentração de  $\text{Na}_2\text{S}$  quando a produção for  $x = 910$  t/dia.

**11.6.6** O Exercício 11.2.6 apresentou os dados sobre concentração de cloreto  $y$  e o percentual de área  $x$  nas encostas dos rios da parte central de Rhode Island. Encontre um intervalo de confiança de 99 % para cada um dos itens:

- $\beta_1$
- $\beta_0$
- Concentração média de cloreto quando o percentual de área analisada for  $x = 1,0$  %.
- Encontre o intervalo de previsão de 99 % para a concentração média de cloreto quando o percentual de área analisada for  $x = 1,0$  %.

**11.6.7** Considere os dados do Exercício 11.2.9 sobre a demanda de oxigênio. Encontre um intervalo de confiança de 99 % para cada um dos itens:

- $\beta_1$
- $\beta_0$
- Encontre o intervalo de confiança de 95 % para a DBO média quando o tempo for igual a oito dias.

**11.6.8** Considere os dados do Exercício 11.2.7 sobre a microestrutura de zircônia. Encontre um intervalo de confiança de 95 % para cada um dos itens:

- Coefficiente angular.
- Coefficiente linear.
- Comprimento médio quando  $x = 1500$ .
- Encontre o intervalo de previsão de 95 % para o comprimento médio quando  $x = 1500$ . Explique por que o intervalo é mais largo do que o intervalo do item (c).

## Exercícios para a Seção 11.7

**11.7.1** Repita o Exercício 11.7.2 usando um termo de erro com um desvio-padrão de 4. Qual o impacto que o aumento do erro do desvio-padrão teve nos valores de  $R^2$ ?

**11.7.2** Considere o modelo de regressão linear simples  $y = 10 + 0,3x + \epsilon$ , em que o termo de erro aleatório é normal e independentemente distribuído, com média zero e desvio-padrão 1. Use um *software* para gerar uma amostra de oito observações, um em cada nível  $x = 10, 12, 14, 16, 18, 20, 22$  e 24.

- Ajuste o modelo de regressão linear por mínimos quadrados e encontre as estimativas dos coeficientes angular e linear.
- Encontre a estimativa de  $\sigma^2$ .

c. Encontre o valor de  $R^2$ .

d. Agora use o *software* para gerar uma amostra de oito observações, sendo cada uma nos níveis de  $x = 10, 14, 18, 22, 26, 30, 34$  e 38. Ajuste o modelo usando os mínimos quadrados.

e. Encontre  $R^2$  para o novo modelo do item (d). Compare esse valor com aquele obtido no item (c). Qual o impacto que o aumento na dispersão do preditor  $x$  teve sobre o valor?

**11.7.3** Consulte os dados de resistência à compressão do Exercício 11.2.3. Use o sumário das grandezas fornecidas para calcular  $R^2$  e forneça uma interpretação prática dessa grandeza.

11.7.4 Considere os dados do Exercício 11.2.2 sobre o preço de venda de casas  $y$  e taxas pagas  $x$ .

- Encontre os resíduos para o modelo de mínimos quadrados.
- Prepare um gráfico de probabilidade normal dos resíduos e interprete esse gráfico.
- Plote os resíduos contra  $\hat{y}$  e *versus*  $x$ . A suposição de variância constante parece ser satisfeita?
- Que proporção da variabilidade total é explicada pelo modelo de regressão?

11.7.5 Consulte o Exercício 11.2.6, que apresenta dados sobre concentração de cloreto  $y$  e área das encostas  $x$ .

- Que proporção da variabilidade total na concentração de cloreto é explicada pelo modelo de regressão?
- Plote os resíduos *versus*  $\hat{y}$  e *versus*  $x$ . Interprete esses gráficos.
- Prepare um gráfico de probabilidade normal dos resíduos. A suposição de normalidade parece ser satisfeita?

11.7.6 Considere o Exercício 11.2.4, que apresenta dados sobre o aumento da pressão sanguínea  $y$  e o nível da pressão sonora  $x$ .

- Que proporção da variabilidade total do aumento da pressão sanguínea é explicada pelo nível da pressão sonora?
- Prepare um gráfico de probabilidade normal dos resíduos provenientes do modelo de mínimos quadrados. Interprete esse gráfico.
- Plote os resíduos *versus*  $\hat{y}$  e *versus*  $x$ . Comente esses gráficos.

11.7.7 Um artigo publicado no *Journal of the American Statistical Association* ["Markov Chain Monte Carlo Methods for Computing Bayes Factors: A Comparative Review" (2001, Vol. 96, pp. 1122-1132)] analisou os dados tabelados sobre a resistência à compressão paralela à semente *versus* a densidade ajustada de resina para os espécimes de pinheiro radiata. Os dados são mostrados a seguir.

Resistência à Compressão	Densidade	Resistência à Compressão	Densidade
3040	29,2	3840	30,7
2470	24,7	3800	32,7
3610	32,3	4600	32,6
3480	31,3	1900	22,1
3810	31,5	2530	25,3
2330	24,5	2920	30,8
1800	19,9	4990	38,9
3110	27,3	1670	22,1
3160	27,1	3310	29,2
2310	24,0	3450	30,1
4360	33,8	3600	31,4
1880	21,5	2850	26,7
3670	32,2	1590	22,1
1740	22,5	3770	30,3
2250	27,5	3850	32,0
2650	25,6	2480	23,2
4970	34,5	3570	30,3
2620	26,2	2620	29,9
2900	26,7	1890	20,8
1670	21,1	3030	33,2
2540	24,1	3030	28,2

- Ajuste um modelo de regressão, relacionando a resistência à compressão com a densidade.
- Teste a significância da regressão com  $\alpha = 0,05$ .

c. Estime  $\sigma^2$  para esse modelo.

d. Calcule  $R^2$  para esse modelo. Forneça uma interpretação prática dessa grandeza.

e. Prepare e interprete o gráfico de probabilidade normal dos resíduos.

f. Plote os resíduos *versus*  $\hat{y}$  e *versus*  $x$ . A suposição de variância constante parece ser satisfeita?

11.7.8 Considere os dados do Exercício 11.2.5 sobre  $y =$  concentração do licor verde de  $\text{Na}_2\text{S}$  e  $x =$  produção da máquina de papel. Suponha que o 14º ponto seja adicionado aos dados originais, em que  $y_{14} = 59$  e  $x_{14} = 855$ .

a. Prepare um diagrama de dispersão de  $y$  *versus*  $x$ . Ajuste o modelo de regressão linear simples às 14 observações.

b. Teste a significância da regressão com  $\alpha = 0,05$ .

c. Estime  $\sigma^2$  para esse modelo.

d. Compare a estimativa de  $\sigma^2$  obtida no item (c) com a estimativa de  $\sigma^2$  obtida a partir dos 13 pontos originais. Qual é a maior estimativa e por quê?

e. Calcule os resíduos para esse modelo. O valor de  $e_{14}$  parece não usual?

f. Prepare e interprete o gráfico de probabilidade normal dos resíduos.

g. Plote os resíduos *versus*  $\hat{y}$  e *versus*  $x$ . Comente esses gráficos.

11.7.9 Mostre que uma maneira equivalente de definir o teste para significância de regressão em uma regressão linear simples é se basear no teste de  $R^2$ , como segue: para testar  $H_0: \beta_1 = 0$  *versus*  $H_0: \beta_1 \neq 0$ , calcule

$$F_0 = \frac{R^2(n-2)}{1-R^2}$$

e rejeite  $H_0: \beta_1 = 0$  se o valor calculado  $f_0 > f_{\alpha,1,n-2}$ . Suponha que o modelo de regressão linear simples tenha sido ajustado para  $n = 25$  observações e  $R^2 = 0,90$ .

a. Teste a significância da regressão para  $\alpha = 0,05$ .

b. Qual é o menor valor de  $R^2$  que levaria à conclusão de uma regressão significativa se  $\alpha = 0,05$ ?

11.7.10 Resíduos na Forma de Student. Mostre que a variância do  $i$ -ésimo resíduo é

$$V(e_i) = \sigma^2 \left[ 1 - \left( \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right) \right]$$

Sugestão:

$$\text{cov}(Y_i, \hat{Y}_i) = \sigma^2 \left[ \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right].$$

O  $i$ -ésimo resíduo na forma de Student é definido como

$$r_i = \frac{e_i}{\sqrt{\hat{\sigma}^2 \left[ 1 - \left( \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right) \right]}}$$

a. Explique por que  $r_i$  tem unidade de desvio-padrão.

b. Os resíduos padronizados têm unidade de desvio-padrão?

c. Discuta o comportamento do resíduo padronizado quando o valor da amostra  $x_i$  estiver muito próximo do meio da faixa de  $x$ .

d. Discuta o comportamento do resíduo na forma de Student quando o valor  $x_i$  da amostra estiver muito próximo de uma extremidade da faixa de  $x$ .

## Exercícios para a Seção 11.8

11.8.1 Suponha que dados sejam obtidos a partir de 20 pares de  $(x, y)$  e que o coeficiente de correlação para a amostra seja 0,8.

- Teste a hipótese que  $H_0: \rho = 0$  contra  $H_1: \rho \neq 0$ , com  $\alpha = 0,05$ ? Calcule o valor  $P$ .
- Teste a hipótese de que  $H_0: \rho = 0,5$  contra  $H_1: \rho \neq 0,5$ , com  $\alpha = 0,05$ . Calcule o valor  $P$ .
- Construa um intervalo bilateral de confiança de 95 % para o coeficiente de correlação. Explique como as questões nos itens (a) e (b) poderiam ser respondidas com um intervalo de confiança.

11.8.2 Uma amostra aleatória de  $n = 25$  observações foi obtida relativa ao tempo de falha de um componente eletrônico e à temperatura do ambiente em que o componente foi usado.

- Dado que  $\hat{\rho} = 0,83$ , teste a hipótese de que  $\rho = 0$ , usando  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?
- Encontre um intervalo de confiança de 95 % para  $\rho$ .
- Teste a hipótese  $H_0: \rho = 0,8$  versus  $H_1: \rho \neq 0,8$ , usando  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?

11.8.3 Os dados da tabela a seguir expressam  $x =$  o conteúdo de água na neve, em 1º de abril, e  $y =$  o campo (em polegadas), de abril a julho, nas encostas do *Snake River* em Wyoming, de 1919 a 1935. (Os dados foram retirados de um artigo em *Research Notes*, Vol. 61, 1950, Estação Experimental da Floresta Situada na Faixa Noroeste do Pacífico – *Pacific Northwest Forest Range Experiment Station*, Oregon.)

$x$	$y$	$x$	$y$
23,1	10,5	37,9	22,8
32,8	16,7	30,5	14,1
31,8	18,2	25,1	12,9
32,0	17,0	12,4	8,8
30,4	16,3	35,1	17,4
24,0	10,5	31,5	14,9
39,5	23,1	21,1	10,5
24,2	12,4	27,6	16,1
52,5	24,9		

- Estime a correlação entre  $Y$  e  $X$ .
- Teste a hipótese de que  $\rho = 0$ , usando  $\alpha = 0,05$ .
- Ajuste um modelo de regressão linear simples e teste a significância da regressão usando  $\alpha = 0,05$ . Que conclusões você pode tirar? Como está relacionado o teste para a significância da regressão com o teste para  $\rho$  no item(b)?
- Análise os resíduos e comente a adequação do modelo.

11.8.4 Os dados referentes ao peso e à pressão sanguínea sistólica de 26 homens selecionados aleatoriamente, na faixa etária de 25 a 30 anos, são mostrados na tabela seguinte. Considere que o peso e a pressão sanguínea estejam distribuídos normal e conjuntamente.

Indivíduo	Peso	PS Sistólica	Indivíduo	Peso	PS Sistólica
1	165	130	14	172	153
2	167	133	15	159	128
3	180	150	16	168	132
4	155	128	17	174	149
5	212	151	18	183	158

Indivíduo	Peso	PS Sistólica	Indivíduo	Peso	PS Sistólica
6	175	146	19	215	150
7	190	150	20	195	163
8	210	140	21	180	156
9	200	148	22	143	124
10	149	125	23	240	170
11	158	133	24	235	165
12	169	135	25	192	160
13	170	150	26	187	159

- Encontre a linha de regressão relacionando a pressão sanguínea sistólica com o peso.
- Teste a significância da regressão, usando  $\alpha = 0,05$ .
- Estime o coeficiente de correlação.
- Teste a hipótese de que  $\rho = 0$ , usando  $\alpha = 0,05$ .
- Teste a hipótese de que  $\rho = 0,5$ , usando  $\alpha = 0,05$ .
- Construa um intervalo de confiança de 95 % para o coeficiente de correlação.

11.8.5 Em um artigo publicado em *IEEE Transactions on Instrumentation and Measurement* (2001, Vol. 50, pp. 986-990), pesquisadores estudaram os efeitos de reduzir, por meios eletrônicos, a perda de corrente em um núcleo magnético. Realizando um experimento emparelhado, eles mediram a corrente em um enrolamento magnético com e sem a eletrônica. Os dados para o caso sem eletrônicos são apresentados na tabela a seguir.

Suprimento de Voltagem	Corrente sem Eletrônicos (mA)
0,66	7,32
1,32	12,22
1,98	16,34
2,64	23,66
3,3	28,06
3,96	33,39
4,62	34,12
3,28	39,21
5,94	44,21
6,6	47,48

- Faça um gráfico e ajuste uma linha de regressão para prever a corrente sem eletrônicos em função do suprimento de voltagem. Há uma regressão significativa para  $\alpha = 0,05$ ? Qual é o valor  $P$ ?
- Estime o coeficiente de correlação.
- Teste a hipótese de que  $\rho = 0$  contra a alternativa  $\rho \neq 0$ , usando  $\alpha = 0,05$ . Qual é o valor  $P$ ?
- Calcule um intervalo de confiança de 95 % para o coeficiente de correlação.

11.8.6 A estimativa absoluta mensal dos índices globais (terra e mar combinados) da temperatura (graus °C) em 2000 e 2001 são ([www.ncdc.noaa.gov/oa/climate/](http://www.ncdc.noaa.gov/oa/climate/)):

2000: 12,28; 12,63; 13,22; 14,21; 15,13; 15,82; 16,05; 16,02; 15,29; 14,29; 13,16; 12,47  
 2001: 12,44; 12,55; 13,35; 14,22; 15,28; 15,99; 16,23; 16,17; 15,44; 14,52; 13,52; 12,61

- a. Faça um gráfico e ajuste uma linha de regressão para prever as temperaturas de 2001 a partir daquelas de 2000. Há uma regressão significativa para  $\alpha = 0,05$ ? Qual é o valor  $P$ ?
- b. Estime o coeficiente de correlação.

- c. Teste a hipótese de que  $\rho = 0,9$  contra a alternativa  $\rho \neq 0,9$ , usando  $\alpha = 0,05$ . Qual é o valor  $P$ ?
- d. Calcule um intervalo de confiança de 95 % para o coeficiente de correlação.

## Exercícios para a Seção 11.9

**11.9.1** Determine se os seguintes modelos são intrinsecamente lineares. Se afirmativo, determine a transformação apropriada para gerar o modelo linear.

- a.  $Y = \beta_0 x^{\beta_1} \varepsilon$
- b.  $Y = \frac{3 + 5x}{x} + \varepsilon$
- c.  $Y = \beta_0 \beta_1^x \varepsilon$
- d.  $Y = \frac{x}{\beta_0 x + \beta_1 + x\varepsilon}$

**11.9.2** A pressão de vapor d'água em várias temperaturas é dada na tabela a seguir:

Número da Observação	Temperatura (K)	Pressão de Vapor (mm Hg)
1	273	4,6
2	283	9,2
3	293	17,5
4	303	31,8
5	313	55,3
6	323	92,5
7	333	149,4
8	343	233,7
9	353	355,1
10	363	525,8
11	373	760,0

- a. Desenhe um diagrama de dispersão desses dados. Que tipo de relação parece apropriado relacionando  $y$  com  $x$ ?
- b. Ajuste um modelo de regressão linear a esses dados.
- c. Teste a significância da regressão, usando  $\alpha = 0,05$ . Que conclusões você pode tirar?
- d. Faça um gráfico dos resíduos do modelo de regressão *versus*  $\hat{y}_i$ . O que você conclui acerca da adequação do modelo?
- e. A relação de Clausius-Clapeyron estabelece que  $\ln(P_v) \propto -\frac{1}{T}$ , em que  $P_v$  é a pressão de vapor d'água. Repita os itens (a)-(d), usando uma transformação apropriada.

**11.9.3** Um serviço elétrico está interessado em desenvolver um modelo relacionando a demanda na hora de pico ( $y$ , em quilowatts) e o consumo mensal total de energia ( $x$ , em quilowatts-hora). Dados para 50 consumidores residenciais são mostrados na tabela a seguir.

Consumidor	$x$	$y$	Consumidor	$x$	$y$
1	679	0,79	26	1434	0,31
2	292	0,44	27	837	4,20
3	1012	0,56	28	1748	4,88
4	493	0,79	29	1381	3,48
5	582	2,70	30	1428	7,58
6	1156	3,64	31	1255	2,63
7	997	4,73	32	1777	4,99
8	2189	9,50	33	370	0,59
9	1097	5,34	34	2316	8,19
10	2078	6,85	35	1130	4,79
11	1818	5,84	36	463	0,51
12	1700	5,21	37	770	1,74
13	747	3,25	38	724	4,10
14	2030	4,43	39	808	3,94
15	1643	3,16	40	790	0,96
16	414	0,50	41	783	3,29
17	354	0,17	42	406	0,44
18	1276	1,88	43	1242	3,24
19	745	0,77	44	658	2,14
20	795	3,70	45	1746	5,71
21	540	0,56	46	895	4,12
22	874	1,56	47	1114	1,90
23	1543	5,28	48	413	0,51
24	1029	0,64	49	1787	8,33
25	710	4,00	50	3560	14,94

## Exercícios para a Seção 11.10

**11.10.1** Um estudo foi conduzido tentando relacionar a posse de casa própria com a renda familiar. Vinte lares foram selecionados, e a renda familiar foi estimada, juntamente com a informação relativa à posse de casa própria ( $y = 1$  indica *sim* e  $y = 0$  indica *não*). Os dados são mostrados na tabela a seguir.

- Ajuste um modelo de regressão logística para a variável de resposta  $y$ . Use um modelo de regressão linear simples como a estrutura para o preditor linear.
- O modelo de regressão logística usado no item (a) é adequado?
- Dê uma interpretação do parâmetro  $\beta_1$  nesse modelo.

Lar	Renda	Status de Possuir Casa
1	38.000	0
2	51.200	1
3	39.600	0
4	43.400	1
5	47.700	0
6	53.000	0
7	41.500	1
8	40.800	0
9	45.400	1
10	52.400	1
11	38.700	1
12	40.100	0
13	49.500	1
14	38.000	0
15	42.000	1
16	54.000	1
17	51.700	1
18	39.400	0
19	40.900	0
20	52.800	1

**11.10.2** A resistência à compressão de um parafuso usado na construção de aeronaves está sendo estudada. Dez cargas foram selecionadas na faixa de 2500-4300 psi e um número de parafusos foi testado nessas cargas. Os números de parafusos que falharam em cada carga foram registrados. Os dados do teste completo são mostrados na tabela a seguir.

Carga, $x$ (psi)	Tamanho da Amostra, $n$	Número de Falhas, $r$
2500	50	10
2700	70	17
2900	100	30
3100	60	21
3300	40	18
3500	85	43
3700	90	54
3900	50	33
4100	80	60
4300	65	51

a. Ajuste um modelo de regressão logística para os dados. Use um modelo de regressão linear simples como a estrutura para o preditor linear.

b. O modelo de regressão logística usado no item (a) é adequado?

**11.10.3** O departamento de pesquisa de mercado de um fabricante de refrigerantes está investigando a eficiência de um cupom de desconto no preço quando da compra de um produto de dois litros da bebida. A uma amostra de 5500 consumidores, foram dados cupons com preços variando entre 5 e 25 centavos. A variável de resposta foi o número de cupons em cada categoria de preço de desconto resgatado depois de um mês. Os dados são mostrados a seguir.

Desconto, $x$	Tamanho da Amostra, $n$	Número Resgatado, $r$
5	500	100
7	500	122
9	500	147
11	500	176
13	500	211
15	500	244
17	500	277
19	500	310
21	500	343
23	500	372
25	500	391

a. Ajuste o modelo de regressão logística aos dados. Use um modelo de regressão linear simples como a estrutura para o preditor linear.

b. O modelo de regressão logística do item (a) é adequado?

c. Desenhe um gráfico dos dados e do modelo ajustado de regressão logística.

d. Expanda o preditor linear para incluir o termo quadrático. Existe alguma evidência de que esse termo quadrático seja necessário no modelo?

e. Desenhe um gráfico desse novo modelo no mesmo gráfico que você preparou no item (c). O modelo expandido fornece visualmente um melhor ajuste dos dados em relação ao modelo do item (a)?

**11.10.4** Um estudo foi feito para investigar compras de veículos novos. Uma amostra de 20 famílias foi selecionada. A cada família foi perguntado sobre a idade do seu veículo mais velho e a renda familiar total. Uma continuação da pesquisa foi conduzida seis meses depois, de modo a determinar se eles realmente haviam comprado um veículo novo durante aquele período ( $y = 1$  indica *sim* e  $y = 0$  indica *não*). Os dados desse estudo são mostrados na tabela a seguir.

Renda, $x_1$	Idade, $x_2$	$y$	Renda, $x_1$	Idade, $x_2$	$y$
45.000	2	0	37.000	5	1
40.000	4	0	31.000	7	1
60.000	3	1	40.000	4	1
50.000	2	1	75.000	2	0
55.000	2	0	43.000	9	1
50.000	5	1	49.000	2	0
35.000	7	1	37.500	4	1
65.000	2	1	71.000	1	0
53.000	2	0	34.000	5	0
48.000	1	0	27.000	6	0

- a. Ajuste o modelo de regressão logística aos dados.
- b. O modelo de regressão logística do item (a) é adequado?
- c. Interprete os coeficientes  $\beta_1$  e  $\beta_2$  do modelo.
- d. Qual é a probabilidade estimada de que uma família com uma renda de US\$ 45.000 e um carro com cinco anos de uso comprará um veículo novo nos próximos seis meses?
- e. Expanda o preditor linear para incluir o termo de interação. Existe alguma evidência de que esse termo seja necessário no modelo?

11.10.5 A Organização Mundial de Saúde define *obesidade* em adultos como tendo um índice de massa corporal (IMC) maior do que 30. Dos 250 homens do estudo mencionado no Exercício 11.2.1, 23 são obesos, por essa definição. Quão boa a medida de cintura (tamanho em polegadas) é como um preditor de obesidade? Um modelo de regressão logística foi ajustado aos dados:

$$\log\left(\frac{p}{1-p}\right) = -41,828 + 0,9864 \text{ cintura}$$

- em que  $p$  é a probabilidade de ser classificado como obeso.
- a. A probabilidade de ser classificado como obeso aumenta ou diminui como função do tamanho da cintura? Explique.
  - b. Qual é a probabilidade estimada de ser classificado como obeso um homem com um tamanho de cintura de 36 polegadas?
  - c. Qual é a probabilidade estimada de ser classificado como obeso um homem com um tamanho de cintura de 42 polegadas?
  - d. Qual é a probabilidade estimada de ser classificado como obeso um homem com um tamanho de cintura de 48 polegadas?
  - e. Faça um gráfico da probabilidade estimada de ser classificado como obeso em função do tamanho de cintura.

## Exercícios Suplementares para o Capítulo 11

11.S7 Um artigo publicado em *IEEE Transactions on Instrumentation and Measurement* ["Direct, Fast, and Accurate Measurement of  $V_T$  and  $K$  of MOS Transistor Using  $V_T$ -Sift Circuit" (Vol. 40, 1991, pp. 951-955)] descreveu o uso de um modelo de regressão linear simples para expressar a corrente (em miliampères)  $y$  como uma função da diferença de voltagem  $x$  (em volts). Os dados são apresentados a seguir:

$y$	$x$	$y$	$x$
0,734	1,1	1,50	1,6
0,886	1,2	1,66	1,7
1,04	1,3	1,81	1,8
1,19	1,4	1,97	1,9
1,35	1,5	2,12	2,0

- a. Desenhe um diagrama de dispersão desses dados. Uma relação linear parece plausível?
- b. Ajuste um modelo de regressão linear simples a esses dados.
- c. Teste a significância da regressão, usando  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?
- d. Encontre um intervalo de confiança de 95 % para o coeficiente angular estimado.
- e. Teste a hipótese  $H_0: \beta_0 = 0$  versus  $H_1: \beta_0 \neq 0$ , usando  $\alpha = 0,05$ . Que conclusões você pode tirar?

11.S8 Mostre que, para o modelo de regressão linear simples, as seguintes afirmações são verdadeiras:

- a.  $\sum_{i=1}^n (y_i - \hat{y}_i) = 0$
- b.  $\sum_{i=1}^n (y_i - \hat{y}_i)x_i = 0$
- c.  $\frac{1}{n} \sum_{i=1}^n \hat{y}_i = \bar{y}$

11.S9 Considere os dados mostrados na tabela a seguir. Suponha que a relação entre  $Y$  e  $x$  seja proposta como  $Y = (\beta_0 + \beta_1 x + \epsilon)^{-1}$ . Ajuste um modelo apropriado para os dados. A forma do modelo proposto parece razoável?

$x$	10	15	18	12	9	8	11	6
$y$	0,1	0,13	0,09	0,15	0,20	0,21	0,18	0,24

11.S10 A resistência do papel usado na fabricação de caixas de papelão ( $y$ ) está relacionada com a percentagem da concentração de madeira de lei na polpa original ( $x$ ). Sob condições controladas, uma planta piloto fabrica 16 amostras, cada uma sendo proveniente de uma batelada diferente de polpa. Mede-se a resistência à tração. Os dados são mostrados a seguir:

$y$	101,4	117,4	117,1	106,2
$x$	1,0	1,5	1,5	1,5
$y$	131,9	146,9	146,8	133,9
$x$	2,0	2,0	2,2	2,4
$y$	111,0	123,0	125,1	145,2
$x$	2,5	2,5	2,8	2,8
$y$	134,3	144,5	143,7	146,9
$x$	3,0	3,0	3,2	3,3

- a. Ajuste um modelo de regressão linear simples para os dados.
- b. Teste a significância da regressão, usando  $\alpha = 0,05$ .
- c. Construa um intervalo de confiança de 90 % para o coeficiente angular  $\beta_1$ .
- d. Construa um intervalo de confiança de 90 % para o coeficiente linear  $\beta_0$ .
- e. Construa um intervalo de confiança de 95 % para a resistência média em  $x = 2,5$ .
- f. Analise os resíduos e comente a adequação do modelo.

11.S11 Os dados da tabela a seguir, adaptados de Montgomery, Peck e Vining (2012), apresentam o número comprovado de deficientes mentais por 10.000 pessoas da população estimada do Reino Unido ( $y$ ) e o número (em milhões) de licenças emitidas ( $x$ ) de receptores de ondas de rádio pela BBC, para os anos 1924 a 1937. Ajuste um modelo de regressão relacionando  $y$  e  $x$ . Comente o modelo. Especificamente, a existência de uma forte correlação implica uma relação de causa e efeito?

Ano	y	x	Ano	y	x
1924	8	1.350	1931	16	4.620
1925	8	1.960	1932	18	5.497
1926	9	2.270	1933	19	6.260
1927	10	2.483	1934	20	7.012
1928	11	2.730	1935	21	7.618
1929	11	3.091	1936	22	8.131
1930	12	3.674	1937	23	8.593

**11.S12** Um artigo publicado em *Air and Waste* ["Update on Ozone Trends in California's South Coast Air Basin" (1993, Vol. 43, pp. 226-240)] estudou os níveis de ozônio em bacias aéreas da costa sul da Califórnia, durante os anos 1976-1991. O autor acredita que o número de dias em que o nível de ozônio excede 0,20 parte por milhão depende do índice sazonal meteorológico (a temperatura sazonal média a 850 milibar). Os dados estão na tabela a seguir.

Ano	Dias	Índice	Ano	Dias	Índice
1976	91	16,7	1984	81	18,0
1977	105	17,1	1985	65	17,2
1978	106	18,2	1986	61	16,9
1979	108	18,1	1987	48	17,1
1980	88	17,2	1988	61	18,2
1981	91	18,2	1989	43	17,3
1982	58	16,0	1990	33	17,5
1983	82	17,2	1991	36	16,6

- Construa um diagrama de dispersão dos dados.
- Ajuste um modelo de regressão linear simples a esses dados. Teste a significância da regressão.
- Encontre um IC de 95 % para o coeficiente angular  $\beta_1$ .
- Analise os resíduos e comente a adequação do modelo.

**11.S13** Um artigo publicado no *Journal of Applied Polymer Science* ["Preparation and Characterizations of Thermotropic Copolyester of *p*-Hydroxybenzoic Acid, Sebacic Acid, and Hydroquinone" (1995, Vol. 56, pp. 471-476)] estudou o efeito da razão molar de ácido sebácico sobre a viscosidade intrínseca de copoliésteres. Os dados são:

Fração molar, $x$	1,0	0,9	0,8	0,7	0,6	0,5	0,4	0,3
Viscosidade, $y$	0,45	0,20	0,34	0,58	0,70	0,57	0,55	0,44

- Construa um diagrama de dispersão dos dados.
- Ajuste um modelo de regressão linear simples a esses dados.
- Teste a significância da regressão. Calcule  $R^2$  para o modelo.
- Analise os resíduos e comente a adequação do modelo.

**11.S14** Dois métodos diferentes podem ser usados para medir a temperatura da solução em uma célula de Hall usada na fusão de alumínio: um termopar implantado na célula e uma medição indireta produzida em um dispositivo de IV. O método indireto é preferível

em razão da destruição dos termopares pela solução. Considere as 10 medidas seguintes:

Termopar	921	935	916	920	940
IV	918	934	924	921	945
Termopar	936	925	940	933	927
IV	930	919	943	932	935

- Construa um diagrama de dispersão para esses dados, sendo  $x$  = medida do termopar e  $y$  = medida do IV.
- Ajuste um modelo de regressão linear simples a esses dados.
- Teste a significância da regressão e calcule  $R^2$ . Que conclusões você pode tirar?
- Há evidência para sustentar a afirmação de que ambos os dispositivos produzem medidas equivalentes de temperatura? Formule e teste uma hipótese apropriada para sustentar essa afirmação.
- Analise os resíduos e comente a adequação do modelo.

**11.S15** Os dados a seguir relacionam os quilates de um diamante com os preços de compra. Eles aparecem no jornal de Cingapura *Business Times*, edição de 18 de fevereiro de 2000.

Quilate	Preço	Quilate	Preço
0,3	1302	0,33	1327
0,3	1510	0,33	1098
0,3	1510	0,34	1693
0,3	1260	0,34	1551
0,31	1641	0,34	1410
0,31	1555	0,34	1269
0,31	1427	0,34	1316
0,31	1427	0,34	1222
0,31	1126	0,35	1738
0,31	1126	0,35	1593
0,32	1468	0,35	1447
0,32	1202	0,35	1255
0,36	1635	0,45	1572
0,36	1485	0,46	2942
0,37	1420	0,48	2532
0,37	1420	0,5	3501
0,4	1911	0,5	3501
0,4	1525	0,5	3501
0,41	1956	0,5	3293
0,43	1747	0,5	3016

- Faça o gráfico. Qual é a relação entre o quilate e o preço? Há um *outlier*?
- O que você diria para aquela pessoa que comprou o diamante que foi um *outlier*?
- Ajuste dois modelos de regressão: um, com todos os dados, e o outro, com os dados não usuais omitidos. Estime o coeficiente de inclinação, com um intervalo de confiança de 95 % em ambos os casos. Comente qualquer diferença.

# Capítulo 12 Exercícios

## Exercícios para a Seção 12.1

**12.1.1** O Exercício 11.2.1 descreveu um modelo de regressão entre o percentual de gordura corporal (% GC), medido por imersão, e o IMC proveniente de um estudo em 250 homens. Os pesquisadores mediram também 13 características físicas de cada homem, incluindo sua idade (anos), altura (polegadas) e tamanho da cintura (polegadas).

Uma regressão do percentual de gordura corporal com a altura e o tamanho da cintura como preditores mostra a seguinte saída computacional:

	Estimativa	Erro-Padrão	Valor t	Pr(> t )
(Interseção)	-3,10088	7,68611	-0,403	0,687
Altura	-0,60154	0,10994	-5,472	1,09e-07
Cintura	1,77309	0,07158	24,770	< 2e -16

Erro-padrão residual: 4,46 para 247 graus de liberdade  
 R<sup>2</sup> múltiplo: 0,7132, R<sup>2</sup> ajustado: 0,7109  
 Estatística F: 307,1 para 2 e 247 GL, valor p < 2,2e<sup>-16</sup>

a. Escreva o modelo de regressão se

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 2,9705 & -4,0042E-2 & -4,1679E-2 \\ -0,04004 & 6,0774E-4 & -7,3875E-5 \\ -4,1679E-3 & -7,3875E-5 & 2,5766E-4 \end{bmatrix}$$

e

$$(\mathbf{X}'\mathbf{y}) = \begin{bmatrix} 4757,9 \\ 334335,8 \\ 179706,7 \end{bmatrix}$$

b. Verifique que o modelo encontrado da tecnologia é correto no mínimo na segunda casa decimal.

c. Qual é a gordura corporal prevista de um homem de 6 pés de altura e 34 polegadas de cintura?

**12.1.2** Hsuie, Ma e Tsai [“Separation and Characterizations of Thermotropic Copolyesters of p-Hydroxybenzoic Acid, Sebacic Acid, and Hydroquinone”, *Journal of Applied Polymer Science* (1995, Vol. 56, pp. 471-476)] estudaram o efeito da razão molar de ácido sebácico (o regressor) sobre a viscosidade intrínseca de copoliésteres (a resposta). A seguir, os dados são apresentados:

Razão	Viscosidade
1,0	0,45
0,9	0,20
0,8	0,34
0,7	0,58
0,6	0,70
0,5	0,57
0,4	0,55
0,3	0,44

- a. Construa um diagrama de dispersão dos dados.
- b. Ajuste uma equação de previsão de segunda ordem.

**12.1.3** A porcentagem do mercado de trabalho para engenheiros em cada estado dos EUA pode ser prevista pela quantidade de dinheiro gasto em ensino superior (como uma porcentagem do produto interno bruto), pelo capital de risco (dólares por \$ 1000 do produto interno bruto) para ideias de negócios em alta tecnologia e pelo financiamento do estado (em dólares por estudante) nas maiores universidades de pesquisa? Dados para todos os 50 estados e um *software* revelaram os seguintes resultados:

	Estimativa	Erro-Padrão	Valor t	Pr(> t )
(Interseção)	1,051e+00	1,567e-01	6,708	2,5e-08 ***
Capital de risco	9,514e-02	3,910e-02	2,433	0,0189 *
Financiamento do estado	4,106e-06	1,437e-05	0,286	0,7763
Ensino superior	-1,673e-01	2,595e-01	-0,645	0,5223

Erro-padrão residual: 0,3007 para 46 graus de liberdade  
 R<sup>2</sup> múltiplo: 0,1622, R<sup>2</sup> ajustado: 0,1075  
 Estatística F: 2,968 para 3 e 46 GL, valor p: 0,04157

- a. Escreva a equação para prever o percentual de engenheiros no mercado de trabalho.
- b. Qual o percentual de engenheiros que você espera ver no mercado de trabalho em um estado que tem \$ 1 por \$ 1000 em capital de risco, gasta \$ 10.000 por estudante para financiar seus estudos na universidade e gasta 0,5 % de seu PIB em educação superior?
- c. Se o estado no item (b) realmente tivesse 1,5 % de engenheiros na força de trabalho, qual seria o resíduo?

**12.1.4** Um modelo de regressão deve ser desenvolvido para prever a habilidade do solo de absorver contaminantes químicos. Foram obtidas dez observações do índice de absorção pelo solo (y) e dois regressores: x<sub>1</sub> = quantidade que pode ser extraída de minério de ferro e x<sub>2</sub> = quantidade de bauxita. Desejamos ajustar o modelo y = β<sub>0</sub> + β<sub>1</sub>x<sub>1</sub> + β<sub>2</sub>x<sub>2</sub> + ε. Algumas grandezas necessárias são:

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 1.17991 & -7,30982 E-3 & 7,3006 E-4 \\ -7,30982 E-3 & 7,9799 E-5 & -1,23713 E-4 \\ 7,3006 E-4 & -1,23713 E-4 & 4,6576 E-4 \end{bmatrix}$$

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} 220 \\ 36.768 \\ 9.965 \end{bmatrix}$$

- a. Estime os coeficientes de regressão do modelo especificado anteriormente.
- b. Qual é o valor previsto do índice de absorção y, quando x<sub>1</sub> = 200 e x<sub>2</sub> = 50?

**12.1.5** Os dados provenientes de uma pesquisa de satisfação de pacientes em hospitais são mostrados a seguir:

Observação	Idade	Gravidade	Médico-Cirúrgico	Ansiedade	Satisfação
1	55	50	0	2,1	68
2	46	24	1	2,8	77
3	30	46	1	3,3	96
4	35	48	1	4,5	80
5	59	58	0	2,0	43

(continua)

(continuação)

Observação	Idade	Gravidade	Médico-Cirúrgico	Ansiedade	Satisfação
6	61	60	0	5,1	44
7	74	65	1	5,5	26
8	38	42	1	3,2	88
9	27	42	0	3,1	75
10	51	50	1	2,4	57
11	53	38	1	2,2	56
12	41	30	0	2,1	88
13	37	31	0	1,9	88
14	24	34	0	3,1	102
15	42	30	0	3,0	88
16	50	48	1	4,2	70
17	58	61	1	4,6	52
18	60	71	1	5,3	43
19	62	62	0	7,2	46
20	68	38	0	7,8	56
21	70	41	1	7,0	59
22	79	66	1	6,2	26
23	63	31	1	4,1	52
24	39	42	0	3,5	83
25	49	40	1	2,1	75

As variáveis regressoras são idade do paciente, índice de gravidade da doença (valores maiores indicam maior gravidade), uma variável indicativa denotando se o paciente é médico (0) ou cirúrgico (1), e um índice de ansiedade (valores maiores indicam grande ansiedade).

a. Ajuste um modelo de regressão linear múltipla para a resposta de satisfação do paciente, usando a idade, a gravidade da doença e o índice de ansiedade como os regressores.

b. Estime  $\sigma^2$ .

c. Encontre os erros-padrão dos coeficientes de regressão.

d. Estão todos os parâmetros do modelo estimados com aproximadamente a mesma precisão? Por que sim ou por que não?

**12.1.6** A energia elétrica consumida mensalmente por uma indústria química está relacionada com a temperatura média ambiente ( $x_1$ ), o número de dias no mês ( $x_2$ ), a pureza média do produto ( $x_3$ ) e as toneladas do produto produzido ( $x_4$ ). Os dados históricos do ano passado estão disponíveis e são apresentados na tabela a seguir.

y	$x_1$	$x_2$	$x_3$	$x_4$
240	25	24	91	100
236	31	21	90	95
270	45	24	88	110
274	60	25	87	88
301	65	25	91	94
316	72	26	94	99
300	80	25	87	97
296	84	25	86	96
267	75	24	88	110
276	60	25	91	105
288	50	25	90	100
261	38	23	89	98

a. Ajuste um modelo de regressão linear múltipla a esses dados.

b. Estime  $\sigma^2$ .

c. Calcule os erros-padrão dos coeficientes de regressão. Estão todos os parâmetros do modelo estimados com aproximadamente a mesma precisão? Por que sim ou por que não?

d. Preveja o consumo de energia para um mês em que  $x_1 = 75$  °F,  $x_2 = 24$  dias,  $x_3 = 90$  % e  $x_4 = 98$  toneladas.

**12.1.7** Um artigo publicado em *Electronic Packaging and Production* (2002, Vol. 42) considerou o efeito da inspeção de raios X de circuitos integrados. As doses de radiação (expressa em rads) foram estudadas em função da corrente (em miliampères) e do tempo (em minutos) de exposição. Os dados estão na seguinte tabela:

Dose de Radiação	mAmps	Tempo de Exposição
7,4	10	0,25
14,8	10	0,5
29,6	10	1
59,2	10	2
88,8	10	3
296	10	10
444	10	15
592	10	20
11,1	15	0,25
22,2	15	0,5
44,4	15	1
88,8	15	2
133,2	15	3
444	15	10
666	15	15
888	15	20
14,8	20	0,25
29,6	20	0,5
59,2	20	1
118,4	20	2
177,6	20	3
592	20	10
888	20	15
1184	20	20
22,2	30	0,25
44,4	30	0,5
88,8	30	1
177,6	30	2
266,4	30	3
888	30	10
1332	30	15
1776	30	20
29,6	40	0,25
59,2	40	0,5
118,4	40	1
236,8	40	2
355,2	40	3
1184	40	10
1776	40	15
2368	40	20

- a. Ajuste um modelo de regressão linear múltipla a esses dados, com dose de radiação como resposta.
- b. Estime  $\sigma^2$  e os erros-padrão dos coeficientes de regressão.
- c. Use o modelo para prever a dose de radiação, quando a corrente for de 15 miliampères e o tempo de exposição for de cinco segundos.

12.1.8 A resistência à tração de um fio colado é uma característica importante. A tabela a seguir fornece informação sobre resistência à tração ( $y$ ), altura da garra ( $x_1$ ), altura da coluna ( $x_2$ ), altura da alça ( $x_3$ ), comprimento do fio ( $x_4$ ), largura da parte colada segurada pela garra ( $x_5$ ) e largura da parte colada segurada pela coluna ( $x_6$ ).

$y$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
8,0	5,2	19,6	29,6	94,9	2,1	2,3
8,3	5,2	19,8	32,4	89,7	2,1	1,8
8,5	5,8	19,6	31,0	96,2	2,0	2,0
8,8	6,4	19,4	32,4	95,6	2,2	2,1
9,0	5,8	18,6	28,6	86,5	2,0	1,8
9,3	5,2	18,8	30,6	84,5	2,1	2,1
9,3	5,6	20,4	32,4	88,8	2,2	1,9
9,5	6,0	19,0	32,6	85,7	2,1	1,9
9,8	5,2	20,8	32,2	93,6	2,3	2,1
10,0	5,8	19,9	31,8	86,0	2,1	1,8
10,3	6,4	18,0	32,6	87,1	2,0	1,6
10,5	6,0	20,6	33,4	93,1	2,1	2,1
10,8	6,2	20,2	31,8	83,4	2,2	2,1
11,0	6,2	20,2	32,4	94,5	2,1	1,9
11,3	6,2	19,2	31,4	83,4	1,9	1,8
11,5	5,6	17,0	33,2	85,2	2,1	2,1
11,8	6,0	19,8	35,4	84,1	2,0	1,8
12,3	5,8	18,8	34,0	86,9	2,1	1,8
12,5	5,6	18,6	34,2	83,0	1,9	2,0

- a. Ajuste o modelo de regressão linear múltipla, usando  $x_2, x_3, x_4$  e  $x_5$  como regressores.
- b. Estime  $\sigma^2$ .
- c. Encontre  $ep(\hat{\beta}_j)$ . Em sua opinião, quão precisamente os coeficientes de regressão são estimados?
- d. Use o modelo do item (a) para prever a resistência à tração, quando  $x_2 = 20, x_3 = 30, x_4 = 90$  e  $x_5 = 2,0$ .

12.1.9 Em um artigo publicado em *IEEE Transactions on Instrumentation and Measurement* ["Measurement and Calculation of Powered Mixture Permittivities" (2001, Vol. 50, pp. 1066-1070)], misturas em pó de carvão e calcário foram analisadas em relação à permissividade. A variável de resposta foi erros na medida de densidade. Os dados são reportados a seguir.

Densidade	Constante Dielétrica	Fator de Perda
0,749	2,05	0,016
0,798	2,15	0,02
0,849	2,25	0,022
0,877	2,3	0,023
0,929	2,4	0,026
0,963	2,47	0,028
0,997	2,54	0,031

Densidade	Constante Dielétrica	Fator de Perda
1,046	2,64	0,034
1,133	2,85	0,039
1,17	2,94	0,042
1,215	3,05	0,045

- a. Ajuste um modelo de regressão linear múltipla para esses dados, tendo a densidade como a resposta.
- b. Estime  $\sigma^2$  e os erros-padrão dos coeficientes de regressão.
- c. Use o modelo para prever a densidade quando a constante dielétrica for 2,5 e o fator de perda for 0,03.

12.1.10 Um artigo publicado em *Biotechnology Progress* ["Optimization of Conditions for Bacteriocin Extraction in PEG/Salt Aqueous Two-Phase Systems Using Statistical Experimental Designs" (2001, Vol. 17, pp. 366-368)] reportou um experimento para investigar e otimizar a extração de nisina em sistemas bifásicos aquosos (SBA). A recuperação de nisina foi a variável dependente ( $y$ ). As duas variáveis regressoras foram concentração (%) de PEG 4000 (denotada como  $x_1$ ) e concentração (%) de  $\text{Na}_2\text{SO}_4$  (denotada como  $x_2$ ). Os dados estão na tabela a seguir.

$x_1$	$x_2$	$y$
13	11	62.8739
15	11	76.1328
13	13	87.4667
15	13	102.3236
14	12	76.1872
14	12	77.5287
14	12	76.7824
14	12	77.4381
14	12	78.7417

- a. Ajuste um modelo de regressão linear múltipla a esses dados.
- b. Estime  $\sigma^2$  e os erros-padrão dos coeficientes de regressão.
- c. Use o modelo para prever a recuperação de nisina, quando  $x_1 = 14,5$  e  $x_2 = 12,5$ .

12.1.11 Um artigo publicado em *Optical Engineering* ["Operating Curve Extraction of a Correlator's Filter" (2004, Vol. 43, pp. 2775-2779)] reportou o uso de um correlador óptico para fazer um experimento, variando o brilho e o contraste. A modulação resultante é caracterizada pela faixa útil de níveis de cinza. Os dados são mostrados a seguir:

Brilho (%):	54	61	65	100	100	100	50	57	54
Contraste (%):	56	80	70	50	65	80	25	35	26
Faixa útil (ng):	96	50	50	112	96	80	155	144	255

- a. Ajuste um modelo de regressão linear múltipla a esses dados.
- b. Estime  $\sigma^2$ .
- c. Calcule os erros-padrão dos coeficientes de regressão.
- d. Preveja a faixa útil, quando o brilho for 80 e o contraste for 75.

12.1.12 A tabela a seguir apresenta a estatística para os times da Liga Nacional de Hockey para a temporada 2008-2009 (*The Sports Network*). Ajuste um modelo de regressão linear múltipla relacionando as Vitórias com as variáveis  $GF$  e  $JMI$ . Uma vez que os times jogam 82 jogos,  $V = 82 - D - E - DTE$ ; porém, tal modelo não ajuda a construir um time melhor. Estime  $\sigma^2$  e encontre os erros-padrão dos coeficientes de regressão para seu modelo.

Time	V	D	DTE	PTS	GF	GC	VT	GFPIN	PI	TMP	TMPM	MED	TTIN	GCPIN	PMP	GFTIN	GCTIN	JMP
Anaheim	42	33	7	91	238	235	309	73	23,6	1418	8	17,4	385	78	79,7	6	6	43
Atlanta	35	41	6	76	250	279	357	69	19,3	1244	12	15,3	366	88	76	13	9	39
Boston	53	19	10	116	270	190	313	74	23,6	1016	12	12,5	306	54	82,4	8	7	47
Buffalo	41	32	9	91	242	229	358	75	21	1105	16	13,7	336	61	81,8	7	4	44
Carolina	45	30	7	97	236	221	374	70	18,7	786	16	9,8	301	59	80,4	8	7	39
Columbus	41	31	10	92	220	223	322	41	12,7	1207	20	15	346	62	82,1	8	9	41
Calgary	46	30	6	98	251	246	358	61	17	1281	18	15,8	349	58	83,4	6	13	37
Chicago	46	24	12	104	260	209	363	70	19,3	1129	28	14,1	330	64	80,6	10	5	43
Colorado	32	45	5	69	190	253	318	50	15,7	1044	18	13	318	64	79,9	4	5	31
Dallas	36	35	11	83	224	251	351	54	15,4	1134	10	14	327	70	78,6	2	2	38
Detroit	51	21	10	112	289	240	353	90	25,5	810	14	10	327	71	78,3	6	4	46
Edmonton	38	35	9	85	228	244	354	60	17	1227	20	15,2	338	76	77,5	3	8	39
Florida	41	30	11	93	231	223	308	51	16,6	884	16	11	311	54	82,6	7	6	39
Los Angeles	34	37	11	79	202	226	360	69	19,2	1191	16	14,7	362	62	82,9	4	7	39
Minnesota	40	33	9	89	214	197	328	66	20,1	869	20	10,8	291	36	87,6	9	6	39
Montreal	41	30	11	93	242	240	374	72	19,2	1223	6	15	370	65	82,4	10	10	38
New Jersey	51	27	4	106	238	207	307	58	18,9	1038	20	12,9	324	65	79,9	12	3	44
Nashville	40	34	8	88	207	228	318	50	15,7	982	12	12,1	338	59	82,5	9	8	41
NY Islanders	26	47	9	61	198	274	320	54	16,9	1198	18	14,8	361	73	79,8	12	5	37
NY Rangers	43	30	9	95	200	212	346	48	13,9	1175	24	14,6	329	40	87,8	9	13	42
Ottawa	36	35	11	83	213	231	339	66	19,5	1084	14	13,4	346	64	81,5	8	5	46
Philadelphia	44	27	11	99	260	232	316	71	22,5	1408	26	17,5	393	67	83	16	1	43
Phoenix	36	39	7	79	205	249	344	50	14,5	1074	18	13,3	293	68	76,8	5	4	36
Pittsburgh	45	28	9	99	258	233	360	62	17,2	1106	8	13,6	347	60	82,7	7	11	46
San Jose	53	18	11	117	251	199	360	87	24,2	1037	16	12,8	306	51	83,3	12	10	46
St. Louis	41	31	10	92	227	227	351	72	20,5	1226	22	15,2	357	58	83,8	10	8	35
Tampa Bay	24	40	18	66	207	269	343	61	17,8	1280	26	15,9	405	89	78	4	8	34
Toronto	34	35	13	81	244	286	330	62	18,8	1113	12	13,7	308	78	74,7	6	7	40
Vancouver	45	27	10	100	243	213	357	67	18,8	1323	28	16,5	371	69	81,4	7	5	47
Washington	50	24	8	108	268	240	337	85	25,2	1021	20	12,7	387	75	80,6	7	9	45

V	Vitórias.	TMP	Total de minutos de penalidade, incluindo minutos que o jogador fica no banco.
D	Derrotas no tempo regulamentar.	TMPM	Total de minutos de penalidades menores no banco.
DTE	Derrotas no tempo extra.	MED	Média de minutos de penalidade por jogo.
PTS	Pontos. Dois pontos para ganhar um jogo, um ponto para um empate ou derrota no tempo extra e zero ponto para derrota em tempo regulamentar.	TTIN	Tempos totais em que o time está com menor número de jogadores na quadra. Mede as oportunidades do oponente.
GF	Gols a favor.	GCPIN	Gols contra durante o período de inferioridade numérica.
GC	Gols contra.	PMP	Porcentagem de morte por penalidades. Mede a habilidade do time de prevenir gols, enquanto seu oponente está em desvantagem numérica. Oportunidades do oponente menos gols durante o período de inferioridade numérica divididas pelas oportunidades do oponente.
VT	Vantagens totais, Oportunidades durante o período em que se joga em inferioridade numérica.	GFTIN	Gols a favor durante o período em que o time está com menor número de jogadores na quadra.
GFPIN	Gols a favor marcados durante o período em que se joga em inferioridade numérica.	GCTIN	Gols contra durante o período em que o time está com menor número de jogadores na quadra.
PI	Porcentagem de inferioridade numérica, Gols durante o período em que se joga em inferioridade numérica divididos pelas vantagens totais.	JMP	Jogos em que o time marcou primeiro.

12.1.13 Um estudo foi realizado sobre o desgaste de um mancal e sua relação com  $x_1 =$  viscosidade do óleo e  $x_2 =$  carga. Os dados seguintes foram obtidos.

y	$x_1$	$x_2$
293	1,6	851
230	15,5	816
172	22,0	1058
91	43,0	1201
113	33,0	1357
125	40,0	1115

- Ajuste um modelo de regressão linear múltipla a esses dados.
- Estime  $\sigma^2$  e os erros-padrão dos coeficientes de regressão.
- Use o modelo para prever o desgaste, quando  $x_1 = 25$  e  $x_2 = 1000$ .
- Ajuste a esses dados um modelo de regressão linear múltipla, com um termo de interação.
- Estime  $\sigma^2$  e  $ep(\hat{\beta}_j)$  para esse novo modelo. Como essas grandezas variam? Isso diz a você alguma coisa sobre o valor de adicionar o termo de interação ao modelo?
- Use o modelo do item (d) para prever o desgaste quando  $x_1 = 25$  e  $x_2 = 1000$ . Compare essa previsão com o valor previsto no item (c).

## Exercícios para a Seção 12.2

12.2.1 Lembre-se da regressão do percentual de gordura corporal em função da altura e do tamanho da cintura, proveniente do Exercício 12.1.1. O modelo de regressão simples do percentual de gordura corporal em função somente da altura mostra o seguinte:

	Estimativa	Erro-padrão	valor $t$	$\Pr(> t )$
(Interseção)	25,58078	14,15400	1,807	0,0719
Altura	-0,09316	0,20119	-0,463	0,6438

- Teste se o coeficiente da altura é estatisticamente significativo.
- Olhando o modelo contendo o tamanho da cintura e altura, teste se o coeficiente da altura é significativo nesse modelo.
- Explique a discrepância em suas duas respostas.

12.2.2 Considere o modelo de regressão linear do Exercício 12.1.2. O termo de segunda ordem é necessário no modelo de regressão?

12.2.3 Considere o modelo de regressão do Exercício 12.1.3 e tente prever o percentual de engenheiros no mercado de trabalho a partir de várias variáveis de gastos.

- Existe alguma variável útil para a previsão? (Teste uma hipótese apropriada.)
- Que percentual de variação no percentual de engenheiros é considerado pelo modelo?
- O que você deve fazer em seguida para criar um modelo melhor?

12.2.4 Você ajustou um modelo de regressão, com dois regressores, a um conjunto de dados que tem 20 observações. A soma total dos quadrados é 1000 e a soma dos quadrados do modelo é 750.

- Qual é o valor de  $R^2$  para esse modelo?
- Qual é o valor ajustado de  $R^2$  para esse modelo?
- Qual é o valor da estatística  $F$  para testar a significância da regressão? Que conclusões você tiraria sobre esse modelo, se  $\alpha = 0,05$ ? E se  $\alpha = 0,01$ ?
- Suponha que você adicione um terceiro regressor ao modelo e, como resultado, a soma dos quadrados do modelo seja agora 785. Parece para você que a adição desse fator melhorou o modelo?

12.2.5 Considere os dados do índice de absorção no Exercício 12.1.4. A soma total dos quadrados para  $y$  é  $SQ_T = 742,00$ .

- Teste a significância da regressão, usando  $\alpha = 0,01$ . Qual é o valor  $P$  para esse teste?
- Teste a hipótese  $H_0: \beta_1 = 0$  versus  $H_1: \beta_1 \neq 0$ , usando  $\alpha = 0,01$ . Qual é o valor  $P$  para esse teste?
- Que conclusão você pode tirar acerca da utilidade de  $x_1$  como um regressor nesse modelo?

12.2.6 Considere os dados do consumo de energia elétrica no Exercício 12.1.6.

- Teste a significância da regressão, usando  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?
- Use o teste  $t$  com a finalidade de quantificar a contribuição de cada regressor para o modelo. Usando  $\alpha = 0,05$ , que conclusões você pode tirar?

12.2.7 Considere o modelo de regressão que ajustou os dados de inspeção de raios  $X$  no Exercício 12.1.7. Use a dose de radiação como a resposta.

- Teste a significância da regressão, usando  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?
- Construa um teste  $t$  para cada coeficiente de regressão. Que conclusões você pode tirar acerca das variáveis nesse modelo? Use  $\alpha = 0,05$ .

12.2.8 Considere os dados da resistência à tração de um fio colado no Exercício 12.1.8.

- Teste a significância da regressão, usando  $\alpha = 0,05$ . Encontre o valor  $P$  para esse teste. Que conclusões você pode tirar?
- Calcule a estatística  $t$  para cada coeficiente de regressão. Usando  $\alpha = 0,05$ , que conclusões você pode tirar? Todas as variáveis contribuem para o modelo?

12.2.9 Considere o modelo de regressão que ajustou os dados de modulação na faixa do cinza no Exercício 12.1.11. Use a faixa útil como a resposta.

- Teste a significância da regressão, usando  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?
- Construa um teste  $t$  para cada coeficiente de regressão. Que conclusões você pode tirar acerca das variáveis nesse modelo? Use  $\alpha = 0,05$ .

12.2.10 Considere o modelo de regressão que ajustou os dados de extração de nisina do Exercício 12.1.10. Use a extração de nisina como a resposta.

- Teste a significância da regressão, usando  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste?
- Construa um teste  $t$  para cada coeficiente de regressão. Que conclusões você pode tirar acerca das variáveis nesse modelo? Use  $\alpha = 0,05$ .

c. Comente o efeito de uma amostra de tamanho pequeno sobre os testes dos itens anteriores.

12.2.11 Considere os dados de desgaste do mancal no Exercício 12.1.13.

- Para o modelo sem interação, teste a significância da regressão, usando  $\alpha = 0,05$ . Qual é o valor  $P$  para esse teste? Quais são as suas conclusões?
- Para o modelo sem interação, calcule a estatística  $t$  para cada coeficiente de regressão. Usando  $\alpha = 0,05$ , que conclusões você pode tirar?
- Para o modelo sem interação, use o método da soma extra dos quadrados para investigar a utilidade de adicionar ao modelo  $x_2 =$  carga, que já contém  $x_1 =$  viscosidade do óleo. Use  $\alpha = 0,05$ .
- Reajuste o modelo com um termo de interação. Teste a significância da regressão, usando  $\alpha = 0,05$ .
- Use o método da soma extra dos quadrados para determinar se o termo de interação contribui significativamente para o modelo. Use  $\alpha = 0,05$ .
- Estime  $\sigma^2$  para o modelo com interação. Compare esse valor com a estimativa de  $\sigma^2$  do modelo do item (a).

12.2.12 Dados sobre o desempenho dos times da Liga Nacional de Hóquei (*National Hockey League*) foram apresentados no Exercício 12.1.12.

- Teste o modelo desse exercício em relação à significância da regressão, usando  $\alpha = 0,05$ . Que conclusões você pode tirar?
- Use o teste  $t$  para avaliar a contribuição de cada regressor ao modelo. Parece que todos os regressores são necessários? Use  $\alpha = 0,05$ .
- Ajuste um modelo de regressão relacionando o número de jogos ganhos com o número de gols feitos e o número dos gols feitos quando o time estava em vantagem numérica. Considerando a sua resposta no item (b), a escolha dos regressores parece ser lógica? Teste esse modelo novo em relação à significância da regressão e avalie a contribuição de cada regressor ao modelo, usando o teste  $t$ . Use  $\alpha = 0,05$ .

12.2.13 Dados da pesquisa sobre a satisfação de pacientes de hospitais foram apresentados no Exercício 12.1.5.

a. Teste o modelo desse exercício em relação à significância da regressão. Que conclusões você pode tirar, se  $\alpha = 0,05$ ? E se  $\alpha = 0,01$ ?

b. Teste a contribuição de regressores individuais, usando o teste  $t$ . Parece que todos os regressores usados no modelo são realmente necessários?

## Exercícios para as Seções 12.3 e 12.4

12.4.1 Usando o modelo de regressão do Exercício 12.1.1,

a. Encontre um intervalo de confiança de 95 % para o coeficiente da altura.

b. Encontre um intervalo de confiança de 95 % para o percentual médio da gordura corporal para um homem com uma altura de 72 polegadas e cintura de 34 polegadas.

c. Encontre um intervalo de previsão de 95 % para o percentual médio da gordura corporal para um homem com a mesma altura e cintura do item (b).

d. Que intervalo é mais largo: o intervalo de confiança ou o intervalo de previsão? Explique brevemente.

e. Dada a sua resposta do item (c), você acredita que esse seja um modelo útil para prever a gordura corporal? Explique brevemente.

12.4.2 Use o modelo de regressão polinomial de segunda ordem do Exercício 12.1.2

a. Encontre um intervalo de confiança de 95 % para os termos de primeira e segunda ordem nesse modelo.

b. O zero está no intervalo de confiança que você encontrou para o termo de segunda ordem no item (a)? O que esse fato lhe diz sobre a contribuição do termo de segunda ordem para o modelo?

c. Refaça o modelo com somente o termo de primeira ordem. Encontre um intervalo de confiança de 95 % para esse termo. Esse intervalo é maior ou menor que o intervalo de confiança que você encontrou para esse termo no item (a)?

12.4.3 Considere os dados de absorção do solo no Exercício 12.1.4.

a. Encontre intervalos de confiança de 95 % para os coeficientes de regressão.

b. Encontre um intervalo de confiança de 95 % para o índice de absorção média do solo, quando  $x_1 = 200$  e  $x_2 = 50$ .

c. Encontre um intervalo de previsão de 95 % para o índice de absorção do solo, quando  $x_1 = 200$  e  $x_2 = 50$ .

12.4.4 Considere os dados sobre consumo de energia elétrica no Exercício 12.1.6.

a. Encontre os intervalos de confiança de 95 % para  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  e  $\beta_4$ .

b. Encontre um intervalo de confiança de 95 % para a média de  $Y$ , quando  $x_1 = 75$ ,  $x_2 = 24$ ,  $x_3 = 90$  e  $x_4 = 98$ .

c. Encontre um intervalo de previsão de 95 % para o consumo de energia, quando  $x_1 = 75$ ,  $x_2 = 24$ ,  $x_3 = 90$  e  $x_4 = 98$ .

12.4.5 Considere o ajuste do modelo de regressão aos dados de inspeção de raios X no Exercício 12.1.7. Use dose de radiação como a resposta.

a. Calcule o intervalo de confiança de 95 % para cada coeficiente de regressão.

b. Calcule um intervalo de confiança de 99 % para a dose média de radiação a 15 miliampères e 1 segundo de tempo de exposição.

c. Calcule um intervalo de previsão de 99 % para a dose de radiação para os mesmos valores dos regressores usados no item (b).

12.4.6 Considere os dados sobre resistência à tração de um fio colado no Exercício 12.1.8.

a. Encontre o intervalo de confiança de 95 % para os coeficientes de regressão.

b. Encontre um intervalo de confiança de 95 % para a resistência média à tração, quando  $x_2 = 20$ ,  $x_3 = 30$ ,  $x_4 = 90$  e  $x_5 = 2,0$ .

c. Encontre um intervalo de previsão de 95 % para a resistência à tração, quando  $x_2 = 20$ ,  $x_3 = 30$ ,  $x_4 = 90$  e  $x_5 = 2,0$ .

12.4.7 Considere o ajuste do modelo de regressão aos dados de modulação na faixa do cinza no Exercício 12.1.11. Use a faixa útil como a resposta.

a. Calcule os intervalos de confiança de 99 % para cada coeficiente de regressão.

b. Calcule um intervalo de confiança de 99 % para a faixa útil média, quando brilho = 70 e contraste = 80.

c. Calcule um intervalo de previsão para a faixa útil considerando os mesmos valores dos regressores usados no item anterior.

d. Calcule um intervalo de confiança de 99 % e um intervalo de previsão de 99 % para a faixa útil, quando brilho = 50 e contraste = 25. Compare as larguras desses intervalos com aqueles calculados nos itens (b) e (c). Explique qualquer diferença nas larguras.

12.4.8 Considere o ajuste do modelo de regressão aos dados da extração de nisina no Exercício 12.1.10.

a. Calcule os intervalos de confiança de 95 % para cada coeficiente de regressão.

b. Calcule um intervalo de confiança de 95 % para a extração média de nisina, quando  $x_1 = 15,5$  e  $x_2 = 16$ .

c. Calcule um intervalo de previsão para a extração de nisina para os mesmos valores dos regressores usados no item (b).

d. Comente o efeito de uma amostra de tamanho pequeno sobre a largura desses intervalos.

12.4.9 Considere os dados de LAH no Exercício 12.1.12.

a. Encontre um intervalo de confiança de 95 % para o coeficiente de regressão da variável  $GF$ .

b. Ajuste um modelo de regressão linear simples, relacionando a variável de resposta  $V$  com o regressor  $GF$ .

c. Encontre um intervalo de confiança de 95 % para a inclinação do modelo de regressão linear simples do item (b).

d. Compare os comprimentos dos dois intervalos de confiança calculados nos itens (a) e (c). Qual é o menor intervalo? Isso diz a você alguma coisa acerca de qual modelo seja preferível?

## Exercícios para a Seção 12.5

12.5.1 Considere o modelo de regressão para os dados de inspeção de raios X no Exercício 12.1.7. Use dose de radiação (rads) como a resposta.

- Que proporção da variabilidade total é explicada por esse modelo?
- Construa um gráfico de probabilidade normal dos resíduos. Que conclusão você pode tirar desse gráfico?
- Plote os resíduos contra  $\hat{y}$ , contra cada regressor e comente a adequação do modelo.
- Calcule a distância de Cook para as observações nesse conjunto de dados. Existe alguma observação influente nesses dados?

12.5.2 Considere os dados sobre o consumo de energia elétrica no Exercício 12.1.6.

- Calcule  $R^2$  para esse modelo. Interprete essa grandeza.
- Plote os resíduos contra  $\hat{y}$  e contra cada regressor. Interprete esse gráfico.
- Construa um gráfico de probabilidade normal dos resíduos e comente a suposição de normalidade.

12.5.3 Considere o modelo de regressão para os dados da mistura de carvão e de calcário no Exercício 12.1.9. Use densidade como a resposta.

- Que proporção da variabilidade total é explicada por esse modelo?
- Construa um gráfico de probabilidade normal dos resíduos. Que conclusão você pode tirar desse gráfico?
- Plote os resíduos contra  $\hat{y}$ , contra cada regressor e comente a adequação do modelo.
- Calcule a distância de Cook para as observações nesse conjunto de dados. Existe alguma observação influente nesses dados?

12.5.4 Considere o modelo de regressão para os dados de extração de nisina no Exercício 12.1.10.

- Que proporção da variabilidade total é explicada por esse modelo?
- Construa um gráfico de probabilidade normal dos resíduos. Que conclusão você pode tirar desse gráfico?
- Plote os resíduos contra  $\hat{y}$ , contra cada regressor e comente a adequação do modelo.
- Calcule a distância de Cook para as observações nesse conjunto de dados. Existe alguma observação influente nesses dados?

12.5.5 Considere o modelo de regressão para os dados de modulação da faixa cinza no Exercício 12.1.11. Use a faixa útil como a resposta.

- Que proporção da variabilidade total é explicada por esse modelo?
- Construa um gráfico de probabilidade normal dos resíduos. Que conclusão você pode tirar desse gráfico?
- Plote os resíduos contra  $\hat{y}$ , contra cada regressor e comente a adequação do modelo.
- Calcule a distância de Cook para as observações nesse conjunto de dados. Existe alguma observação influente nesses dados?

12.5.6 Considere o modelo de regressão para os dados da LAH no Exercício 12.1.12.

- Ajuste um modelo usando  $GF$  como o único regressor.
- Quanta variabilidade é explicada por esse modelo?
- Plote os resíduos contra  $\hat{y}$  e comente a adequação do modelo.
- Plote os resíduos do item (a) contra  $GFPIN$ , os pontos obtidos durante o jogo em vantagem. Isso indica que o modelo seria melhor se essa variável fosse incluída?

12.5.7 Considere os dados de desgaste no Exercício 12.1.13.

- Encontre o valor de  $R^2$ , quando o modelo usa os regressores  $x_1$  e  $x_2$ .
- O que acontece com o valor de  $R^2$ , quando um termo de interação  $x_1x_2$  for adicionado ao modelo? Isso implica necessariamente que a adição de um termo de interação seja uma boa ideia?

## Exercícios para a Seção 12.6

12.6.1 Um artigo intitulado "A Method for Improving the Accuracy of Polynomial Regression Analysis" publicado no *Journal of Quality Technology* (1971, Vol. 3(4), pp. 149-155) reportou os seguintes dados sobre  $y$  = tensão cisalhante máxima (psi) de uma borracha e  $x$  = temperatura de cura ( $^{\circ}F$ ).

$y$	770	800	840	810
$x$	280	284	292	295
$y$	735	640	590	560
$x$	298	305	308	315

- Ajuste um polinômio de segundo grau a esses dados.
- Teste a significância da regressão, usando  $\alpha = 0,05$ .
- Teste a hipótese de que  $\beta_{11} = 0$ , usando  $\alpha = 0,05$ .
- Calcule os resíduos do item (a) e use-os para avaliar a adequação do modelo.

12.6.2 Considere os seguintes dados, que resultam de um experimento para determinar o efeito de  $x$  = tempo de teste, em horas, em uma temperatura particular, sobre  $y$  = mudança na viscosidade do óleo:

$y$	-1,42	-1,39	-1,55	-1,89	-2,43
$x$	,25	,50	,75	1,00	1,25
$y$	-3,15	-4,05	-5,15	-6,43	-7,89
$x$	1,50	1,75	2,00	2,25	2,50

- Ajuste um polinômio de segundo grau a esses dados.
- Teste a significância da regressão, usando  $\alpha = 0,05$ .
- Teste a hipótese de que  $\beta_{11} = 0$ , usando  $\alpha = 0,05$ .
- Calcule os resíduos do item (a) e use-os para avaliar a adequação do modelo.

12.6.3 Os dados mostrados a seguir foram coletados durante um experimento para determinar a variação na eficiência de impulsão ( $y$ , em porcentagem) à medida que varia o ângulo de divergência do bocal de um foguete ( $x$ ):

$y$	24,60	24,71	23,90	39,50	39,60	57,12
$x$	4,0	4,0	4,0	5,0	5,0	6,0
$y$	67,11	67,24	67,15	77,87	80,11	84,67
$x$	6,5	6,5	6,75	7,0	7,1	7,3

- Teste a significância da regressão, usando  $\alpha = 0,05$ .
- Teste a significância da regressão e a falta de ajuste, usando  $\alpha = 0,05$ .
- Teste a hipótese de que  $\beta_{11} = 0$ , usando  $\alpha = 0,05$ .
- Plote os resíduos e comente a adequação do modelo.
- Ajuste um modelo cúbico e teste a significância do termo cúbico, usando  $\alpha = 0,05$ .

**12.6.4** Um artigo publicado no *Journal of Pharmaceuticals Sciences* ["Statistical Analysis of the Extended Hansen Method Using the Bootstrap Technique" (1991, Vol. 80, pp. 971-977)] apresentou dados sobre a solubilidade observada (expressa em fração molar) de um soluto, a uma temperatura constante, e sobre os parâmetros de solubilidade do modelo de Hansen, referentes à dispersão, ao dipolo e às ligações de hidrogênio. Os dados são mostrados na tabela seguinte, em que  $y$  é o logaritmo negativo da solubilidade,  $x_1$  é a solubilidade causada pela dispersão,  $x_2$  é a solubilidade em razão do dipolo, e  $x_3$  é a solubilidade em razão da ligação de hidrogênio.

Número de Observação	$y$	$x_1$	$x_2$	$x_3$
1	0,22200	7,3	0,0	0,0
2	0,39500	8,7	0,0	0,3
3	0,42200	8,8	0,7	1,0
4	0,43700	8,1	4,0	0,2
5	0,42800	9,0	0,5	1,0
6	0,46700	8,7	1,5	2,8
7	0,44400	9,3	2,1	1,0
8	0,37800	7,6	5,1	3,4
9	0,49400	10,0	0,0	0,3
10	0,45600	8,4	3,7	4,1
11	0,45200	9,3	3,6	2,0
12	0,11200	7,7	2,8	7,1
13	0,43200	9,8	4,2	2,0
14	0,10100	7,3	2,5	6,8
15	0,23200	8,5	2,0	6,6
16	0,30600	9,5	2,5	5,0
17	0,09230	7,4	2,8	7,8
18	0,11600	7,8	2,8	7,7
19	0,07640	7,7	3,0	8,0
20	0,43900	10,3	1,7	4,2
21	0,09440	7,8	3,3	8,5
22	0,11700	7,1	3,9	6,6
23	0,07260	7,7	4,3	9,5
24	0,04120	7,4	6,0	10,9
25	0,25100	7,3	2,0	5,2
26	0,00002	7,6	7,8	20,7

- Ajuste o modelo  $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \varepsilon$ .
- Teste a significância da regressão, usando  $\alpha = 0,05$ .
- Plote os resíduos e comente a adequação dos modelos.
- Use o método da soma extra dos quadrados para testar a contribuição dos termos de segundo grau, usando  $\alpha = 0,05$ .

**12.6.5** Considere os dados de potência elétrica no Exercício 12.1.6. Construa modelos de regressão para os dados, usando as seguintes técnicas:

- Todas as regressões possíveis. Encontre as equações que minimizam  $C_p$  e  $MQ_E$ .
- Regressão em etapas.
- Seleção progressiva.
- Eliminação regressiva.
- Comente os vários modelos obtidos. Que modelo você prefere?

**12.6.6** Considere os dados de inspeção de raios X no Exercício 12.1.7. Use dose de radiação como a resposta. Construa modelos de regressão, usando as seguintes técnicas:

- Todas as regressões possíveis.
- Regressão em etapas.
- Seleção progressiva.
- Eliminação regressiva.
- Comente os vários modelos obtidos. Que modelo você prefere? Por quê?

**12.6.7** Considere os dados de resistência à tração de um fio no Exercício 12.1.8. Construa modelos de regressão para os dados, usando os seguintes métodos:

- Todas as regressões possíveis. Encontre as equações que minimizam  $C_p$  e  $MQ_E$ .
- Regressão em etapas.
- Seleção progressiva.
- Eliminação regressiva.
- Comente os vários modelos obtidos. Que modelo você prefere?

**12.6.8** Considere o ajuste de um modelo de regressão aos dados da mistura de carvão e calcário no Exercício 12.1.9. Use densidade como a resposta. Construa modelos de regressão para os dados, usando as seguintes técnicas:

- Todas as regressões possíveis.
- Regressão em etapas.
- Seleção progressiva.
- Eliminação regressiva.
- Comente os vários modelos obtidos. Que modelo você prefere? Por quê?

**12.6.9** Considere os dados de extração de nisina no Exercício 12.1.10. Construa modelos de regressão para os dados, usando as seguintes técnicas:

- Todas as regressões possíveis.
- Regressão em etapas.
- Seleção progressiva.
- Eliminação regressiva.
- Comente os vários modelos obtidos. Que modelo você prefere? Por quê?

**12.6.10** Considere os dados de modulação na faixa cinza no Exercício 12.1.11. Use a faixa útil como a resposta. Construa modelos de regressão para os dados, usando as seguintes técnicas:

- Todas as regressões possíveis.
- Regressão em etapas.
- Seleção progressiva.
- Eliminação regressiva.
- Comente a respeito dos vários modelos obtidos. Que modelo você prefere? Por quê?

**12.6.11** Use os dados do desempenho dos times da LAH, no Exercício 12.1.12. Construa modelos de regressão para esses dados, com os regressores  $GF$  até  $JMP$ , usando os seguintes métodos:

- Todas as regressões possíveis. Encontre as equações que minimizam  $C_p$  e  $MQ_E$ .
- Regressão em etapas.
- Seleção progressiva.
- Eliminação regressiva.
- Que modelo você prefere?

**12.6.12** Quando ajustamos modelos de regressão polinomial, frequentemente subtraímos  $\bar{x}$  de cada valor  $x$  para produzir um regressor “centralizado”  $x' = x - \bar{x}$ . Isso reduz o efeito das dependências entre os termos do modelo e frequentemente conduz a estimativas mais exatas dos coeficientes de regressão. Usando os dados do Exercício 12.6.1, ajuste o modelo  $Y = \beta_0^* + \beta_1^*x' + \beta_{11}^*(x')^2 + \varepsilon$ .

- Use os resultados para estimar os coeficientes no modelo não centralizado  $Y = \beta_0 + \beta_1x + \beta_{11}x^2 + \varepsilon$ . Preveja  $y$  quando  $x = 285^\circ\text{F}$ . Suponha que usemos a variável padronizada  $x' = (x - \bar{x})/s_x$ , em que  $s_x$  é o desvio-padrão de  $x$ , na construção de um modelo de regressão polinomial. Ajuste o modelo  $Y = \beta_0^* + \beta_1^*x' + \beta_{11}^*(x')^2 + \varepsilon$ .
- Qual é o valor de  $y$  previsto quando  $x = 285^\circ\text{F}$ ?
- Estime os coeficientes de regressão no modelo não padronizado  $Y = \beta_0 + \beta_1x + \beta_{11}x^2 + \varepsilon$ .
- O que você pode dizer acerca da relação entre  $SQ_E$  e  $R^2$  para os modelos padronizado e não padronizado?
- Suponha que a variável  $y' = (y - \bar{y})/s_y$  seja usada no modelo, juntamente com  $x'$ . Ajuste o modelo e comente a relação entre  $SQ_E$  e  $R^2$  nos modelos padronizado e não padronizado.

**12.6.13** Considere os dados no Exercício 12.6.4. Use todos os termos do modelo quadrático completo como regressores candidatos.

- Use a seleção progressiva para identificar o modelo.
- Use a eliminação regressiva para identificar o modelo.
- Compare os dois modelos obtidos nos itens (a) e (b). Qual modelo você prefere e por quê?

**12.6.14** Usamos uma amostra de 30 observações para ajustar um modelo de regressão. O modelo completo tem nove regressores, sendo a variância estimada  $\hat{\sigma}^2 = MQ_E = 100$  e  $R^2 = 0,92$ .

- Calcule a estatística  $F$  para testar a significância de regressão. Usando  $\alpha = 0,05$ , o que você concluiria?
- Suponha que ajustemos outro modelo usando somente quatro dos regressores originais e que a soma dos quadrados do erro para esse novo modelo seja 2200. Encontre a estimativa de  $\sigma^2$  para esse novo modelo reduzido. Você concluiria que o modelo reduzido é superior ao velho? Por quê?
- Encontre o valor de  $C_p$  para o modelo reduzido no item (b). Você concluiria que o modelo reduzido é melhor do que o velho?

**12.6.15** Uma amostra de 25 observações é usada para ajustar um modelo de regressão com sete variáveis. A estimativa de  $\sigma^2$  para esse modelo completo é  $MQ_E = 10$ .

- Um algoritmo de seleção progressiva colocou três dos sete regressores originais no modelo. A soma dos quadrados do erro para o modelo com três variáveis é  $SQ_E = 300$ . Baseado no valor de  $C_p$ , você concluiria que o modelo com três variáveis tem qual-quer tendenciosidade restante?
- Depois de olhar o modelo de seleção progressiva no item (a), suponha que você pudesse adicionar mais um regressor ao modelo. Esse regressor reduzirá a soma dos quadrados do erro para 275. A adição dessa variável melhorará o modelo? Por quê?

## Exercícios Suplementares para o Capítulo 12

**12.S8** Considere a saída computacional a seguir.

A equação de regressão é

$$Y = 517 + 11,5 x_1 - 8,14 x_2 + 10,9 x_3$$

Preditor	Coef.	EP do Coef.	T	P
Constante	517,46	11,76	?	?
$x_1$	11,4720	?	36,50	?
$x_2$	-8,1378	0,1969	?	?
$x_3$	10,8565	0,6652	?	?
$S = 10,2560$	$R^2 = ?$	$R^2(\text{ajustado}) = ?$		

Análise de Variância

Fonte	GL	SQ	MQ	F	P
Regressão	?	347300	115767	?	?
Erro residual	16	?	105		
Total	19	348983			

- Preencha os valores que faltam. Use os limites para os valores  $P$ .
- O modelo é significativo para  $\alpha = 0,05$ ? É significativo para  $\alpha = 0,01$ ?
- Discuta a contribuição dos regressores individuais para o modelo.

**12.S9** Considere os dados do inversor no Exercício 12.S10. Remova a observação 2 dos dados originais. Defina as novas variáveis como segue:  $y^* = \ln y$ ,  $x_1^* = 1/\sqrt{x_1}$ ,  $x_2^* = \sqrt{x_2}$ ,  $x_3^* = 1/\sqrt{x_3}$  e  $x_4^* = \sqrt{x_4}$ .

- Ajuste o modelo de regressão, usando esses regressores transformados (não use  $x_5$  ou  $x_6$ ).
- Teste o modelo para a significância da regressão, usando  $\alpha = 0,05$ . Use o teste  $t$  para investigar a contribuição de cada variável ao modelo ( $\alpha = 0,05$ ). Quais são as suas conclusões?
- Plote os resíduos contra  $\hat{y}^*$  e contra cada um dos regressores transformados. Comente esses gráficos.

**12.S10** Os pontos transientes de um inversor eletrônico são influenciados por muitos fatores. A tabela a seguir apresenta os dados sobre o ponto transiente ( $y$ , em volts) dos inversores PMOS-NMOS e cinco regressores candidatos:  $x_1$  = largura do dispositivo NMOS,  $x_2$  = comprimento do dispositivo NMOS,  $x_3$  = largura do dispositivo PMOS,  $x_4$  = comprimento do dispositivo PMOS e  $x_5$  = temperatura ( $^\circ\text{C}$ ).

Número de Observação	Número de					$y$
	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	
1	3	3	3	3	0	0,787
2	8	30	8	8	0	0,293
3	3	6	6	6	0	1,710
4	4	4	4	12	0	0,203
5	8	7	6	5	0	0,806
6	10	20	5	5	0	4,713
7	8	6	3	3	25	0,607
8	6	24	4	4	25	9,107
9	4	10	12	4	25	9,210

(continua)

(continuação)

Número de Observação	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$y$
10	16	12	8	4	25	1,365
11	3	10	8	8	25	4,554
12	8	3	3	3	25	0,293
13	3	6	3	3	50	2,252
14	3	8	8	3	50	9,167
15	4	8	4	8	50	0,694
16	5	2	2	2	50	0,379
17	2	2	2	3	50	0,485
18	10	15	3	3	50	3,345
19	15	6	2	3	50	0,208
20	15	6	2	3	75	0,201
21	10	4	3	3	75	0,329
22	3	8	2	2	75	4,966
23	6	6	6	4	75	1,362
24	2	3	8	6	75	1,515
25	3	3	8	8	75	0,751

- a. Ajuste o modelo de regressão linear múltipla a esses dados, usando todos os regressores. Teste a significância da regressão, usando  $\alpha = 0,01$ . Encontre o valor  $P$  para esse teste e use-o para tirar suas conclusões.
- b. Teste a contribuição de cada variável ao modelo, usando o teste  $t$ , com  $\alpha = 0,05$ . Quais são as suas conclusões?
- c. Remova  $x_5$  do modelo. Teste o novo modelo com relação à significância da regressão. Teste também, com o teste  $t$ , a contribuição relativa de cada regressor para o novo modelo. Usando  $\alpha = 0,05$ , quais são as suas conclusões?
- d. Note que a  $MQ_E$  para o modelo no item (c) é menor do que a  $MQ_E$  para o modelo completo no item (a). Explique por que isso ocorreu.
- e. Calcule os resíduos na forma de Student. Algum desses resíduos parece excepcionalmente grande?
- f. Suponha que você descubra que a segunda observação estava registrada incorretamente. Remova essa observação e reajuste o modelo usando  $x_1$ ,  $x_2$ ,  $x_3$  e  $x_4$  como regressores. Note que  $R^2$  para esse modelo é consideravelmente maior do que  $R^2$  para os modelos ajustados previamente. Explique por que o  $R^2$  para esse modelo aumentou.
- g. Teste o modelo do item (f) com relação à significância da regressão, usando  $\alpha = 0,05$ . Investigue também a contribuição de cada regressor para o modelo, usando o teste  $t$  com  $\alpha = 0,05$ . Que conclusões você pode tirar?
- h. Plote os resíduos do modelo no item (f) contra  $\hat{y}$  e contra cada um dos regressores  $x_1$ ,  $x_2$ ,  $x_3$  e  $x_4$ . Comente esses gráficos.
- 12.S11 Um modelo de regressão múltipla foi usado para relacionar  $y$  = viscosidade de um produto químico com  $x_1$  = temperatura e  $x_2$  = tempo de reação. O conjunto de dados consistiu em  $n = 15$  observações.
- a. Os coeficientes estimados de regressão foram  $\hat{\beta}_0 = 300,00$ ,  $\hat{\beta}_1 = 0,85$  e  $\hat{\beta}_2 = 10,40$ . Calcule uma estimativa da viscosidade média, quando  $x_1 = 100$  °F e  $x_2 = 2$  horas.

- b. As somas dos quadrados foram  $SQ_T = 1230,50$  e  $SQ_E = 120,30$ . Teste a significância da regressão, usando  $\alpha = 0,05$ . Que conclusão você pode tirar?
- c. Que proporção da variabilidade total na viscosidade é considerada pelas variáveis nesse modelo?
- d. Suponha que outro regressor,  $x_3$  = taxa de agitação, seja adicionado ao modelo. O novo valor da soma dos quadrados do erro é  $SQ_E = 117,20$ . A adição da nova variável resultou em um valor menor de  $MQ_E$ ? Discuta a significância desse resultado.
- e. Calcule uma estatística  $F$  para verificar a contribuição de  $x_3$  para o modelo. Usando  $\alpha = 0,05$ , que conclusões você pode alcançar?

12.S12 Considere os dados do inversor eletrônico nos Exercícios 12.S9 e 12.S10. Defina a variável de resposta e as variáveis regressoras como no Exercício 12.S9 e remova a segunda observação da amostra.

- a. Use o método de todas as regressões possíveis para encontrar a equação que minimiza  $C_p$ .
- b. Use o método de todas as regressões possíveis para encontrar a equação que minimiza  $MQ_E$ .
- c. Use a regressão em etapas para selecionar um modelo de regressão para o subconjunto.
- d. Compare os modelos que você obteve.

12.S13 Um artigo publicado no *Journal of the American Ceramics Society* ["In Situ Vitrification and the Effects of Soil Additives – A Mixture Experiment Case Study" (1992, Vol. 75, pp. 112-116)] descreveu um processo para imobilizar lixos químicos ou nucleares em solo, pela dissolução do solo contaminado em um bloco de vidro. Os autores misturam CaO e Na<sub>2</sub>O com solo e modelam a viscosidade e a condutividade elétrica. O modelo da condutividade elétrica envolve seis regressores, com a amostra consistindo em  $n = 14$  observações.

- a. Para o modelo com seis regressores, suponha que  $SQ_T = 0,50$  e  $R^2 = 0,94$ . Encontre  $SQ_E$  e  $SQ_R$  e use essa informação para testar a significância da regressão, com  $\alpha = 0,05$ . Quais são suas conclusões?
- b. Suponha que um dos regressores originais seja removido do modelo, resultando em  $R^2 = 0,92$ . O que você pode concluir acerca da contribuição da variável que foi removida? Responda essa pergunta calculando uma estatística  $F$ .
- c. A remoção da variável regressora no item (b) resulta em um valor menor de  $MQ_E$  para o modelo com cinco variáveis, em comparação ao modelo original com seis variáveis? Comente a significância de sua resposta.

12.S14 O Exercício 12.1.5 introduziu dados de uma pesquisa sobre a satisfação de pacientes em hospitais. Uma das variáveis nesse conjunto de dados é uma variável categórica que indica se o paciente é um caso médico ou cirúrgico. Ajuste um modelo incluindo essa variável indicativa aos dados, usando todos os outros três regressores. Há alguma evidência de que o serviço em que o paciente se encontra (médico versus cirúrgico) tenha um impacto na satisfação reportada?

12.S15 Considere a matriz inversa do modelo, mostrada a seguir.

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 0,125 & 0 & 0 & 0 \\ 0 & 0,125 & 0 & 0 \\ 0 & 0 & 0,125 & 0 \\ 0 & 0 & 0 & 0,125 \end{bmatrix}$$

- a. Quantos regressores estão nesse modelo?
- b. Qual é o tamanho da amostra?
- c. Note a estrutura especial da diagonal da matriz. O que isso diz a você sobre as colunas na matriz original  $\mathbf{X}$ ?

# Capítulo 13 Exercícios

## Exercícios para a Seção 13.2

13.2.1 Considere a saída computacional a seguir.

Fonte	GL	SQ	MQ	F	Valor P
Fator	?	117,4	39,1	?	?
Erro	16	396,8	?		
Total	19	514,2			

- Quantos níveis do fator foram usados nesse experimento?
- Quantas réplicas a pessoa usou?
- Complete a informação que falta na tabela de ANOVA. Use os limites para o valor  $P$ .
- Que conclusões você pode tirar sobre as diferenças entre as médias dos fatores?

13.2.2 Considere a saída computacional a seguir.

Fonte	GL	SQ	MQ	F	Valor P
Fator	?	?	330,4716	4,42	?
Erro	?	?	?		
Total	31	?			

- Quantas réplicas a pessoa usou?
- Complete a informação que falta na tabela de ANOVA. Use os limites para o valor  $P$ .
- Que conclusões você pode tirar sobre as diferenças entre as médias dos fatores?

13.2.3 No livro *Planejamento e Análise de Experimentos (Design and Analysis of Experiments)*, 9ª edição (Wiley, 2017), D. C. Montgomery descreveu um experimento em que um fabricante estava interessado na resistência à tração de uma fibra sintética. Suspeitava-se que a resistência estivesse relacionada com a porcentagem do algodão na fibra. Cinco níveis de porcentagem de algodão foram usados e cinco replicatas foram corridas em uma ordem aleatória, resultando nos dados a seguir.

Porcentagem de Algodão	Resistência à Tração				
15	7	7	15	11	9
20	12	17	12	18	18
25	14	18	18	19	19
30	19	25	22	19	23
35	7	10	11	15	11

- A porcentagem de algodão afetou a resistência à ruptura do fio? Desenhe diagramas de caixas comparativos e faça uma análise de variância. Use  $\alpha = 0,05$ .
- Plote a resistência média à tração contra a porcentagem de algodão e interprete os resultados.
- Analise os resíduos e comente a adequação do modelo.

13.2.4 Um artigo publicado na *Nature* descreve um experimento para investigar o efeito de consumir chocolate sobre a saúde cardiovascular (“*Plasma Antioxidants from Chocolate*”, 2003, Vol. 424, p. 1013). O experimento consistiu em usar três diferentes tipos de chocolate: 100 g de chocolate amargo (CA), 100 g de chocolate amargo com

200 mL de chocolate ao leite integral (CA+CL) e 200 g de chocolate ao leite (CL). Doze pessoas foram usadas: sete mulheres e cinco homens, com idade média de  $32,2 \pm 1$  anos, peso médio de  $65,8 \pm 3,1$  kg e índice de massa corporal de  $21,9 \pm 0,4$  kg m<sup>-2</sup>. Em dias diferentes, uma pessoa consumiu um dos níveis de chocolate, e uma hora mais tarde a capacidade total de oxidação de seu sangue foi medida em um experimento. Dados similares àqueles resumidos no artigo são mostrados a seguir.

CA		CA + CL		CL	
118,8	115,8	105,4	100,0	102,1	102,8
122,6	115,1	101,1	99,8	105,8	98,7
115,6	116,9	102,7	102,6	99,6	94,7
113,6	115,4	97,1	100,9	102,7	97,8
119,5	115,6	101,9	104,5	98,8	99,7
115,9	107,9	98,9	93,5	100,9	98,6

- Construa diagramas de caixa (*box plot*) comparativos e estude os dados. Que impressão visual você tem a partir do exame desses diagramas?
- Analise os dados experimentais usando uma ANOVA. Se  $\alpha = 0,05$ , que conclusões você poderia tirar? O que você concluiria se  $\alpha = 0,01$ ?
- Existe alguma evidência de que o chocolate escuro aumenta a capacidade antioxidante média do sangue das pessoas?
- Analise os resíduos desse experimento.

13.2.5 No artigo “Orthogonal Design for Process Optimization and Its Application to Plasma Etching” (*Solid State Technology*, maio de 1987), G. Z. Yin e D. W. Jillie descreveram um experimento para determinar o efeito da taxa de escoamento de C<sub>2</sub>F<sub>6</sub> sobre a uniformidade do ataque químico em uma pastilha de silicone usada na fabricação de um circuito integrado. Três taxas de escoamento são usadas no experimento, e a uniformidade (em porcentagem) resultante, para seis replicatas, é mostrada a seguir.

Escoamento de C <sub>2</sub> F <sub>6</sub> (cm <sup>3</sup> /min)	Uniformidade (%)					
	125	2,7	4,6	2,6	3,0	3,2
160	4,9	4,6	5,0	4,2	3,6	4,2
200	4,6	3,4	2,9	3,5	4,1	5,1

- A taxa de escoamento de C<sub>2</sub>F<sub>6</sub> afeta a uniformidade do ataque químico? Construa diagramas de caixa para comparar os níveis do fator e faça uma análise de variância. Use  $\alpha = 0,05$ .
- Os resíduos indicam qualquer problema com as suposições propostas?

13.2.6 Um artigo publicado em *Environment International* [“Influence of Water Temperature and Shower Head Orifice Size on the Release of Radon During Showering” (1992, Vol. 18(4), pp. 363-369)] descreveu um experimento em que se investigou a quantidade de radônio liberado em chuveiros. Água enriquecida com radônio foi usada no experimento e seis tipos diferentes de diâmetros do orifício foram testados nas cabeças do chuveiro. Os dados do experimento são mostrados na seguinte tabela:

Diâmetro do Orifício	Radônio Liberado (%)			
0,37	80	83	83	85
0,51	75	75	79	79
0,71	74	73	76	77
1,02	67	72	74	74
1,40	62	62	67	69
1,99	60	61	64	66

- O tamanho do orifício afeta a percentagem média de radônio liberado? Use  $\alpha = 0,05$ .
- Encontre o valor  $P$  para a estatística  $F$  calculada no item (a).
- Análise os resíduos provenientes desse experimento.
- Encontre um intervalo de confiança de 95 % para a percentagem média de radônio liberado, quando o diâmetro do orifício for 1,40.

13.2.7 Um artigo publicado no *ACI Materials Journal* (1987, Vol. 84, pp. 213-216) descreveu vários experimentos investigando a vibração da massa de concreto, a fim de remover o ar preso. Um cilindro de 3 polegadas  $\times$  6 polegadas foi usado, e o número de vezes que esse bastão foi usado é a variável de planejamento. A resposta é a resistência compressiva resultante do corpo de prova de concreto. Os dados são mostrados na seguinte tabela:

Nível de Giro	Resistência à Compressão		
10	1530	1530	1440
15	1610	1650	1500
20	1560	1730	1530
25	1500	1490	1510

- Há alguma diferença na resistência compressiva em razão do nível de vibração?
- Encontre o valor  $P$  para a estatística  $F$  calculada no item (a).
- Análise os resíduos provenientes do experimento. Que conclusões você pode tirar acerca das suposições propostas no modelo?

13.2.8 Um artigo publicado em *Materials Research Bulletin* [1991, Vol. 26(11), pp. 1151-1165] investigou quatro métodos diferentes de preparar o composto supercondutor  $PbMo_6S_8$ . Os autores afirmam que a presença de oxigênio durante o processo de preparação afeta a temperatura de transição,  $T_c$ , da supercondutividade do material. Os métodos de preparação 1 e 2 usam técnicas que são planejadas para eliminar a presença de oxigênio, enquanto os métodos 3 e 4 permitem a presença de oxigênio. Cinco observações de  $T_c$  (em °K) foram feitas para cada material, sendo os resultados mostrados a seguir.

Método de Preparação	Temperatura de Transição $T_c$ (°K)				
1	14,8	14,8	14,7	14,8	14,9
2	14,6	15,0	14,9	14,8	14,7
3	12,7	11,6	12,4	12,7	12,1
4	14,2	14,4	14,4	12,2	11,7

- Há qualquer evidência que confirme a afirmação de que a presença de oxigênio durante a preparação afeta a temperatura média de transição? Use  $\alpha = 0,05$ .
- Qual é o valor  $P$  para o teste  $F$  no item (a)?
- Análise os resíduos provenientes desse experimento.
- Encontre um intervalo de confiança de 95 % para a  $T_c$  média, quando o método 1 for usado para preparar o material.

13.2.9 Um trabalho publicado no *Journal of the Association of Asphalt Paving Technologists* (1990, Vol. 59) descreveu um experimento com o objetivo de determinar o efeito de bolhas de ar sobre a percentagem da resistência preservada do asfalto. Para finalidades do experimento, bolhas de ar são controladas em três níveis: baixo (2-4 %), médio (4-6 %) e alto (6-8 %). Os dados são mostrados na seguinte tabela:

Bolhas de Ar	Resistência Preservada (%)							
Baixo	106	90	103	90	79	88	92	95
Médio	80	69	94	91	70	83	87	83
Alto	78	80	62	69	76	85	69	85

- Os diferentes níveis de bolhas de ar afetam significativamente a resistência média preservada? Use  $\alpha = 0,01$ .
- Encontre o valor  $P$  para a estatística  $F$  calculada no item (a).
- Análise os resíduos provenientes desse experimento.
- Encontre um intervalo de confiança de 95 % para a resistência preservada, em que haja um nível alto de bolhas de ar.
- Encontre um intervalo de confiança de 95 % para a diferença na resistência média preservada, nos níveis baixo e alto de bolhas de ar.

13.2.10 Um artigo publicado em *Quality Engineering* ["Estimating Sources of Variation: a Case Study from Polyurethane Product Research" (1999-2000, Vol. 12, pp. 89-96)] estudou os efeitos de aditivos nas propriedades finais de polímeros. Nesse caso, aditivos à base de poliuretano foram usados como agentes endurecedores. O espaço médio entre domínios foi a medida da propriedade do polímero. Os dados são mostrados a seguir.

Nível do Endurecedor	Espaço entre Domínios (nm)					
-1	8,2	8	8,2	7,9	8,1	8
-0,75	8,3	8,4	8,3	8,2	8,3	8,1
-0,5	8,9	8,7	8,9	8,4	8,3	8,5
0	8,5	8,7	8,7	8,7	8,8	8,8
0,5	8,8	9,1	9,0	8,7	8,9	8,5
1	8,6	8,5	8,6	8,7	8,8	8,8

- Há alguma diferença no nível do endurecedor? Desenhe diagramas comparativos de caixa e faça uma análise de variância. Use  $\alpha = 0,05$ .
- Encontre o valor  $P$  do teste. Estime a variabilidade em razão do erro aleatório.
- Faça um gráfico do espaço entre domínios contra o nível de endurecedor e interprete os resultados.
- Análise os resíduos provenientes do experimento e comente a adequação do modelo.

13.2.11 Um artigo publicado no *Journal of Food Science* ["Prevention of Potato Spoilage During Storage by Chlorine Dioxide" (2001, Vol. 66(3), pp. 472-477)] estudou a deterioração de batata, baseando-se em diferentes condições de oxina acidificada (OA), que é uma mistura de cloreto e dióxido de cloro. Os dados são mostrados a seguir:

Solução de OA (ppm)	% de Deterioração		
50	100	50	60
100	60	30	30
200	60	50	29
400	25	30	15

- a. As soluções de OA diferem na porcentagem de deterioração? Use  $\alpha = 0,05$ .
- b. Encontre o valor  $P$  do teste. Estime a variabilidade em razão do erro aleatório.
- c. Faça um gráfico da deterioração contra solução de OA e interprete os resultados. Que solução OA você recomendaria para usar na prática?
- d. Analise os resíduos desse experimento.

13.2.12 Um artigo publicado em *Scientia Iranica* ["Tuning the Parameters of an Artificial Neural Network (ANN) Using Central Composite Design and Genetic Algorithm" (2011, Vol. 18(6), pp. 1600-1608)] descreveu uma série de experimentos para ajustar os parâmetros nas redes neurais artificiais. Um experimento considerou a relação entre o ajuste do modelo [medido pela raiz quadrada do erro médio quadrático (RQEMQ) em um teste separado para um conjunto de dados] e a complexidade do modelo, sendo controlada pelo número de nós em duas camadas escondidas. A seguinte tabela de dados (extraídos de um conjunto de dados muito maior) contém três diferentes ANNs: ANN1 tem 33 nós na camada 1 e 30 nós na camada 2; ANN2 tem 49 nós na camada 1 e 45 nós na camada 2; e ANN3 tem 17 nós na camada 1 e 15 nós na camada 2.

ANN1	ANN2	ANN3
0,0121	0,0031	0,1562
0,0132	0,0006	0,2227
0,0011	0	0,0953

0,0023	0	0,8911
0,0391	0,022	1,3892
0,0054	0,0019	0,0154
0,0003	0,0007	1,7916
0,0014	0	0,1992

- a. Construa um diagrama de caixa (*box plot*) para comparar os diferentes ANNs.
- b. Faça a análise de variância com  $\alpha = 0,05$ . Qual é o valor  $P$ ?
- c. Analise os resíduos do experimento.
- d. Calcule um intervalo de confiança de 95 % para RQEMQ, considerando o tipo ANN2.

Para cada um dos seguintes exercícios, use os dados prévios para completar esses itens.

- a. Use o método MDS de Fisher, com  $\alpha = 0,05$ , e determine quais são os níveis dos fatores que diferem.
- b. Use o método gráfico para comparar as médias descritas nesta seção e compare suas conclusões com aquelas do método MDS de Fisher.

- 13.2.13 Porcentagem de algodão no Exercício 13.2.3. Use  $\alpha = 0,05$ .
- 13.2.14 Tipo de chocolate no Exercício 13.2.4. Use  $\alpha = 0,05$ .
- 13.2.15 Taxa de escoamento no Exercício 13.2.5. Use  $\alpha = 0,01$ .
- 13.2.16 Método de preparação no Exercício 13.2.8. Use  $\alpha = 0,05$ .
- 13.2.17 Vazios no Exercício 13.2.9. Use  $\alpha = 0,05$ .
- 13.2.18 Endurecedor no Exercício 13.2.10. Use  $\alpha = 0,05$ .

### Exercícios para a Seção 13.3

13.3.1 Um artigo publicado no *Journal of the Electrochemical Society* [1992, Vol. 139(2), pp. 524-532] descreve um experimento para investigar a baixa pressão da deposição de vapor de polissilicone. O experimento foi executado em um reator de alta capacidade, em Sematech, em Austin, Texas. O reator tem várias posições da pastilha, e quatro dessas posições são selecionadas ao acaso. A variável de resposta é a uniformidade da espessura do filme. Três réplicas do experimento foram corridas, sendo os dados apresentados a seguir:

Posição da Pastilha	Uniformidade		
1	2,76	5,67	4,49
2	1,43	1,70	2,19
3	2,34	1,97	1,47
4	0,94	1,36	1,65

- a. Existe uma diferença nas posições da pastilha? Use  $\alpha = 0,05$ .
- b. Estime a variabilidade em razão das posições da pastilha.
- c. Estime a componente do erro aleatório.
- d. Analise os resíduos desse experimento e comente a adequação do modelo.

13.3.2 Um moinho têxtil tem um grande número de teares. Supõe-se que cada tear forneça a mesma quantidade de tecido por minuto. Para investigar essa suposição, cinco teares são escolhidos ao acaso e suas produções são medidas em diferentes tempos. Os seguintes dados são obtidos:

Tear	Produção (lb/min)				
1	4,0	4,1	4,2	4,0	4,1
2	3,9	3,8	3,9	4,0	4,0
3	4,1	4,2	4,1	4,0	3,9
4	3,6	3,8	4,0	3,9	3,7
5	3,8	3,6	3,9	3,8	4,0

- a. Os teares são similares com relação à produção? Use  $\alpha = 0,05$ .
- b. Estime a variabilidade entre os teares.
- c. Estime a variância do erro experimental.
- d. Analise os resíduos desse experimento e verifique a adequação do modelo.

13.3.3 No livro *Bayesian Inference in Statistical Analysis* (Wiley, 1973), de Box e Tiao, o rendimento total do produto foi determinado para cinco amostras selecionadas aleatoriamente, provenientes de cada uma das seis bateladas escolhidas de matéria-prima.

Batelada	Rendimento (em gramas)				
1	1545	1440	1440	1520	1580
2	1540	1555	1490	1560	1495
3	1595	1550	1605	1510	1560
4	1445	1440	1595	1465	1545
5	1595	1630	1515	1635	1625
6	1520	1455	1450	1480	1445

- a. As diferentes bateladas de matéria-prima afetam significativamente o rendimento médio? Use  $\alpha = 0,01$ .
- b. Estime a variabilidade entre as bateladas.
- c. Estime a variabilidade entre amostras intrabateladas.
- d. Analise os resíduos desse experimento e verifique a adequação do modelo.

13.3.4 Um artigo publicado no *Journal of Quality Technology* [1981, Vol. 13(2), pp. 111-114] descreveu um experimento que investiga os efeitos, no brilho da polpa, de quatro produtos químicos de branqueamento. Esses quatro produtos químicos foram selecionados ao acaso, a partir de uma população grande de agentes potenciais para o branqueamento. Os dados são mostrados a seguir.

Produto Químico	Brilho da Polpa				
1	77,199	74,466	92,746	76,208	82,876
2	80,522	79,306	81,914	80,346	73,385
3	79,417	78,017	91,596	80,802	80,626
4	78,001	78,358	77,544	77,364	77,386

- a. Existe uma diferença nos tipos de produtos químicos? Use  $\alpha = 0,05$ .

- b. Estime a variabilidade em razão dos tipos de produtos químicos.
- c. Estime a variabilidade em razão do erro aleatório.
- d. Analise os resíduos desse experimento e comente a adequação do modelo.

13.3.5 Considere o experimento da deposição de vapor descrito no Exercício 13.3.1

- a. Estime a variabilidade total na resposta uniformidade.
- b. Quanto da variabilidade total na resposta uniformidade é em razão da diferença entre as posições no reator?
- c. Para qual nível a variabilidade na resposta uniformidade poderia ser reduzida, se a variabilidade entre as posições no reator pudesse ser eliminada? Você acredita que essa redução seja significativa?

13.3.6 Considere o experimento do tecido descrito no Exercício 13.3.2.

- a. Estime a variabilidade total na resposta produção.
- b. Quanto da variabilidade total na resposta produção é em razão da diferença entre os teares?
- c. Para qual nível a variabilidade na resposta produção poderia ser reduzida, se a variabilidade entre os teares pudesse ser eliminada? Você acredita que essa redução seja significativa?

## Exercícios para a Seção 13.4

13.4.1 Considere a seguinte saída computacional para PBCA.

Fonte	GL	SQ	MQ	F	P
Fator	?	193,800	64,600	?	?
Bloco	3	464,218	154,739		
Erro	?	?	4,464		
Total	15	698,190			

- a. Quantos níveis do fator foram usados nesse experimento?
- b. Quantos blocos foram usados nesse experimento?
- c. Complete a informação que falta. Use limites para o valor  $P$ .
- d. Que conclusões você tiraria se  $\alpha = 0,05$ ? O que você concluiria se  $\alpha = 0,01$ ?

13.4.2 O Exercício 13.2.4 introduziu você a um experimento para investigar o efeito potencial de consumir chocolate sobre a saúde cardiovascular. O experimento foi conduzido como um planejamento completamente aleatorizado; o exercício lhe pediu para usar a ANOVA de modo a analisar os dados e tirar conclusões. Agora, considere que o experimento tenha sido conduzido como um PBCA, com as pessoas tendo sido consideradas como blocos. Analise os dados usando essa suposição. Que conclusões você tiraria (usando  $\alpha = 0,05$ ) acerca do efeito dos diferentes tipos de chocolate sobre a saúde cardiovascular? Sua conclusão mudaria se  $\alpha = 0,01$ ?

13.4.3 No artigo intitulado “The Effect of Nozzle Design on the Stability and Performance of Turbulent Water Jets” [*Fire Safety Journal*, agosto de 1981, Vol. 4(1), pp. 1-13], C. Theobald descreveu um experimento em que uma medida da forma foi determinada para vários tipos diferentes de bocais, com níveis diferentes de velocidade do jato de saída. O interesse nesse experimento está principalmente no tipo de bocal, sendo a velocidade um fator de ruído. Os dados são apresentados a seguir:

Tipo de Bocal	Velocidade do Jato de Saída (m/s)					
	11,73	14,37	16,59	20,43	23,46	28,74
1	0,78	0,80	0,81	0,75	0,77	0,78
2	0,85	0,85	0,92	0,86	0,81	0,83
3	0,93	0,92	0,95	0,89	0,89	0,83
4	1,14	0,97	0,98	0,88	0,86	0,83
5	0,97	0,86	0,78	0,76	0,76	0,75

- a. O tipo de bocal afeta a medida da forma? Compare os bocais, usando os diagramas de caixas e a análise de variância.
- b. Use o método MSD de Fisher para determinar diferenças específicas entre os bocais. Um gráfico da média (ou desvio-padrão) das medidas da forma *versus* o tipo de bocal ajuda nas conclusões?
- c. Analise os resíduos desse experimento.

13.4.4 Um artigo publicado em *Quality Engineering* [“Designed Experiment to Stabilize Blood Glucose Levels” (1999-2000, Vol. 12, pp. 83-87)] descreveu um experimento para minimizar variações em níveis de glicose no sangue. O tratamento foi o tempo (10 ou 20 minutos) de exercício de um esquiador de competições. Para o experimento, utilizou-se o período do dia como a variável bloco. Os dados são fornecidos a seguir:

Exercício (min)	Período do Dia	Glicose Média no Sangue
10	tarde	71,5
10	manhã	103,0
20	manhã	83,5
20	tarde	126,0
10	manhã	125,5
10	tarde	129,5
20	tarde	95,0
20	manhã	93,0

- a. Há um efeito do tempo de exercício sobre a glicose média no sangue? Use  $\alpha = 0,05$ .
- b. Encontre o valor  $P$  para o teste no item (a).
- c. Analise os resíduos desse experimento.

13.4.5 Um artigo publicado no *American Industrial Hygiene Association Journal* ["A Rapid Method for the Determination of Arsenic Concentrations in Urine at Field Locations" (1976, Vol. 37(7), pp. 418-422)] descreveu um teste de campo para detectar a presença de arsênio em amostras de urina. O teste foi proposto para uso entre trabalhadores florestais, por causa do uso crescente de arsênio orgânico naquela indústria. O experimento comparou os testes feitos pelo estagiário e pelo laboratorista experiente com uma análise de um laboratório remoto. Quatro indivíduos foram selecionados para se submeterem ao teste, sendo considerados como blocos. A variável de resposta é o conteúdo (em ppm) de arsênio na urina do indivíduo. Os dados são mostrados a seguir:

Teste	Indivíduo			
	1	2	3	4
Estagiário	0,05	0,05	0,04	0,15
Analista Experiente	0,05	0,05	0,04	0,17
Laboratório	0,04	0,04	0,03	0,10

- a. Há alguma diferença no procedimento de teste do arsênio?
- b. Analise os resíduos desse experimento.

13.4.6 No livro "Planejamento e Análise de Experimentos" (*Design and Analysis of Experiments*), 9ª edição (Wiley, 2017), D. C. Montgomery descreve um experimento que determinou o efeito de quatro tipos diferentes de ponteiras em um teste de dureza de uma liga metálica. Quatro espécimes da liga foram obtidos e cada ponteira foi testada uma vez em cada espécime, produzindo os seguintes dados:

Tipo de Ponteira	Espécime			
	1	2	3	4
1	9,3	9,4	9,6	10,0
2	9,4	9,3	9,8	9,9
3	9,2	9,4	9,5	9,7
4	9,7	9,6	10,0	10,2

- a. Há alguma diferença nas medidas de dureza entre as ponteiras?
- b. Use o método MSD de Fisher para investigar diferenças específicas entre as ponteiras.
- c. Analise os resíduos desse experimento.

Tempo de Estocagem	Lote								
	1	2	3	4	5	6	7	8	9
0 dia	1694,0	989,0	917,3	346,1	1260,0	965,6	1123,0	1106,0	1116,0
7 dias	1802,0	1074,0	278,8	1375,0	544,0	672,2	818,0	406,8	461,6
14 dias	1568,0	646,2	1820,0	1150,0	983,7	395,3	422,3	420,0	409,5
21 dias	415,5	845,4	377,6	279,4	447,8	272,1	394,1	356,4	351,2

13.4.7 Um experimento foi conduzido a fim de investigar o escapamento de corrente elétrica em um dispositivo SOS MOSFETS. A finalidade do experimento foi investigar como o escapamento de corrente varia com o comprimento do canal. Quatro comprimentos diferentes do canal foram selecionados. Para cada comprimento do canal, cinco larguras diferentes foram também usadas. A largura deve ser considerada um fator de ruído. Veja os dados a seguir:

Comprimento do Canal	Largura				
	1	2	3	4	5
1	0,7	0,8	0,8	0,9	1,0
2	0,8	0,8	0,9	0,9	1,0
3	0,9	1,0	1,7	2,0	4,0
4	1,0	1,5	2,0	3,0	20,0

- a. Teste a hipótese de que o escapamento médio de voltagem não depende do comprimento do canal, usando  $\alpha = 0,05$ .
- b. Analise os resíduos desse experimento. Comente os gráficos de resíduos.
- c. O escapamento observado de voltagem para o comprimento 4 do canal e largura 5 foi registrado erroneamente. A observação correta é 4,0. Analise os dados corrigidos desse experimento. Há alguma evidência para concluir que o escapamento médio de voltagem aumenta com o comprimento do canal?

13.4.8 Um artigo publicado no *Food Technology Journal* (1956, Vol. 10, pp. 39-42) descreveu um estudo sobre o teor de protopectina em tomates, durante a estocagem. Quatro tempos de estocagem foram selecionados, e amostras de nove lotes de tomates foram analisadas. O teor (expresso como a fração solúvel de ácido clorídrico, mg/kg) de protopectina está na tabela a seguir.

- a. Nesse estudo, os pesquisadores supuseram que o teor médio de protopectina seria diferente para tempos diferentes de estocagem. Você pode confirmar essa hipótese com um teste estatístico, usando  $\alpha = 0,05$ ?
- b. Encontre o valor  $P$  para o teste no item (a).
- c. Que tempos específicos de estocagem são diferentes? Você concordaria com a afirmação de que o teor de protopectina diminui à medida que o tempo de estocagem aumenta?
- d. Analise os resíduos desse experimento.

## Exercícios Suplementares para o Capítulo 13

13.S9 Considere a saída computacional a seguir.

Fonte	GL	SQ	MQ	F	P
Fator	?	126,880	63,4401	?	?
Bloco	?	54,825	18,2751		
Erro	6	?	2,7403		
Total	11	198,147			

- Quantos níveis do fator foram usados nesse experimento?
- Quantos blocos foram usados?
- Complete a informação que falta. Use limites para o valor  $P$ .
- Que conclusões você tiraria se  $\alpha = 0,05$ ? E se  $\alpha = 0,01$ ?

13.S10 Um artigo publicado em *Lubrication Engineering* (dezembro de 1990) descreveu os resultados de um experimento planejado para investigar os efeitos das propriedades de materiais à base de carbono sobre a evolução de defeitos na superfície de lacres desses materiais. Esses lacres são usados extensivamente em equipamentos como aqueles usados na partida de turbina a ar. Cinco materiais diferentes à base de carbono foram testados, e a rugosidade da superfície foi medida. Os dados são apresentados a seguir:

Tipo de Material à Base de Carbono	Rugosidade da Superfície				
EC10	0,50	0,55	0,55	0,36	
EC10A	0,31	0,07	0,25	0,18	0,56 0,20
EC4	0,20	0,28	0,12		
EC1	0,10	0,16			

- O tipo de material à base de carbono tem um efeito na rugosidade média da superfície? Use  $\alpha = 0,05$ .
- Encontre os resíduos desse experimento. Um gráfico da probabilidade normal dos resíduos indica qualquer problema com a suposição de normalidade?
- Plote os resíduos *versus*  $\hat{y}_{ij}$ . Comente o gráfico.
- Encontre um intervalo de confiança de 95 % para a diferença entre a rugosidade média na superfície quando são usados os graus de carbono EC10 e EC1.
- Aplique o método MDS de Fisher para esse experimento. Resuma suas conclusões em relação ao efeito do tipo de material sobre a rugosidade da superfície.

13.S11 Um artigo publicado em *IEEE Transactions on Components, Hybrids, and Manufacturing Technology* [(1992, Vol. 15(2), pp. 146-153)] descreveu um experimento em que a resistência de contato de um relé foi estudada para três materiais diferentes (todos eram ligas à base de prata). Os dados são mostrados a seguir:

Liga	Resistência de Contato				
1	95	97	99	98	99
	99	99	94	95	98
2	104	102	102	105	99
	102	111	103	100	103
3	119	130	132	136	141
	172	145	150	144	135

- O tipo de liga afeta a resistência média de contato? Use  $\alpha = 0,01$ .
- Use o método MDS de Fisher para determinar qual a média que difere.
- Para a liga 3, encontre um intervalo de confiança de 99 % para a resistência média de contato.
- Análise os resíduos para esse experimento.

13.S12 Um artigo publicado no *Journal of Quality Technology* [(1982, Vol. 14(2), pp. 80-89)] descreveu um experimento em que três métodos diferentes de preparar peixe são avaliados, com base no critério sensorial, sendo dada uma pontuação de qualidade. Considere que esses métodos tenham sido selecionados aleatoriamente a partir de uma grande população de métodos de preparação. Os dados estão na tabela seguinte:

Método	Pontuação			
1	24,4	23,2	25,0	19,7
	22,2	24,4	23,8	18,0
2	22,1	19,5	17,3	19,7
	22,3	23,2	21,4	22,6
3	23,3	22,8	22,4	23,7
	20,4	23,5	20,8	24,1

- Há alguma diferença nos métodos de preparação? Use  $\alpha = 0,05$ .
- Calcule o valor  $P$  para a estatística  $F$  no item (a).
- Análise os resíduos provenientes do experimento e comente a adequação do modelo.
- Estime os componentes de variância.

13.S13 Um artigo publicado no *Journal of Agricultural Engineering Research* (1992, Vol. 52, pp. 53-76) descreveu um experimento para investigar o efeito da temperatura de secagem do grão de trigo sobre a qualidade do processo de assar pão. Três níveis diferentes de temperatura foram usados, e a variável de resposta medida foi o volume de pão produzido. Os dados são mostrados a seguir:

Temperatura (°C)	Volume (cm³)				
70,0	1245	1235	1285	1245	1235
75,0	1235	1240	1200	1220	1210
80,0	1225	1200	1170	1155	1095

- A temperatura de secagem afeta o volume médio do pão? Use  $\alpha = 0,01$ .
- Encontre o valor  $P$  para esse teste.
- Use o método MDS de Fisher para determinar quais são as médias diferentes.
- Análise os resíduos desse experimento e comente a adequação do modelo.

13.S14 Um artigo publicado em *Nature Genetics*, ["Treatment-Specific Changes in Gene Expression Discriminate in vivo Drug Response in Human Leukemia Cells", (2003, Vol. 34(1), pp. 85-90)] estudou a expressão genética como uma função de tratamentos diferentes para leucemia. Três grupos de tratamentos são: somente mercaptopurina (MP); baixa dose de metotrexato (BDM) e MP; e alta dose de metotrexato (ADM) e MP. Cada grupo era formado de dez pessoas. As respostas provenientes de um gene específico são mostradas a seguir.

Tratamentos	Expressão Genética				
	334,5	31,6	701	41,2	61,2
Somente MP	69,6	67,5	66,6	120,7	881,9
MP + ADM	919,4	404,2	1024,8	54,1	62,8
	671,6	882,1	354,2	321,9	91,1
MP + BDM	108,4	26,1	240,8	191,1	69,7
	242,8	62,7	396,9	23,6	290,4

- a. Verifique a normalidade dos dados. Você pode considerar essas amostras provenientes de populações normais?
- b. Tire o logaritmo dos dados originais e verifique a normalidade dos dados transformados. Há alguma evidência para sustentar a afirmação de que as médias dos tratamentos diferem para os dados transformados? Use  $\alpha = 0,1$ .
- c. Analise os resíduos dos dados transformados e comente a adequação do modelo.

13.S15 Um artigo publicado em *Communications of the ACM* [(1987, Vol. 30(5), pp. 53-76)] estudou diferentes algoritmos para estimar os custos de desenvolvimento de *softwares*. Seis algoritmos foram aplicados a oito projetos de desenvolvimento de programas *softwares* e observou-se o erro percentual na estimação do custo de desenvolvimento. Os dados observados estão na tabela a seguir.

- a. Os algoritmos diferem na exatidão da estimação de seus custos médios? Use  $\alpha = 0,05$ .
- b. Analise os resíduos desse experimento.
- c. Que algoritmo você recomendaria para uso na prática?

Algoritmo	Projeto							
	1	2	3	4	5	6	7	8
1(SLIM)	1244	21	82	2221	905	839	527	122
2(COCOMO-A)	281	129	396	1306	336	910	473	199
3(COCOMO-R)	220	84	458	543	300	794	488	142
4(COCOMO-C)	225	83	425	552	291	826	509	153
5(FUNCTION POINTS)	19	11	-34	121	15	103	87	-17
6(ESTIMALS)	-20	35	-53	170	104	199	142	41

13.S16 Um artigo publicado no *Journal of Hazardous Materials* ["Toxicity Assessment from Electro-Coagulation Treated-Textile Dye Waste Waters by Bioassays" (2009, Vol. 172(1), pp. 330-337)] discutiu um estudo de remoção de poluentes de água residual de tinta de indústria têxtil com uma técnica de eletrocoagulação. A demanda química de oxigênio (DQO) (uma medida comum de poluição de água) foi usada como a resposta e três valores diferentes para o tempo de eletrólise foram considerados. Os seguintes dados foram extraídos de um estudo mais amplo. Suponha que um experimento com blocos completos aleatorizados foi conduzido com três blocos baseados em valores iniciais de pH.

Tempo de Eletrólise (min)	pH Inicial		
	3	7	11
15	77,1	75,2	42,2
30	80,1	76,8	45,0
45	82,8	75,2	46,8

- a. Existe algum efeito do tempo de eletrólise com  $\alpha = 0,05$ ? Calcule o valor  $P$ .
- b. Analise os resíduos a partir do experimento.
- c. Calcule um intervalo de confiança de 95 % para a remoção média de DQO, quando o tempo de eletrólise for 15 minutos.
- d. Faça uma ANOVA considerando que todos os dados sejam coletados em um único valor de pH. Comente as diferenças em relação ao item (a).

## Capítulo 14 Exercícios

### Exercícios para a Seção 14.3

14.3.1 Um artigo publicado em *Industrial Quality Control* (1956, pp. 5-8) descreve um experimento para investigar o efeito de dois fatores (tipo de vidro e tipo de fósforo) sobre o brilho de um tubo de televisão. A variável de resposta medida é a corrente (em microampères) necessária para obter um nível especificado de brilho. Os dados são mostrados na tabela ao lado.

- a. Estabeleça as hipóteses de interesse nesse experimento.
- b. Teste as hipóteses anteriores (a) e tire conclusões, usando a análise de variância com  $\alpha = 0,05$ .
- c. Analise os resíduos desse experimento.

Tipo de Vidro	Tipo de Fósforo		
	1	2	3
1	280	300	290
	290	310	285
	285	295	290
2	230	260	220
	235	240	225
	240	235	230

14.3.2 Um engenheiro suspeita que o acabamento de uma superfície de peças metálicas seja influenciado pelo tipo de tinta usada e pelo tempo de secagem. Ele selecionou três tempos de secagem – 20, 25 e 30 minutos – e usou dois tipos de tinta. Três peças são testadas com cada combinação de tipo de tinta e tempo de secagem. Os dados são apresentados na tabela a seguir.

Tinta	Tempo de Secagem (min)		
	20	25	30
1	74	73	78
	64	61	85
	50	44	92
2	92	98	66
	86	73	45
	68	88	85

- Estabeleça as hipóteses de interesse nesse experimento.
- Teste as hipóteses anteriores (a) e tire conclusões, usando a análise de variância com  $\alpha = 0,05$ .
- Análise os resíduos a partir desse experimento.

14.3.3 Um artigo publicado em *Technometrics* ["Exact Analysis of Means with Unequal Variances" (2002, Vol. 44, pp. 152-160)] descreveu a técnica da análise de médias (ANOM) e apresentou os resultados de um experimento de isolamento. Quatro tipos de isolante foram testados em três temperaturas diferentes. Os dados são mostrados na tabela a seguir:

Isolante	Temperatura (°F)					
	1		2		3	
1	6,6	4	4,5	2,2	2,3	0,9
	2,7	6,2	5,5	2,7	5,6	4,9
	6	5	4,8	5,8	2,2	3,4
2	3	3,2	3	1,5	1,3	3,3
	2,1	4,1	2,5	2,6	0,5	1,1
	5,9	2,5	0,4	3,5	1,7	0,1
3	5,7	4,4	8,9	7,7	2,6	9,9
	3,2	3,2	7	7,3	11,5	10,5
	5,3	9,7	8	2,2	3,4	6,7
4	7	8,9	12	9,7	8,3	8
	7,3	9	8,5	10,8	10,4	9,7
	8,6	11,3	7,9	7,3	10,6	7,4

- Escreva um modelo para esse experimento.
- Teste as hipóteses apropriadas e tire conclusões, usando a análise de variância com  $\alpha = 0,05$ .
- Análise graficamente a interação.
- Análise os resíduos do experimento.
- Use o método MDS de Fisher para investigar as diferenças entre os efeitos médios do tipo de isolante. Use  $\alpha = 0,05$ .

14.3.4 Um experimento foi conduzido para determinar se a temperatura de queima ou a posição da fornalha afetam a densidade de um ânodo de carbono. Os dados são mostrados na coluna ao lado.

- Estabeleça as hipóteses de interesse.
- Teste as hipóteses anteriores (a), usando a análise de variância com  $\alpha = 0,05$ . Quais são as suas conclusões?
- Análise os resíduos desse experimento.
- Usando o método MDS de Fisher, investigue as diferenças entre a densidade média do ânodo nos três níveis diferentes de temperatura. Use  $\alpha = 0,05$ .

Posição	Temperatura (°C)		
	800	825	850
1	570	1063	565
	565	1080	510
	583	1043	590
2	528	988	526
	547	1026	538
	521	1004	532

14.3.5 Um artigo publicado em *IEEE Transactions on Electron Devices* (novembro de 1986, Vol. 33, p. 1754) descreveu um estudo sobre os efeitos de duas variáveis – tratamento com polissilicone e condições de recozimento (tempo e temperatura) – sobre a corrente de um transistor bipolar. Os dados desse experimento estão na tabela a seguir:

Tratamento com Polissilicone	Recozimento (temperatura/tempo)				
	900	900	950	1000	1000
$1 \times 10^{20}$	60	180	60	15	30
	4,40	8,30	10,15	10,29	11,01
$2 \times 10^{20}$	4,60	8,90	10,20	10,30	10,58
	3,20	7,81	9,38	10,19	10,81
	3,50	7,75	10,02	10,10	10,60

- Há qualquer evidência confirmando a afirmação de que os níveis do tratamento com polissilicone ou das condições de recozimento afetam a corrente? Essas variáveis interagem? Use  $\alpha = 0,05$ .
- Análise graficamente a interação.
- Análise os resíduos desse experimento.
- Use o método MDS de Fisher para isolar os efeitos das condições de recozimento sobre a corrente, com  $\alpha = 0,05$ .

14.3.6 Um artigo publicado no *Journal of Testing and Evaluation* (1988, Vol. 16, pp. 508-515) investigou os efeitos da frequência cíclica de carregamento e das condições ambientais sobre o crescimento da fratura por fadiga, a uma tensão constante de 22 MPa para um material particulado. Os dados do experimento são mostrados a seguir. A variável de resposta é a taxa de crescimento da fratura por fadiga.

Frequência	Ambiente		
	Ar	H <sub>2</sub> O	Água Salgada H <sub>2</sub> O
10	2,29	2,06	1,90
	2,47	2,05	1,93
	2,48	2,23	1,75
	2,12	2,03	2,06
	2,65	3,20	3,10
1	2,68	3,18	3,24
	2,06	3,96	3,98
	2,38	3,64	3,24
	2,24	11,00	9,96
	2,71	11,00	10,01
0,1	2,81	9,06	9,36
	2,08	11,30	10,40

- a. Há qualquer indicação de que os fatores afetam a taxa de crescimento da fratura? Há qualquer indicação de interação? Use  $\alpha = 0,05$ .
- b. Analise os resíduos desse experimento.
- c. Repita a análise do item (a), usando  $\ln(y)$  como resposta. Analise os resíduos dessa nova variável de resposta e comente os seus resultados.

14.3.7 Um artigo publicado em *Bioresource Technology* ["Quantitative Response of Cell Growth and Tuber Polysaccharides Biosynthesis by Medicinal Mushroom Chinese Truffle Tuber Sinense to Metal Ion in Culture Medium" (2008, Vol. 99(16), pp. 7606-7615)] descreveu um experimento para investigar o efeito da concentração de íon de metal na produção de polissacarídeos extracelulares (PEC). Suspeita-se que  $Mg^{2+}$  e  $K^+$  (em milimolar) estão relacionados com a PEC. Os dados de um planejamento fatorial completo estão dados a seguir.

Corrida	$Mg^{2+}$ (mM)	$K^+$ (mM)	PEC (g/L)
1	40	5	3,88
2	50	15	4,23
3	40	10	4,67
4	30	5	5,86
5	50	10	4,50
6	50	5	3,62
7	30	15	3,84
8	40	15	3,25
9	30	10	4,18

- a. Estabeleça as hipóteses de interesse.
- b. Teste suas hipóteses com  $\alpha = 0,05$ .
- c. Analise os resíduos e faça um gráfico dos resíduos *versus* o rendimento previsto.

### Exercícios para a Seção 14.4

14.4.1 A concentração, em porcentagem, de madeira de lei na polpa virgem, a facilidade de drenagem e o tempo de cozimento da polpa estão sendo investigados pelos seus efeitos sobre a resistência do papel. Os dados de um experimento fatorial com três fatores são mostrados na seguinte tabela.

Concentração de Madeira de Lei (%)	Tempo de Cozimento de 1,5 hora			Tempo de Cozimento de 2,0 horas		
	Facilidade de Drenagem			Facilidade de Drenagem		
	350	500	650	350	500	650
10	96,6	97,7	99,4	98,4	99,6	100,6
	96,0	96,0	99,8	98,6	100,4	100,9
15	98,5	96,0	98,4	97,5	98,7	99,6
	97,2	96,9	97,6	98,1	96,0	99,0
20	97,5	95,6	97,4	97,6	97,0	98,5
	96,6	96,2	98,1	98,4	97,8	99,8

- a. Analise os dados usando a análise de variância, considerando que todos os fatores sejam fixos. Use  $\alpha = 0,05$ .
- b. Encontre os valores  $P$  para as razões  $F$  no item (a).
- c. Os resíduos são encontrados por  $e_{ijkl} = y_{ijkl} - \bar{y}_{ijk..}$ . Analise graficamente os resíduos desse experimento.

14.4.2 Em uma fábrica de tecidos, o setor de qualidade está estudando os efeitos de vários fatores na tintura para um tecido misto de fibra de algodão e de fibra sintética, usado na fabricação de camisas.

Três operadores, três tempos de ciclo e duas temperaturas foram selecionados e três pequenos pedaços de tecido foram tingidos sob cada conjunto de condições. O tecido acabado foi comparado com um padrão, e uma escala numérica foi atribuída. Os resultados são mostrados na tabela seguinte.

Tempo do Ciclo	Temperatura					
	300°			350°		
	Operador			Operador		
	1	2	3	1	2	3
40	23	27	31	24	38	34
	24	28	32	23	36	36
	25	26	28	28	35	39
50	36	34	33	37	34	34
	35	38	34	39	38	36
	36	39	35	35	36	31
60	28	35	26	26	36	28
	24	35	27	29	37	26
	27	34	25	25	34	34

- a. Estabeleça e teste as hipóteses apropriadas, usando a análise de variância com  $\alpha = 0,05$ .
- b. Os resíduos podem ser obtidos de  $e_{ijkl} = y_{ijkl} - \bar{y}_{ijk..}$ . Analise graficamente os resíduos desse experimento.

## Exercícios para a Seção 14.5

**14.5.1** Pensa-se que quatro fatores influenciam o sabor de um refrigerante: tipo de adoçante (*A*), razão entre xarope e água (*B*), nível de carbonatação (*C*) e temperatura (*D*). Cada fator pode ser corrido em dois níveis, produzindo um planejamento  $2^4$ . Em cada corrida no planejamento, amostras de refrigerante são dadas a 20 pessoas para testar. Cada pessoa atribui uma pontuação de 1 a 10 ao refrigerante. A pontuação total é a variável de resposta, e o objetivo é encontrar uma formulação que maximize a pontuação total. Duas réplicas desse planejamento são corridas, e os resultados são mostrados na tabela. Analise os dados e tire as conclusões. Use  $\alpha = 0,05$  nos testes estatísticos.

Combinação de Tratamentos	Réplica	
	I	II
(1)	159	163
<i>a</i>	168	175
<i>b</i>	158	163
<i>ab</i>	166	168
<i>c</i>	175	178
<i>ac</i>	179	183
<i>bc</i>	173	168
<i>abc</i>	179	182
<i>d</i>	164	159
<i>ad</i>	187	189
<i>bd</i>	163	159
<i>abd</i>	185	191
<i>cd</i>	168	174
<i>acd</i>	197	199
<i>bcd</i>	170	174
<i>abcd</i>	194	198

**14.5.2** Um engenheiro está interessado no efeito da velocidade de corte (*A*), na dureza do metal (*B*) e no ângulo de corte (*C*) sobre a vida de uma ferramenta de corte. Dois níveis de cada fator são escolhidos e duas réplicas de um planejamento fatorial  $2^3$  são corridas. Os dados da vida (em horas) da ferramenta são mostrados na seguinte tabela:

Combinação de Tratamentos	Réplica	
	I	II
(1)	221	311
<i>a</i>	325	435
<i>b</i>	354	348
<i>ab</i>	552	472
<i>c</i>	440	453
<i>ac</i>	406	377
<i>bc</i>	605	500
<i>abc</i>	392	419

- Analisar os dados desse experimento.
- Encontrar um modelo apropriado de regressão, que explique a vida da ferramenta, em termos das variáveis usadas no experimento.
- Analisar os resíduos desse experimento.

**14.5.3** Um artigo publicado em *IEEE Transactions on Semiconductor Manufacturing* (1992, Vol. 5, pp. 214-222) descreve um experimento para investigar a carga na superfície de uma pastilha de silicone. Os fatores possíveis de influenciar a carga induzida na superfície são o método de limpeza (secagem giratória com enxágue ou SGE e secagem giratória ou SG) e a posição na pastilha onde a carga foi medida. Os dados de resposta da carga ( $\times 10^{11}$  C/cm<sup>3</sup>) na superfície são mostrados a seguir.

Método de Limpeza	Posição de Teste	
	E	D
SG	1,66	1,84
	1,90	1,84
SGE	1,92	1,62
	-4,21	-7,58
	-1,35	-2,20
	-2,08	-5,36

- Estimar os efeitos dos fatores.
- Que fatores parecem ser importantes? Use  $\alpha = 0,05$ .
- Analisar os resíduos desse experimento.

**14.5.4** Um artigo publicado em *Talanta* (2005, Vol. 65, pp. 895-899) apresentou um planejamento fatorial  $2^3$  (com duas réplicas) para encontrar o nível de chumbo, usando a espectrometria de absorção atômica de chama (EAAC). Os dados são mostrados na seguinte tabela.

Corrida	Fatores			Recuperação de Chumbo (%)	
	TA	pH	CR	R1	R2
1	-	-	-	39,8	42,1
2	+	-	-	51,3	48
3	-	+	-	57,9	58,1
4	+	+	-	78,9	85,9
5	-	-	+	78,9	84,2
6	+	-	+	84,2	84,2
7	-	+	+	94,4	90,9
8	+	+	+	94,7	105,3

Os fatores e níveis são mostrados na tabela seguinte.

Fator	Inferior (-)	Superior (+)
Concentração de reagente (RC) (mol l <sup>-1</sup> )	5 × 10 <sup>-6</sup>	5 × 10 <sup>-5</sup>
pH	6,0	8,0
Tempo de agitação (TA) (min)	10	30

- Construa um gráfico de probabilidade normal das estimativas de efeitos. Quais efeitos aparentam ser grandes?
- Conduza uma análise de variância para confirmar o que você achou no item (a).
- Analise os resíduos desse experimento. Existe algum problema com a adequação do modelo?

14.5.5 Um artigo publicado no *Journal of Construction Engineering and Management* [“Analysis of Earth-Moving Systems Using Discrete-Event Simulation” (1995, Vol. 121(4), pp. 388-396)] considerou um experimento fatorial com dois níveis, com réplicas, para estudar os fatores mais importantes para a saída em um sistema de terraplenagem. O planejamento fatorial 2<sup>4</sup> tem quatro réplicas e tem como resposta a taxa de produção (m<sup>3</sup>/h). Os dados são mostrados na seguinte tabela.

- Estime os efeitos dos fatores. Baseado no gráfico de probabilidade normal da estimativa dos efeitos, identifique um modelo para os dados desse experimento.
- Conduza uma ANOVA para confirmar o que você achou no item (a). Quais são suas conclusões?
- Analise os resíduos desse experimento e plote os resíduos *versus* a produção prevista.
- Comente a adequação do modelo.

Corrida	Número de Caminhões	Passagens por Carga	Tempo de Passagem da Carga	Tempo de Viagem	Produção (m <sup>3</sup> /h)			
					1	2	3	4
1	-	-	-	-	179,6	179,8	176,3	173,1
2	+	-	-	-	373,1	375,9	372,4	361,1
3	-	+	-	-	153,2	153,6	150,8	148,6
4	+	+	-	-	226,1	220,0	225,7	218,5
5	-	-	+	-	156,9	155,4	154,2	152,2
6	+	-	+	-	242,0	233,5	242,3	233,6
7	-	+	+	-	122,7	119,6	120,9	118,6
8	+	+	+	-	135,7	130,9	135,5	131,6
9	-	-	-	+	44,2	44,0	43,5	43,6
10	+	-	-	+	124,2	123,3	122,8	121,6
11	-	+	-	+	42,0	42,4	42,5	41,0
12	+	+	-	+	116,3	117,3	115,6	114,7
13	-	-	+	+	42,1	42,6	42,8	42,9
14	+	-	+	+	119,1	119,5	116,9	117,2
15	-	+	+	+	39,6	39,7	39,5	39,2
16	+	+	+	+	107,0	105,3	104,2	103,0

Nível	Número de Caminhões	Passagens por Carga	Tempo de Passagem da Carga	Tempo de Viagem
-1	2	4	12 s	100 s
1	6	7	22 s	800 s

### Exercícios para a Seção 14.6

14.6.1 Considere os dados do Exercício 14.5.2. Suponha que os dados da segunda réplica não estivessem disponíveis. Analise os dados somente da réplica I e comente o que você achou.

14.6.2 Os dados mostrados aqui representam uma única réplica de um planejamento 2<sup>5</sup> que é usado em um experimento para estudar a resistência à compressão do concreto. Os fatores são: mistura (A), tempo (B), laboratório (C), temperatura (D) e tempo de secagem (E).

- |           |            |
|-----------|------------|
| (1) = 700 | e = 800    |
| a = 900   | ae = 1200  |
| b = 3400  | be = 3500  |
| ab = 5500 | abe = 6200 |
| c = 600   | ce = 600   |
| ac = 1000 | ace = 1200 |

$bc = 3000$	$bce = 3006$
$abc = 5300$	$abce = 5500$
$d = 1000$	$de = 1900$
$ad = 1100$	$ade = 1500$
$bd = 3000$	$bde = 4000$
$abd = 6100$	$abde = 6500$
$cd = 800$	$cde = 1500$
$acd = 1100$	$acde = 2000$
$bcd = 3300$	$bcde = 3400$
$abcd = 6000$	$abcde = 6800$

- Estime os efeitos dos fatores.
- Que efeitos parecem importantes? Use um gráfico de probabilidade normal.
- Se for desejado maximizar a resistência, em qual direção você deverá ajustar as variáveis de processo?
- Análise os resíduos desse experimento.

14.6.3 Um experimento correu uma única réplica de um planejamento  $2^4$  e calculou os seguintes efeitos dos fatores:

$A = 80,25$	$AB = 53,25$	$ABC = -2,95$
$B = -65,50$	$AC = 11,00$	$ABD = -8,00$
$C = -9,25$	$AD = 9,75$	$ACD = 10,25$
$D = -20,50$	$BC = 18,36$	$BCD = -7,95$
	$BD = 15,10$	$ABCD = -6,25$
	$CD = -1,25$	

- Construa um gráfico de probabilidade normal dos efeitos.
- Identifique um modelo-tentativa, baseando-se no gráfico dos efeitos do item (a).
- Estime os coeficientes de regressão nesse modelo, considerando que  $\bar{y} = 400$ .

14.6.4 Um experimento descrito por M. G. Natrella no *National Bureau of Standards' Handbook of Experimental Statistics* (1963, nº 91) envolve testar tecidos com relação à chama, depois de aplicar tratamentos com retardantes de fogo. Os quatro fatores considerados são o tipo de tecido ( $A$ ), o tipo de tratamento de retardante de fogo ( $B$ ), condição de lavagem ( $C$  – o nível inferior é sem lavagem; o nível superior é depois de uma lavagem) e o método de condução do teste de chama ( $D$ ). Todos os fatores são corridos em dois níveis, e a variável de resposta são as polegadas de tecido queimado na amostra de teste de tamanho-padrão. Os dados são:

$(1) = 42$	$d = 40$
$a = 31$	$ad = 30$
$b = 45$	$bd = 50$
$ab = 29$	$abd = 25$
$c = 39$	$cd = 40$
$ac = 28$	$acd = 25$
$bc = 46$	$bcd = 50$
$abc = 32$	$abcd = 23$

- Estime e prepare um gráfico normal dos efeitos.
- Construa uma tabela de análise de variância, baseando-se no modelo-tentativa identificado no item (a).
- Construa um gráfico de probabilidade normal dos resíduos e comente os resultados.

14.6.5 Considere a seguinte saída computacional para o experimento fatorial  $2^4$ .

#### Efeitos Estimados e Coeficientes

Termo	Efeito	Coef.	EP do Coef.	t	P
Constante		35,250	?	39,26	0,000
A	2,250	?	?	?	?
B	24,750	12,375	?	13,78	0,000
C	1,000	0,500	?	0,56	0,602
D	10,750	5,375	?	5,99	0,002
A*B	-10,500	-5,250	?	-5,85	0,002
A*C	4,250	2,125	?	2,37	0,064
A*D	-5,000	-2,500	?	-2,78	0,039
B*C	5,250	2,625	?	2,92	0,033
B*D	4,000	2,000	?	2,23	0,076
C*D	-0,750	-0,375	?	-0,42	0,694
S = 3,59166					

#### Análise de Variância

Fonte	GL	SQ	MQ	F	P
A	?	?	?	?	0,266
B	1	2450,25	2450,25	189,94	0,000
C	1	4,00	4,00	0,31	0,602
D	1	462,25	462,25	35,83	0,002
AB	1	441,00	441,00	34,19	0,002
AC	1	72,25	72,25	5,60	0,064
AD	1	100,00	100,00	7,75	0,039
BC	1	110,25	110,25	8,55	0,033
BD	1	64,00	64,00	4,96	0,076
CD	1	2,25	2,25	0,17	0,694
Erro Residual	?	64,50	?		
Total	?	3791,00			

- Que efeitos são usados para estimar o erro?
- Calcule os termos marcados com “?” na saída.

14.6.6 Um artigo publicado em *Bioresource Technology* [“Influence of Vegetable Oils Fatty-Acid Composition on Biodiesel Optimization”, (2011, Vol. 102(2), pp. 1059-1065)] descreveu um experimento para analisar a influência da composição de ácidos graxos em biodiesel. Os fatores foram a concentração de catalisador, a quantidade de metanol e a temperatura e o tempo de reação; o projeto incluiu três pontos centrais. O teor de metil éster do óleo de milho (MEM) foi registrado como a resposta. Seguem dados.

Corrida	Temperatura (°C)	Tempo (min)	Catalisador (%)	Razão molar metanol/óleo	MEM (% em massa)
1	45	40	0,8	5,4	88,30
2	25	40	1,2	5,4	90,50
3	45	10	0,8	4,2	77,96
4	25	10	1,2	5,4	85,59
5	45	40	1,2	5,4	97,14
6	45	10	1,2	4,2	90,64
7	45	40	1,2	4,2	89,86
8	25	40	0,8	4,2	82,35
9	25	10	0,8	5,4	80,31
10	25	40	0,8	5,4	85,51
11	25	10	0,8	4,2	76,21
12	45	40	0,8	4,2	86,86
13	25	10	1,2	4,2	86,35

(continua)

(continuação)

Corrida	Temperatura (°C)	Tempo (min)	Catalisador (%)	Razão molar metanol/ óleo	MEM (% em massa)
14	45	10	0,8	5,4	84,58
15	25	40	1,2	4,2	89,37
16	45	10	1,2	5,4	90,51
17	35	25	1	4,8	91,40
18	35	25	1	4,8	91,96
19	35	25	1	4,8	91,07

- Identifique os efeitos importantes a partir de um gráfico de probabilidade normal.
- Compare os resultados do item anterior com os resultados que usam um termo de erro baseado nos pontos centrais.
- Teste a curvatura.
- Analise os resíduos do modelo.

14.6.7 Um artigo publicado em *Analytica Chimica Acta* ["Design-of-Experiment Optimization of Exhaled Breath Condensate Analysis Using a Miniature Differential Mobility Spectrometer (DMS)"] (2008, Vol. 628(2), pp. 155-161) examinou quatro parâmetros que afetam a sensibilidade e detecção dos instrumentos analíticos usados para medir amostras clínicas. Os autores otimizaram a função sensor usando amostras do condensado do exalado pulmonar (*exhaled breath condensate* – EBC) adicionadas com acetona, um conhecido indicador clínico em respiração. A tabela seguinte mostra os resultados para uma única réplica de um experimento fatorial  $2^4$  para uma das saídas, a amplitude média do pico de acetona sobre as outras três repetições.

Configuração	A	B	C	D	y
1	+	+	+	+	0,12
2	+	+	+	-	0,1193
3	+	+	-	+	0,1196
4	+	+	-	-	0,1192
5	+	-	+	+	0,1186

Configuração	A	B	C	D	y
6	+	-	+	-	0,1188
7	+	-	-	+	0,1191
8	+	-	-	-	0,1186
9	-	+	+	+	0,121
10	-	+	+	-	0,1195
11	-	+	-	+	0,1196
12	-	+	-	-	0,1191
13	-	-	+	+	0,1192
14	-	-	+	-	0,1194
15	-	-	-	+	0,1188
16	-	-	-	-	0,1188

Os fatores e níveis são mostrados na tabela seguinte.

- A Voltagem RF do sensor DMS (1200 ou 1400 V)
- B Vazão do gás transportador nitrogênio (250 ou 500 mL·min<sup>-1</sup>)
- C Tipo de filtro para microextração em fase sólida (MEFS) (poliacrilato ou PDMS-DVB)
- D Perfil de resfriamento do GC (criogênico e não criogênico)

- Estime os efeitos dos fatores e use o gráfico de probabilidade normal dos efeitos. Identifique quais efeitos parecem ser grandes e identifique um modelo para os dados desse experimento.
- Conduza uma análise de variância para confirmar o que você achou no item (a). Quais são suas conclusões?
- Analise os resíduos desse experimento. Existe algum problema com a adequação do modelo?
- Projete o planejamento neste problema para um planejamento  $2^r$  para  $r < 4$  nos fatores importantes. Esquematize o planejamento e mostre a média e a faixa de rendimentos em cada corrida. Esse esquema ajuda na representação dos dados?

## Exercícios para a Seção 14.7

14.7.1 Considere um experimento fatorial  $2^2$ , com quatro pontos centrais. Os dados são (1) = 21,  $a = 125$ ,  $b = 154$ ,  $ab = 352$  e as respostas no ponto central são 92, 130, 98 e 152. Calcule uma ANOVA com a soma dos quadrados para a curvatura e faça um teste  $F$  para a curvatura. Use  $\alpha = 0,05$ .

14.7.2 Considere o experimento do Exercício 14.6.2. Suponha que um ponto central com cinco réplicas seja adicionado às corridas fatoriais e que as respostas sejam 2800, 5600, 4500, 5400 e 3600. Calcule uma ANOVA com a soma dos quadrados para a curvatura e faça um teste  $F$  para a curvatura. Use  $\alpha = 0,05$ .

## Exercícios para a Seção 14.8

14.8.1 Considere os dados da primeira réplica do Exercício 14.5.2.

- Construa um planejamento com dois blocos de quatro observações cada, com  $ABC$  superposto.
- Analise os dados.

14.8.2 Considere os dados do Exercício 14.6.4.

- Construa o planejamento que teria sido usado para correr esse experimento em dois blocos de oito corridas cada.
- Analise os dados e tire conclusões.

14.8.3 Construa um planejamento 2<sup>3</sup> em dois blocos. Selecione a interação *ABCDE* para ser superposta com blocos.

14.8.4 Considere os dados da primeira réplica do Exercício 14.5.1 e suponha que quatro blocos sejam necessários. Superponha *ABD* e *ABC* (e, conseqüentemente, *CD*) com blocos.

- Construa um planejamento com quatro blocos de quatro observações cada um.
- Analise os dados.

14.8.5 Um artigo publicado em *Quality Engineering* ["Designed Experiment to Stabilize Blood Glucose Levels" (1999-2000, Vol. 12, pp. 83-87)] reportou um experimento para minimizar variações nos níveis de glicose no sangue. Os fatores foram: volume de suco ingerido antes de um exercício (4 ou 8 onças), tempo de exercício em uma estação de esqui (10 ou 20 min) e intervalo de tempo (0 ou 20 min) entre tomar o suco e começar o exercício. O período do dia foi a variável bloco. Os dados seguem.

Corrida	Suco (onça)	Exercício (min)	Atraso (min)	Período do Dia	Teor Médio de Glicose no Sangue
1	4	10	0	Tarde	71,5
2	8	10	0	manhã	103
3	4	20	0	manhã	83,5
4	8	20	0	Tarde	126
5	4	10	20	manhã	125,5
6	8	10	20	Tarde	129,5
7	4	20	20	Tarde	95
8	8	20	20	manhã	93

- Que efeitos são superpostos com os blocos? Comente qualquer preocupação com a superposição nesse planejamento.
- Analise os dados e tire conclusões.

14.8.6 Considere o planejamento fatorial 2<sup>6</sup> em quatro blocos. Mostre que um planejamento que superpõe três das interações de quarta ordem com blocos é o melhor arranjo possível de blocagem.

14.8.7 Um artigo publicado em *Advanced Semiconductor Manufacturing Conference (ASMC)* (maio de 2004, pp. 325-329) estabeleceu que regras de expedição e estratégias de retrabalho são os dois maiores elementos operacionais que impactam a produtividade na planta de fabricação de semicondutores. Um experimento com quatro fatores foi realizado para determinar o efeito do tempo da regra de expedição (5 ou 10 min), do atraso pelo retrabalho (0 ou 15 min), da temperatura de fabricação (60 ou 80 °F) e dos níveis de retrabalho (nível 0 ou nível 1) sobre as medidas cruciais de desempenho da fábrica. A medida de desempenho analisada foi o tempo médio do ciclo. A temperatura de fabricação foi a variável usada como bloco. Os dados modificados a partir do estudo original são mostrados na tabela seguinte.

Corrida	Tempo da Regra de Expedição (min)	Atraso pelo Retrabalho (min)	Nível de Retrabalho	Temperatura de Fabricação (°F)	Tempo Médio do Ciclo (min)
1	5	0	0	60	218
2	10	0	0	80	256,5
3	5	0	1	80	231
4	10	0	1	60	302,5
5	5	15	0	80	298,5
6	10	15	0	60	314
7	5	15	1	60	249
8	10	15	1	80	241

- Que efeitos são superpostos com os blocos? Você vê algum problema com as superposições nesse planejamento? Se sim, comente isso.
- Analise os dados e tire conclusões.

## Exercícios para a Seção 14.9

14.9.1 Um artigo de L. B. Hare ["In the Soup: A Case Study to Identify Contributors to Filling Variability", *Journal of Quality Technology* (1988, Vol. 20, pp. 36-43)] descreveu um experimento fatorial usado para estudar a variabilidade de enchimento de pacotes de misturas secas para sopas. Os fatores são *A* = número de portas de mistura pelas quais o óleo vegetal é adicionado (1, 2), *B* = temperatura em torno do misturador (fria, ambiente), *C* = tempo de mistura (60, 80 s), *D* = massa da batelada (1500, 2000 lb) e *E* = número de dias de atraso entre a mistura e o empacotamento (1, 7). Entre 125 e 150 pacotes de sopa foram amostrados em um período de oito horas para cada corrida do planejamento, e o desvio-padrão do peso do pacote foi usado como a variável de resposta. O planejamento e os dados resultantes seguem.

Ordem Padrão	A Portas do Misturador	B Temp.	C Tempo	D Massa da Batelada	E Atraso	y Desvio-Padrão
1	-	-	-	-	-	1,13
2	+	-	-	-	+	1,25
3	-	+	-	-	+	0,97
4	+	+	-	-	-	1,70
5	-	-	+	-	+	1,47
6	+	-	+	-	-	1,28

Ordem Padrão	A Portas do Misturador	B Temp.	C Tempo	D Massa da Batelada	E Atraso	y Desvio-Padrão
7	-	+	+	-	-	1,18
8	+	+	+	-	+	0,98
9	-	-	-	+	+	0,78
10	+	-	-	+	-	1,36
11	-	+	-	+	-	1,85
12	+	+	-	+	+	0,62
13	-	-	+	+	-	1,09
14	+	-	+	+	+	1,10
15	-	+	+	+	+	0,76
16	+	+	+	+	-	2,10

- Qual é o gerador para esse planejamento?
- Qual é a resolução desse planejamento?
- Estime os efeitos dos fatores. Quais efeitos são grandes?
- Uma análise residual indica qualquer problema com as suposições básicas?
- Tire conclusões sobre esse processo de enchimento.

14.9.2 Um artigo publicado em *Quality Engineering* [“A Comparison of Multi-Response Optimization: Sensitivity to Parameter Selection” (1999, Vol. 11, pp. 405-415)] conduziu metade de uma réplica de um planejamento fatorial  $2^5$ , de modo a otimizar o processo de esterilização em retortas de bife ensopado, utilizado em ração militar; essa refeição é conhecida como RPC (refeição pronta para comer). Os fatores do planejamento são  $x_1$  = viscosidade do molho,  $x_2$  = gás residual,  $x_3$  = razão sólido/líquido,  $x_4$  = peso líquido,  $x_5$  = velocidade de rotação. A variável resposta é o índice de taxa de aquecimento, uma medida de penetração de calor. Há duas réplicas.

Corrida	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	Índice de Taxa de Aquecimento	
						I	II
1	-1	-1	-1	-1	1	8,46	9,61
2	1	-1	-1	-1	-1	15,68	14,68
3	-1	1	-1	-1	-1	14,94	13,09
4	1	1	-1	-1	1	12,52	12,71
5	-1	-1	1	-1	-1	17,0	16,36
6	1	-1	1	-1	1	11,44	11,83
7	-1	1	1	-1	1	10,45	9,22
8	1	1	1	-1	-1	19,73	16,94
9	-1	-1	-1	1	-1	17,37	16,36
10	1	-1	-1	1	1	14,98	11,93
11	-1	1	-1	1	1	8,40	8,16
12	1	1	-1	1	-1	19,08	15,40
13	-1	-1	1	1	1	13,07	10,55
14	1	-1	1	1	-1	18,57	20,53
15	-1	1	1	1	-1	20,59	21,19
16	1	1	1	1	1	14,03	11,31

- a. Estime os efeitos dos fatores. Baseando-se no gráfico de probabilidade normal das estimativas dos efeitos, identifique um modelo para os dados desse experimento.  
b. Conduza uma ANOVA baseada no modelo identificado no item (a). Quais são suas conclusões?

- c. Analise os resíduos e comente a adequação do modelo.  
d. Encontre o modelo de regressão para prever o rendimento em termos dos níveis dos fatores codificados.  
e. Esse experimento foi replicado; logo, uma ANOVA poderia ter sido conduzida sem usar um gráfico normal dos efeitos para tentar identificar um modelo. Qual modelo seria apropriado? Use a ANOVA para analisar esse modelo e compare os resultados com aqueles obtidos a partir da abordagem do gráfico de probabilidade normal.

14.9.3 R. D. Snee (“Experimenting with a Large Number of Variables”, em *Experiments in Industry: Design, Analysis and Interpretation of Results*, por R. D. Snee, L. D. Hare e J. B. Trout, eds., ASQC, 1985) descreve um experimento em que um planejamento  $2^{5-1}$ , com  $I = ABCDE$ , foi usado para investigar os efeitos de cinco fatores sobre a cor de um produto químico. Os fatores são  $A$  = solvente/reagente,  $B$  = catalisador/reagente,  $C$  = temperatura,  $D$  = pureza do reagente e  $E$  = pH do reagente. Os resultados obtidos são:

$$\begin{aligned} e &= -0,63 & d &= 6,79 \\ a &= 2,51 & ade &= 6,47 \\ b &= -2,68 & bde &= 3,45 \\ abe &= 1,66 & abd &= 5,68 \\ c &= 2,06 & cde &= 5,22 \\ ace &= 1,22 & acd &= 4,38 \\ bce &= -2,09 & bcd &= 4,30 \\ abc &= 1,93 & abcde &= 4,05 \end{aligned}$$

- a. Prepare um gráfico de probabilidade normal dos efeitos. Que fatores são ativos?  
b. Calcule os resíduos. Construa um gráfico de probabilidade normal dos resíduos e plote-os contra os valores previstos. Comente os gráficos.  
c. Se qualquer fator for desprezível, reduza o planejamento  $2^{5-1}$  a um fatorial completo nos fatores ativos. Comente o planejamento resultante e interprete os resultados.

## Exercícios para a Seção 14.10

14.10.1 Um artigo publicado em *Industrial and Engineering Chemistry* [“More on Planning Experiments to Increase Research Efficiency” (1970, pp. 60-65)] usa um planejamento  $2^{5-2}$  para investigar o efeito no rendimento do processo de  $A$  = temperatura de condensação,  $B$  = quantidade de material 1,  $C$  = volume do solvente,  $D$  = tempo de condensação e  $E$  = quantidade de material 2. Os resultados obtidos são:

$$\begin{aligned} ae &= 23,2 & cd &= 23,8 \\ ab &= 15,5 & ace &= 23,4 \\ ad &= 16,9 & bde &= 16,8 \\ bc &= 16,2 & abcde &= 18,1 \end{aligned}$$

- a. Verifique se os geradores usados do planejamento foram  $I = ACE$  e  $I = BDE$ .  
b. Escreva a relação completa de definição e os pares associados do planejamento.

- c. Estime os efeitos principais.  
d. Prepare uma tabela de análise de variância. Verifique se as interações  $AB$  e  $AD$  estão disponíveis para uso como erro.  
e. Plote os resíduos contra os valores previstos. Construa também um gráfico de probabilidade normal dos resíduos. Comente os resultados.

14.10.2 Suponha que no Exercício 14.6.2 foi possível correr somente duas meias-frações (ou 1/4) do planejamento  $2^3$ . Construa o planejamento e analise os dados que são obtidos selecionando somente a resposta para as oito corridas em seu planejamento.

14.10.3 Para cada um dos seguintes planejamentos, escreva os pares associados, considerando que somente os efeitos principais e as interações de segunda ordem são de interesse.

a.  $2_{III}^{6-3}$       b.  $2_{IV}^{8-4}$

14.10.4 Considere o planejamento  $2^{6-2}$  na Tabela 14.26.

- a. Suponha que, depois de analisar os dados originais, encontramos que os fatores  $C$  e  $E$  podem ser eliminados. Que tipo de planejamento  $2^k$  é deixado nas variáveis restantes?
- b. Suponha que, depois de analisar os dados originais, encontramos que os fatores  $D$  e  $F$  podem ser eliminados. Que tipo de planejamento  $2^k$  é deixado nas variáveis restantes? Compare os resultados com o item (a). Você pode explicar por que as respostas são diferentes?

14.10.5 Um artigo publicado no *Journal of Marketing Research* (1973, Vol. 10(3), pp. 270-276) apresentou um planejamento fatorial fracionário  $2^{7-4}$  para conduzir a pesquisa de mercado:

Corridas	A	B	C	D	E	F	G	Vendas por um Período de 6 Semanas (em \$ 1000)
1	-1	-1	-1	1	1	1	-1	8,7
2	1	-1	-1	-1	-1	1	1	15,7
3	-1	1	-1	-1	1	-1	1	9,7
4	1	1	-1	1	-1	-1	-1	11,3
5	-1	-1	1	1	-1	-1	1	14,7
6	1	-1	1	-1	1	-1	-1	22,3
7	-1	1	1	-1	-1	1	-1	16,1
8	1	1	1	1	1	1	1	22,1

Os fatores e níveis são mostrados na tabela seguinte.

Fator	-1	+1
A Anúncio na televisão	Sem anúncio	Anúncio
B Anúncio em outdoor	Sem anúncio	Anúncio
C Anúncio em jornal	Sem anúncio	Anúncio
D Planejamento para embalar doces	Planejamento conservativo	Projeto gráfico chamativo
E Planejamento da disposição	Disposição normal na prateleira	Disposição especial no corredor
F Programa de amostra grátis	Sem amostras grátis	Amostras grátis
G Tamanho da barra de doce	Barra de 1 oz.	Barra de 2,5 oz.

- a. Escreva as relações dos pares associados.
- b. Estime os efeitos principais.
- c. Prepare um gráfico de probabilidade normal para os efeitos e interprete os resultados.

## Exercícios para a Seção 14.11

14.11.1 Um artigo publicado em *Rubber Age* (1961, Vol. 89, pp. 453-458) descreve um experimento sobre a fabricação de um produto, em que dois fatores foram variados. Os fatores são o tempo de reação ( $h$ ) e a temperatura ( $^{\circ}\text{C}$ ). Esses fatores são codificados como  $x_1 = (\text{tempo} - 12)/8$  e  $x_2 = (\text{temperatura} - 250)/30$ . Os dados a seguir foram observados, sendo  $y$  o rendimento (em porcentagem):

Número da Corrida	$x_1$	$x_2$	$y$
1	-1	0	83,8
2	1	0	81,7
3	0	0	82,4
4	0	0	82,9
5	0	-1	84,7
6	0	1	75,9
7	0	0	81,2
8	-1,414	-1,414	81,3
9	-1,414	1,414	83,1
10	1,414	-1,414	85,3
11	1,414	1,414	72,7
12	0	0	82,0

- a. Plote os pontos nos quais as corridas experimentais foram feitas.
- b. Ajuste um modelo de segunda ordem aos dados. O modelo de segunda ordem é adequado?
- c. Plote a superfície de resposta do rendimento. Que recomendações você faria acerca das condições operacionais para esse processo?

14.11.2 Um artigo publicado em *Quality Engineering* ["Mean and Variance Modeling with Qualitative Responses: A Case Study" (1998-1999, Vol. 11, pp. 141-148)] estudou como três ingredientes ativos de um alimento popular afetam o gosto global do produto. A medida do gosto global é a graduação média global de satisfação (GMGS). Os três ingredientes são identificados pelas variáveis  $x_1$ ,  $x_2$  e  $x_3$ . Os dados são mostrados na tabela a seguir.

Corrida	$x_1$	$x_2$	$x_3$	GMGS
1	1	1	-1	6,3261
2	1	1	1	6,2444
3	0	0	0	6,5909
4	0	-1	0	6,3409
5	1	-1	1	5,907
6	1	-1	-1	6,488
7	0	0	-1	5,9773
8	0	1	0	6,8605
9	-1	-1	1	6,0455
10	0	0	1	6,3478
11	1	0	0	6,7609
12	-1	-1	-1	5,7727
13	-1	1	-1	6,1805
14	-1	1	1	6,4894
15	-1	0	0	6,8182

- a. Ajuste um modelo de superfície de resposta de segunda ordem aos dados.
- b. Construa gráficos de curvas de nível e de superfícies de resposta para GMGS. Quais são suas conclusões?
- c. Analise os resíduos desse experimento. Sua análise indica qualquer problema potencial?
- d. Esse planejamento tem um único ponto central. Em sua opinião, esse é um bom planejamento?

14.11.3 Considere o modelo de primeira ordem

$$\hat{y} = 50 + 1,5x_1 - 0,8x_2$$

em que  $-1 \leq x_i \leq 1$ . Encontre a direção da ascendente de maior inclinação.

14.11.4 Um fabricante de ferramentas de corte desenvolveu duas equações empíricas para a vida da ferramenta ( $y_1$ ) e o seu custo ( $y_2$ ). Ambos os modelos são funções da dureza da ferramenta ( $x_1$ ) e do tempo de fabricação ( $x_2$ ). As equações são

$$\hat{y}_1 = 10 + 5x_1 + 2x_2$$

$$\hat{y}_2 = 23 + 3x_1 + 4x_2$$

sendo ambas as equações válidas sobre a faixa  $-1,5 \leq x_i \leq 1,5$ . Suponha que a vida da ferramenta tenha de exceder 12 horas e que o custo deva ser abaixo de \$ 27,50.

- a. Existe um conjunto exequível de condições operacionais?
- b. Onde você correria esse processo?

14.11.5 Um artigo em *Tappi* (1960, Vol. 43, pp. 38-44) descreve um experimento que investigou o valor da cinza da polpa de papel (uma medida de impurezas inorgânicas). Duas variáveis, temperatura  $T$ , em graus Celsius, e tempo, em horas, foram estudadas, sendo alguns resultados mostrados na tabela a seguir. As variáveis predictoras codificadas mostradas são

$$x_1 = \frac{(T - 775)}{115}, \quad x_2 = \frac{(t - 3)}{1,5}$$

e a resposta  $y$  é (valor da cinza seca em %)  $\times 10^3$ .

$x_1$	$x_2$	$y$	$x_1$	$x_2$	$y$
-1	-1	211	0	-1,5	168
1	-1	92	0	1,5	179
-1	1	216	0	0	122
1	1	99	0	0	175
-1,5	0	222	0	0	157
1,5	0	48	0	0	146

- a. Que tipo de planejamento foi usado nesse estudo? Esse planejamento é rotacionável?
- b. Ajuste um modelo quadrático aos dados. Esse modelo é satisfatório?
- c. Se for importante minimizar o valor da cinza, onde você correria o processo?

14.11.6 Em seu livro *Empirical Model Building and Response Surfaces* (John Wiley, 1987), Box e Draper descrevem um experimento com três fatores. Os dados mostrados na tabela a seguir representam uma variação do experimento original de seu livro. Suponha que esses dados tenham sido coletados em um processo de fabricação de semicondutores.

$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
-1	-1	-1	24,00	12,49
0	-1	-1	120,33	8,39
1	-1	-1	213,67	42,83
-1	0	-1	86,00	3,46
0	0	-1	136,63	80,41
1	0	-1	340,67	16,17
-1	1	-1	112,33	27,57
0	1	-1	256,33	4,62
1	1	-1	271,67	23,63
-1	-1	0	81,00	0,00
0	-1	0	101,67	17,67
1	-1	0	357,00	32,91
-1	0	0	171,33	15,01
0	0	0	372,00	0,00
1	0	0	501,67	92,50
-1	1	0	264,00	63,50
0	1	0	427,00	88,61
1	1	0	730,67	21,08
-1	-1	1	220,67	133,82
0	-1	1	239,67	23,46
1	-1	1	422,00	18,52
-1	0	1	199,00	29,44
0	0	1	485,33	44,67
1	0	1	673,67	158,21
-1	1	1	176,67	55,51
0	1	1	501,00	138,94
1	1	1	1010,00	142,45

- a. A resposta  $y_1$  é a média de três leituras da resistividade para uma única pastilha. Ajuste um modelo quadrático para essa resposta.
- b. A resposta  $y_2$  é o desvio-padrão das três medidas da resistividade. Ajuste um modelo linear para essa resposta.
- c. Onde você recomendaria que estabelecêssemos  $x_1, x_2$  e  $x_3$ , se o objetivo fosse manter a resistividade média em 500 e minimizar o desvio-padrão?

14.11.7 Considere o modelo de primeira ordem

$$y = 12 + 1,2x_1 - 2,1x_2 + 1,6x_3 - 0,6x_4$$

em que  $-1 \leq x_i \leq 1$ .

- a. Encontre a direção da ascendente de maior inclinação.
- b. Considere que o atual planejamento esteja centralizado no ponto (0, 0, 0, 0). Determine o ponto que está a três unidades do atual ponto central na direção da ascendente de maior inclinação.

14.11.8 Suponha que a resposta  $y_1$  seja uma função de dois fatores,  $x_1$  e  $x_2$ , com  $y_1 = 2x_2^2 - 4x_1^2 - x_1x_2 + 4$ .

- a. Desenhe as curvas de nível dessa função de resposta.
- b. Considere outra resposta:  $y_2 = (x_1 - 2)^2 + (x_2 - 3)^2$ .
- c. Adicione as curvas de níveis para  $y_2$  e discuta quão factível é minimizar ambos  $y_1$  e  $y_2$  com valores para  $x_1$  e  $x_2$ .

14.11.9 Duas respostas,  $y_1$  e  $y_2$ , estão relacionadas com os dois fatores,  $x_1$  e  $x_2$ , pelos modelos  $y_1 = 5 + (x_1 - 2)^2 + (x_2 - 3)^2$  e  $y_2 = x_2 - x_1 + 3$ . Suponha que os objetivos sejam  $y_1 \leq 9$  e  $y_2 \geq 6$ .

a. Existe um conjunto factível de condições operacionais para  $x_1$  e  $x_2$ ? Em caso positivo, plote a região factível no espaço de  $x_1$  e  $x_2$ .

b. Determine os pontos  $(x_1, x_2)$  que resultam em  $y_2 \geq 6$  e que minimizem  $y_1$ .

14.11.10 Um artigo publicado no *Journal of Materials Processing Technology* (1997, Vol. 67, pp. 55-61) usou a metodologia de superfície de resposta para gerar modelos de previsão da rugosidade da superfície para aço EN 24T (290 BHN). Os dados são mostrados na tabela a seguir.

Os fatores e os níveis para o experimento são mostrados na seguinte tabela.

Níveis	Mais				Mais
	Baixo	Inferior	Central	Superior	
Codificação	$-\sqrt{2}$	-1	0	1	$\sqrt{2}$
Velocidade, $V$ (m min <sup>-1</sup> )	28	36	65	117	150
Alimentação, $f$ (mm rev <sup>-1</sup> )	0,12	0,15	0,25	0,40	0,50
Profundidade do corte, $d$ (mm)	0,42	0,50	0,75	1,125	1,33

Tentativa	Velocidade (m min <sup>-1</sup> )	Alimentação (mm rev <sup>-1</sup> )	Profundidade do Corte (mm)	Codificação			Rugosidade da Superfície (µm)
				$x_1$	$x_2$	$x_3$	
1	36	0,15	0,50	-1	-1	-1	1,8
2	117	0,15	0,50	1	-1	-1	1,233
3	36	0,40	0,50	-1	1	-1	5,3
4	117	0,40	0,50	1	1	-1	5,067
5	36	0,15	1,125	-1	-1	1	2,133
6	117	0,15	1,125	1	-1	1	1,45
7	36	0,40	1,125	-1	1	1	6,233
8	117	0,40	1,125	1	1	1	5,167
9	65	0,25	0,75	0	0	0	2,433
10	65	0,25	0,75	0	0	0	2,3
11	65	0,25	0,75	0	0	0	2,367
12	65	0,25	0,75	0	0	0	2,467
13	28	0,25	0,75	$-\sqrt{2}$	0	0	3,633
14	150	0,25	0,75	$\sqrt{2}$	0	0	2,767
15	65	0,12	0,75	0	$-\sqrt{2}$	0	1,153
16	65	0,50	0,75	0	$\sqrt{2}$	0	6,333
17	65	0,25	0,42	0	0	$-\sqrt{2}$	2,533
18	65	0,25	1,33	0	0	$\sqrt{2}$	3,20
19	28	0,25	0,75	$-\sqrt{2}$	0	0	3,233
20	150	0,25	0,75	$\sqrt{2}$	0	0	2,967
21	65	0,12	0,75	0	$-\sqrt{2}$	0	1,21
22	65	0,50	0,75	0	$\sqrt{2}$	0	6,733
23	65	0,25	0,42	0	0	$-\sqrt{2}$	2,833
24	65	0,25	1,33	0	0	$\sqrt{2}$	3,267

- a. Plote os pontos nos quais as corridas experimentais foram feitas.  
 b. Ajuste os modelos de primeira e segunda ordens aos dados. Comente a adequação desses modelos.  
 c. Plote a superfície de resposta para a rugosidade para o modelo de segunda ordem e comente.

14.11.11 Um artigo publicado em *Analytical Biochemistry* ["Application of Central Composite Design for DNA Hybridization Onto Magnetic Microparticles" (2009, Vol. 39(1), pp. 17-23)] considerou os efeitos da concentração da sonda e do alvo e o número de partículas

na imobilização e hibridização sobre um ensaio de hibridização à base de micropartícula de DNA. A fluorescência média é a resposta. A concentração de partículas foi transformada em medidas de área superficial. Outras concentrações foram medidas em micromols por litro ( $\mu\text{M}$ ). Os dados estão a seguir.

- a. Que tipo de planejamento é esse?  
 b. Ajuste um modelo de superfície de resposta de segunda ordem para os dados.  
 c. A análise residual indica algum problema?

Corrida	Área de Imobilização ( $\text{cm}^2$ )	Área do Sensor ( $\mu\text{M}$ )	Hibridização ( $\text{cm}^2$ )	Alvo ( $\mu\text{M}$ )	Fluorescência Média
1	0,35	0,025	0,35	0,025	4,7
2	7	0,025	0,35	0,025	4,7
3	0,35	2,5	0,35	0,025	28,0
4	7	2,5	0,35	0,025	81,2
5	0,35	0,025	3,5	0,025	5,7
6	7	0,025	3,5	0,025	3,8
7	0,35	2,5	3,5	0,025	12,2
8	7	2,5	3,5	0,025	19,5
9	0,35	0,025	0,35	5	4,4
10	7	0,025	0,35	5	2,6
11	0,35	2,5	0,35	5	83,7
12	7	2,5	0,35	5	84,7
13	0,35	0,025	3,5	5	6,8
14	7	0,025	3,5	5	2,4
15	0,35	2,5	3,5	5	76
16	7	2,5	3,5	5	77,9
17	0,35	5	2	2,5	42,6
18	7	5	2	2,5	52,3
19	3,5	0,025	2	2,5	2,6
20	3,5	2,5	2	2,5	72,8
21	3,5	5	0,35	2,5	47,7
22	3,5	5	3,5	2,5	54,4
23	3,5	5	2	0,025	30,8
24	3,5	5	2	5	64,8
25	3,5	5	2	2,5	51,6
26	3,5	5	2	2,5	52,6
27	3,5	5	2	2,5	56,1

14.11.12 Um artigo publicado em *Applied Biochemistry and Biotechnology* ["A Statistical Approach for Optimization of Polyhydroxybutyrate Production by *Bacillus sphaericus* NCIM 5149 under Submerged Fermentation Using Central Composite Design" (2010, Vol. 162(4), pp. 996-1007)] descreveu um experimento para otimizar a produção de polihidroxibutirato (PHB). A idade de inoculação, o pH e o substrato foram selecionados como fatores e o planejamento composto central foi conduzido. Os dados seguem.

Corrida	Idade do Inóculo (h)	pH	Substrato (g/L)	PHB (g/L)
1	12	4	1	0,84
2	24	8	1	0,55
3	18	6	2,5	1,96
4	28	6	2,5	1,2
5	12	4	4	0,783
6	18	6	2,5	1,66
7	18	6	2,5	2,22
8	18	6	5	0,8
9	12	8	4	0,48
10	18	6	2,5	1,97
11	18	6	2,5	2,2

Corrida	Idade do Inóculo (h)	pH	Substrato (g/L)	PHB (g/L)
12	18	6	2,5	2,25
13	18	2	2,5	0,2
14	18	6	0	0,22
15	12	8	1	0,37
16	24	8	4	0,66
17	24	4	1	0,28
18	24	4	4	0,88
19	18	9	2,5	0,3
20	7	6	2,5	0,42

- Plote os pontos nos quais as corridas experimentais foram feitas. [Sugestão: codifique cada variável primeiro.] Que tipo de planejamento é usado?
- Ajuste um modelo de superfície de resposta de segunda ordem aos dados.
- A análise residual indica qualquer problema?
- Construa uma curva de nível e uma superfície de resposta para a quantidade de PHB em termos de dois fatores.
- Você pode recomendar valores para a idade do inóculo, o pH e o substrato de modo a maximizar a produção?

## Exercícios Suplementares para o Capítulo 14

14.S6 O tratamento térmico de peças metálicas é um processo largamente utilizado. Um artigo publicado no *Journal of Metals* ["Optimizing Heat Treatment with Factorial Design" (1989, Vol. 41(3), pp. 52-53)] descreveu um experimento para investigar a distorção do achatamento proveniente do tratamento térmico para três tipos de engrenagens e dois tempos de tratamento térmico. Os dados são mostrados a seguir:

Tipo de Engrenagem	Tempo (minutos)	
	90	120
20 dentes	0,0265	0,0560
	0,0340	0,0650
24 dentes	0,0430	0,0720
	0,0510	0,0880
28 dentes	0,0405	0,0620
	0,0575	0,0825

- Há alguma evidência de que a distorção do achatamento seja diferente para diferentes tipos de engrenagens? Há alguma indicação de que o tempo de tratamento térmico afete a distorção do achatamento? Esses fatores interagem? Use  $\alpha = 0,05$ .
- Construa gráficos dos efeitos dos fatores, que ajudem a tirar conclusões desse experimento.
- Análise os resíduos desse experimento. Comente a validade das suposições em foco.

14.S7 Um artigo publicado em *Process Engineering* (1992, nº 71, pp. 46-47) apresenta um experimento fatorial, com dois fatores, usado para investigar o efeito do pH e da concentração de catalisador sobre a viscosidade (cSt) do produto. Os dados são:

pH	Concentração do Catalisador	
	2,5	2,7
5,6	192, 199, 189, 198	178, 186, 179, 188
5,9	185, 193, 185, 192	197, 196, 204, 204

- Teste os efeitos principais e as interações, usando  $\alpha = 0,05$ . Quais são as suas conclusões?
- Faça um gráfico da interação e discuta a informação fornecida por esse gráfico.
- Análise os resíduos desse experimento.

14.S8 Um artigo publicado em *IEEE Transactions on Components, Hybrids, and Manufacturing Technology* (1992, Vol. 15) descreve um experimento para investigar um método de alinhamento de *chips* ópticos em placas de circuitos. O método envolve colocar pingos de solda no fundo do *chip*. O experimento usou três tamanhos de pingos de solda e três métodos de alinhamento. A variável de resposta é a exatidão do alinhamento (em micrômetros). Os dados são:

Tamanho dos Pingos de Solda (diâmetro em $\mu\text{m}$ )	Método de Alinhamento		
	1	2	3
75	4,60	1,55	1,05
	4,53	1,45	1,00
130	2,33	1,72	0,82
	2,44	1,76	0,95
260	4,95	2,73	2,36
	4,55	2,60	2,46

- a. Há qualquer indicação de que tanto o tamanho dos pingos de solda como o método de alinhamento afetam a exatidão do alinhamento? Há qualquer evidência de interação entre esses fatores? Use  $\alpha = 0,05$ .
- b. Que recomendações você faria a respeito desse processo?
- c. Analise os resíduos desse experimento. Comente a adequação do modelo.

14.S9 Um artigo publicado no *Textile Research Institute Journal* (1984, Vol. 54, pp. 171-179) reportou os resultados de um experimento que estudou os efeitos de tratar tecidos com sais inorgânicos selecionados sobre a inflamabilidade do material. Foram usados dois níveis de aplicação de cada sal, e um teste de queima vertical foi usado em cada amostra. (Esse teste encontra a temperatura na qual cada amostra se incendia.) Os dados do teste de queima são mostrados a seguir.

Nível	Sal					
	Não tratado	MgCl <sub>2</sub>	NaCl	CaCO <sub>3</sub>	CaCl <sub>2</sub>	Na <sub>2</sub> CO <sub>3</sub>
1	812	752	739	733	725	751
	827	728	731	728	727	761
	876	764	726	720	719	755
2	945	794	741	786	756	910
	881	760	744	771	781	854
	919	757	727	779	814	848

- a. Teste as diferenças entre sais, entre níveis de aplicação e entre interações. Use  $\alpha = 0,01$ .
- b. Desenhe um gráfico da interação entre o sal e o nível de aplicação. Que conclusões você pode tirar desse gráfico?
- c. Analise os resíduos desse experimento.

14.S10 Um artigo publicado no *Journal of Coatings Technology* (1988, Vol. 60, pp. 27-32) descreve um planejamento fatorial 2<sup>4</sup>, usado para estudar uma tinta, à base de prata, para automóveis. A variável de resposta é a propriedade de distinguir imagens (PDI). As variáveis usadas no experimento são

- A = Percentagem de poliéster por peso de poliéster/melanina (valor baixo = 50 %; valor alto = 70 %)
- B = Percentagem de acetato butirato carboxilato de celulose (valor baixo = 15 %; valor alto = 30 %)
- C = Percentagem de estearato de alumínio (valor baixo = 1 %; valor alto = 3 %)
- D = Percentagem de catalisador ácido (valor baixo = 0,25 %; valor alto = 0,50 %)

As respostas são: (1) = 63,8; a = 77,6; b = 68,8; ab = 76,5; c = 72,5; ac = 77,2; bc = 77,7; abc = 84,5; d = 60,6; ad = 64,9; bd = 72,7; abd = 73,3; cd = 68,0; acd = 76,3; bcd = 76,0 e abcd = 75,9.

- a. Estime os efeitos dos fatores.
- b. A partir do gráfico de probabilidade normal dos efeitos, tente um modelo para os dados desse experimento.
- c. Usando os fatores aparentemente desprezíveis como uma estimativa de erro, teste a significância dos fatores identificados no item (b). Use  $\alpha = 0,05$ .
- d. Que modelo você usaria para descrever o processo, baseado nesse experimento? Interprete o modelo.
- e. Analise os resíduos do modelo no item (d) e comente as suas descobertas.

14.S11 Um artigo publicado em *Solid State Technology* (1984, Vol. 29, pp. 281-284) descreve o uso de planejamentos fatoriais em fotolitografia, uma importante etapa no processo de fabricação de circuitos integrados. As variáveis nesse experimento (todas com dois níveis) são temperatura de pré-cozimento (A), tempo de pré-cozimento (B) e energia de exposição (C); a variável de resposta é a largura da linha delta, a diferença entre a linha na máscara e a linha impressa no dispositivo. Os dados são: (1) = -2,30; a = -9,87; b = -18,20; ab = -30,20; c = -23,80; ac = -4,30; bc = -3,80 e abc = -14,70.

- a. Estime os efeitos dos fatores.
- b. Use o gráfico de probabilidade normal das estimativas dos efeitos para identificar fatores que podem ser importantes.
- c. Que modelo você recomendaria para prever a resposta largura da linha delta, baseado nos resultados desse experimento?
- d. Analise os resíduos desse experimento e comente a adequação do modelo.

14.S12 Construa um planejamento 2<sup>4-1</sup><sub>IV</sub> para o problema do Exercício 14.S10. Selecione os dados para as oito corridas requeridas para esse planejamento. Analise essas corridas e compare suas conclusões com aquelas obtidas no Exercício 14.S10 para o planejamento completo.

14.S13 Construa um planejamento 2<sup>5-2</sup><sub>III</sub> em oito corridas. Quais são as relações de associação nesse planejamento?

14.S14 Construa um planejamento 2<sup>8-4</sup><sub>IV</sub> em 16 corridas. Quais são as relações de associação nesse planejamento?

14.S15 Um artigo publicado em *Rubber Chemistry and Technology* (1974, Vol. 47, pp. 825-836) descreve um experimento que estuda a viscosidade Mooney da borracha, em relação a muitas variáveis, incluindo a carga de sílica (percentagem) e a carga de óleo (percentagem). Dados típicos reportados nesse experimento são mostrados na sequência, sendo

$$x_1 = \frac{\text{sílica} - 60}{15}, \quad x_2 = \frac{\text{óleo} - 21}{15}$$

**Níveis Codificados**

x <sub>1</sub>	x <sub>2</sub>	y
-1	-1	13,71
1	-1	14,15
-1	1	12,87
1	1	13,53
-1	-1	13,90
1	-1	14,88
-1	1	12,25
-1	1	13,35

- a. Que tipo de planejamento de experimentos foi usado?
- b. Analise os dados e tire conclusões apropriadas.

14.S16 Um artigo publicado em *Plant Disease* [“Effect of Nitrogen and Potassium Fertilizer Rates on Severity of Xanthomonas Blight of *Syngonium Podophyllum*” (1989, Vol. 73(12), pp. 972-975)] mostrou o efeito das variáveis taxas de nitrogênio e de potássio sobre o crescimento da “Borboleta Branca” e a percentagem média da doença. Dados representativos daqueles coletados nesse experimento são fornecidos na tabela seguinte.

Nitrogênio (mg/pote/ semana)	Potássio (mg/pote/semana)					
	30		90		120	
50	61,0	61,3	45,5	42,5	59,5	58,2
150	54,5	55,9	53,5	51,9	34,0	35,9
250	42,7	40,4	36,5	37,4	32,5	33,8

- Estabeleça as hipóteses apropriadas.
- Use a análise de variância para testar essas hipóteses com  $\alpha = 0,05$ .
- Análise graficamente os resíduos desse experimento.
- Estime os componentes apropriados de variância.

14.S17 Um artigo publicado na revista *Journal of Applied Electrochemistry* (maio de 2008, Vol. 38(5), pp. 583-590) apresentou um planejamento fatorial fracionário  $2^{7-3}$  para fazer a otimização dos arranjos de eletrodos, baseados em membranas de polibenzimidazol para células de combustível  $H_2/O_2$ . O planejamento e os dados são mostrados na tabela seguinte.

Corridas	A	B	C	D	E	F	G	Densidade de Corrente
								(em mA cm <sup>2</sup> )
1	-1	-1	-1	-1	-1	-1	-1	160
2	+1	-1	-1	-1	+1	+1	+1	20
3	-1	+1	-1	-1	+1	+1	-1	80
4	+1	+1	-1	-1	-1	-1	+1	317
5	-1	-1	+1	-1	+1	-1	+1	19
6	+1	-1	+1	-1	-1	+1	-1	4
7	-1	+1	+1	-1	-1	+1	+1	20
8	+1	+1	+1	-1	+1	-1	-1	88
9	-1	-1	-1	+1	-1	+1	+1	1100
10	+1	-1	-1	+1	+1	-1	-1	12
11	-1	+1	-1	+1	+1	-1	+1	552
12	+1	+1	-1	+1	-1	+1	-1	880
13	-1	-1	+1	+1	+1	+1	-1	16
14	+1	-1	+1	+1	-1	-1	+1	20
15	-1	+1	+1	+1	-1	-1	-1	8
16	+1	+1	+1	+1	+1	+1	+1	15

Os fatores e os níveis são mostrados na tabela seguinte.

Fator	-1	+1
A Quantidade de ligante na camada de catalisador	0,2 mg cm <sup>2</sup>	1 mg cm <sup>2</sup>
B Carregamento eletrocatalítico	0,1 mg cm <sup>2</sup>	1 mg cm <sup>2</sup>
C Quantidade de carbono na camada de difusão de gás	2 mg cm <sup>2</sup>	4,5 mg cm <sup>2</sup>

D Tempo de compactação a quente	1 min	10 min
E Temperatura de compactação	100 °C	150 °C
F Carga de compactação a quente	0,04 t cm <sup>2</sup>	0,2 t cm <sup>2</sup>
G Quantidade de PTFE na camada de difusão de gás	0,1 mg cm <sup>2</sup>	1 mg cm <sup>2</sup>

- Escreva as relações dos pares associados.
- Estime os efeitos principais.
- Prepare um gráfico de probabilidade normal para os efeitos e interprete os resultados.
- A partir da estimativa do efeito correspondente, calcule a soma dos quadrados para os pares associados que contenham a interação ABG.

14.S18 Considere os seguintes resultados de um experimento com dois fatores e dois níveis para o fator A e três níveis para o fator B. Cada tratamento tem três réplicas.

A	B	Média	Desvio-Padrão
1	1	21,33333	6,027714
1	2	20	7,549834
1	3	32,66667	3,511885
2	1	31	6,244998
2	2	33	6,557439
2	3	23	10

- Calcule a soma dos quadrados para cada fator e para a interação.
- Calcule a soma dos quadrados total e do erro.
- Complete a tabela da ANOVA com a estatística F.

14.S19 O tempo necessário de retrabalho para uma máquina dependeu da velocidade na qual a máquina foi operada (A), do lubrificante usado durante a operação (B) e da dureza do metal usado na máquina (C). Dois níveis de cada fator foram escolhidos e uma única réplica de um experimento  $2^3$  foi corrida. Os dados do tempo (em horas) de retrabalho são mostrados na seguinte tabela.

Combinação de Tratamentos	Tempo (em horas)
(1)	27
a	34
b	38
ab	59
c	44
ac	40
bc	63
abc	37

- Esses tratamentos não podem ser corridos sob as mesmas condições. Estabeleça um planejamento para correr essas observações em dois blocos de quatro observações cada, com ABC superposto com os blocos.
- Análise os dados.

14.S20 Um artigo publicado na revista *Process Biochemistry* (1996, Vol. 31(8), pp. 773-785) apresentou um planejamento fatorial

fracionário 2<sup>7-3</sup> para avaliar a otimização do desempenho do meio de biolixiviação de dióxido de manganês. Os dados são mostrados na tabela seguinte.

Corridas	A	B	C	D	E	F	G	Rendimento da Extração de Manganês (%)
1	-1	-1	-1	-1	-1	-1	-1	99,0
2	1	-1	-1	-1	1	-1	1	97,4
3	-1	1	-1	-1	1	1	1	97,7
4	1	1	-1	-1	-1	1	-1	90,0
5	-1	-1	1	-1	1	1	-1	100,0
6	1	-1	1	-1	-1	1	1	98,0
7	-1	1	1	-1	-1	-1	1	90,0
8	1	1	1	-1	1	-1	-1	93,5
9	-1	-1	-1	1	-1	1	1	100,0
10	1	-1	-1	1	1	1	-1	98,6
11	-1	1	-1	1	1	-1	-1	97,1
12	1	1	-1	1	-1	-1	1	92,4
13	-1	-1	1	1	1	-1	1	93,0
14	1	-1	1	1	-1	-1	-1	95,0
15	-1	1	1	1	-1	1	-1	97,0
16	1	1	1	1	1	1	1	98,0

Os fatores e os níveis são mostrados na tabela seguinte.

Fator	-1	+1
A	Concentração de mineral (%)	10      20
B	Melaço (g/litro)	100      200
C	NH <sub>4</sub> NO <sub>3</sub> (g/litro)	1,25      2,50
D	KH <sub>2</sub> PO <sub>4</sub> (g/litro)	0,75      1,50
E	MgSO <sub>4</sub> (g/litro)	0,5      1,00
F	Etrato de levedura (g/litro)	0,20      0,50
G	NaHCO <sub>3</sub> (g/litro)	2,00      4,00

- Escreva a relação completa de definição do planejamento e os pares associados desse planejamento.
- Estime os efeitos principais.
- Plote as estimativas dos efeitos em um papel de probabilidade normal e interprete os resultados.
- Faça uma análise residual.

14.S21 Considere a tabela seguinte de ANOVA de um experimento fatorial com dois fatores.

**ANOVA com dois fatores: y versus A, B**

Fonte	GL	SQ	MQ	F	P
A	3	1213770	404590	?	0,341
B	2	?	17335441	58,30	0,000
Erro	?	1784195	?		
Total	11	37668847			

- Quantos níveis de cada fator foram usados no experimento?
- Quantas réplicas foram usadas?
- Qual suposição é feita para obter uma estimativa de erro?
- Calcule as informações que faltam (denotadas por “?”) na tabela da ANOVA.

## Capítulo 15 Exercícios

### Exercícios para a Seção 15.3

15.3.1 Gráficos de controle são construídos para amostras de tamanho  $n = 4$ , sendo  $\bar{x}$  e  $s$  calculados para cada uma das 20 amostras preliminares, como segue:

$$\sum_{i=1}^{20} \bar{x}_i = 4460 \quad \sum_{i=1}^{20} s_i = 271,6$$

- Encontre os limites de controle iniciais para os gráficos  $\bar{X}$  e  $S$ .
- Considerando que o processo esteja sob controle, estime sua média e seu desvio-padrão.

15.3.2 Vinte e cinco amostras de tamanho cinco são retiradas de um processo, em intervalos de uma hora, sendo obtidos os seguintes dados:

$$\sum_{i=1}^{25} \bar{x}_i = 362,75 \quad \sum_{i=1}^{25} r_i = 8,60 \quad \sum_{i=1}^{25} s_i = 3,64$$

- Encontre os limites de controle iniciais para os gráficos  $\bar{X}$  e  $R$ .
- Repita o item (a) para os gráficos  $\bar{X}$  e  $S$ .

15.3.3 O nível de colesterol (em mg/dL) é um importante índice para a saúde humana. O tamanho da amostra é  $n = 5$ . O seguinte sumário das estatísticas é obtido a partir das medidas de colesterol:

$$\sum_{i=1}^{30} \bar{x}_i = 140,03, \quad \sum_{i=1}^{30} r_i = 13,63, \quad \sum_{i=1}^{30} s_i = 5,10$$

- Encontre os limites de controle iniciais para os gráficos  $\bar{X}$  e  $R$ .
- Repita o item (a) para os gráficos  $\bar{X}$  e  $S$ .

15.3.4 Um gráfico de controle  $\bar{X}$  com limites de controle 3-sigma, tem  $LSC = 48,75$  e  $LIC = 42,71$ . Suponha que o desvio-padrão do processo seja  $\sigma = 2,25$ . Que tamanho do subgrupo foi usado para o gráfico?

15.3.5 Monitore-se a resistência à tração de um fio de chumbo colado a uma placa de circuito integrado. A seguinte tabela fornece dados para 20 amostras, cada uma tendo um tamanho igual a 3.

Número da Amostra	$x_1$	$x_2$	$x_3$
1	15,4	15,6	15,3
2	15,4	17,1	15,2
3	16,1	16,1	13,5
4	13,5	12,5	10,2
5	18,3	16,1	17,0
6	19,2	17,2	19,4
7	14,1	12,4	11,7
8	15,6	13,3	13,6
9	13,9	14,9	15,5
10	18,7	21,2	20,1
11	15,3	13,1	13,7
12	16,6	18,0	18,0
13	17,0	15,2	18,1
14	16,3	16,5	17,7
15	8,4	7,7	8,4
16	11,1	13,8	11,9
17	16,5	17,1	18,5
18	18,0	14,1	15,9
19	17,8	17,3	12,0
20	11,5	10,8	11,2

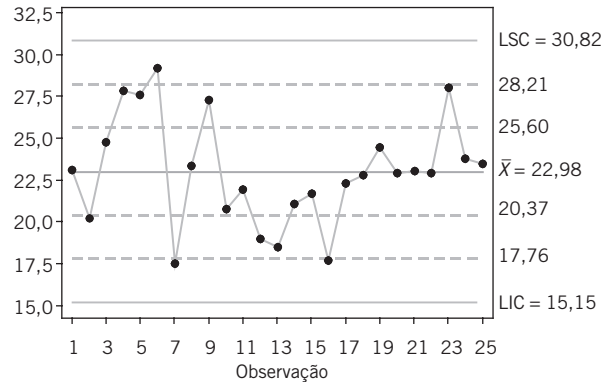
- Use todos os dados para determinar os limites de controle iniciais para os gráficos  $\bar{X}$  e  $R$ , construa o gráfico e plote os dados.
- Use os limites de controle do item (a) para identificar pontos fora de controle. Se necessário, reveja os limites de controle calculados, considerando que qualquer amostra fora dos limites de controle poderá ser eliminada.
- Repita os itens (a) e (b) para os gráficos  $\bar{X}$  e  $S$ .

15.3.6 O teor de cobre de um banho de chapeamento é medido três vezes ao dia, sendo os resultados reportados em ppm. Os valores de  $\bar{x}$  e  $r$  para 25 dias são mostrados na tabela a seguir.

Dia	$\bar{x}$	$r$	Dia	$\bar{x}$	$r$
1	5,45	1,21	14	7,01	1,45
2	5,39	0,95	15	5,83	1,37
3	6,85	1,43	16	6,35	1,04
4	6,74	1,29	17	6,05	0,83
5	5,83	1,35	18	7,11	1,35
6	7,22	0,88	19	7,32	1,09
7	6,39	0,92	20	5,90	1,22
8	6,50	1,13	21	5,50	0,98
9	7,15	1,25	22	6,32	1,21
10	5,92	1,05	23	6,55	0,76
11	6,45	0,98	24	5,90	1,20
12	5,38	1,36	25	5,95	1,19
13	6,03	0,83			

- Usando todos os dados, encontre os limites de controle iniciais para os gráficos  $\bar{X}$  e  $R$ , construa o gráfico e plote os dados. O processo está sob controle estatístico?
- Se necessário, reveja os limites de controle calculados no item (a), considerando que qualquer amostra fora dos limites de controle poderá ser eliminada.
- Repita os itens (a) e (b) para os gráficos  $\bar{X}$  e  $S$ .

15.3.7 Aplique as Regras *Western Electric* ao seguinte gráfico de controle  $\bar{x}$ . Os limites de advertência são mostrados como linhas tracejadas. Descreva qualquer uma das violações das regras.



15.3.8 Os dados seguintes foram considerados em *Quality Engineering* [“An SPC Case Study on Stabilizing Syringe Lengths” (1999-2000, Vol. 12(1), pp. 65-71)]. O comprimento de seringas é medido durante um processo de fabricação farmacêutica. A tabela a seguir fornece dados (em polegadas) para 20 amostras, cada uma de tamanho 5.

Amostra	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
1	4,960	4,946	4,950	4,956	4,958
2	4,958	4,927	4,935	4,940	4,950
3	4,971	4,929	4,965	4,952	4,938
4	4,940	4,982	4,970	4,953	4,960
5	4,964	4,950	4,953	4,962	4,956
6	4,969	4,951	4,955	4,966	4,954
7	4,960	4,944	4,957	4,948	4,951
8	4,969	4,949	4,963	4,952	4,962
9	4,984	4,928	4,960	4,943	4,955
10	4,970	4,934	4,961	4,940	4,965
11	4,975	4,959	4,962	4,971	4,968
12	4,945	4,977	4,950	4,969	4,954
13	4,976	4,964	4,970	4,968	4,972
14	4,970	4,954	4,964	4,959	4,968
15	4,982	4,962	4,968	4,975	4,963
16	4,961	4,943	4,950	4,949	4,957
17	4,980	4,970	4,975	4,978	4,977
18	4,975	4,968	4,971	4,969	4,972
19	4,977	4,966	4,969	4,973	4,970
20	4,975	4,967	4,969	4,972	4,972

- Usando todos os dados, encontre os limites de controle iniciais para os gráficos  $\bar{X}$  e  $R$ , construa o gráfico e plote os dados. Esse processo está sob controle estatístico?
- Use os limites de controle do item (a) para identificar pontos fora de controle. Se necessário, reveja os limites de controle calculados, considerando que qualquer amostra fora dos limites de controle poderá ser eliminada.
- Repita os itens (a) e (b) para os gráficos  $\bar{X}$  e  $S$ .

15.3.9 Considere os dados no Exercício 15.3.5. Calcule o desvio-padrão da amostra de todas as 60 medidas e compare esse resultado à estimativa de  $\sigma$  obtida a partir de seus gráficos revisados de  $\bar{X}$  e  $R$ . Explique qualquer diferença.

15.3.10 O tráfego na internet pode ser medido para ajudar a realçar problemas de segurança ou indicar uma falta de potencial de largura de banda. Dados sobre o tráfego na internet (em milhares de solicitações), provenientes de [http://en.wikipedia.org/wiki/Web\\_traffic](http://en.wikipedia.org/wiki/Web_traffic), são apresentados na seguinte tabela para 25 amostras, cada uma tendo um tamanho igual a 4.

Amostra	$x_1$	$x_2$	$x_3$	$x_4$
1	163,95	164,54	163,87	165,10
2	163,30	162,85	163,18	165,10
3	163,13	165,14	162,80	163,81
4	164,08	163,43	164,03	163,77
5	165,44	163,63	163,95	164,78
6	163,83	164,14	165,22	164,91
7	162,94	163,64	162,30	163,78
8	164,97	163,68	164,73	162,32
9	165,04	164,06	164,40	163,69
10	164,74	163,74	165,10	164,32
11	164,72	165,75	163,07	163,84
12	164,25	162,72	163,25	164,14
13	164,71	162,63	165,07	162,59
14	166,61	167,07	167,41	166,10
15	165,23	163,40	164,94	163,74
16	164,27	163,42	164,73	164,88
17	163,59	164,84	164,45	164,12
18	164,90	164,20	164,32	163,98

Amostra	$x_1$	$x_2$	$x_3$	$x_4$
19	163,98	163,53	163,34	163,82
20	164,08	164,33	162,38	164,08
21	165,71	162,63	164,42	165,27
22	164,03	163,36	164,55	165,77
23	160,52	161,68	161,18	161,33
24	164,22	164,27	164,35	165,12
25	163,93	163,96	165,05	164,52

- a. Use todos os dados para determinar os limites de controle iniciais para os gráficos  $\bar{X}$  e  $R$ , construa o gráfico e plote os dados.
- b. Use os limites de controle do item (a) para identificar pontos fora de controle. Se necessário, reveja seus limites de controle, considerando que qualquer amostra fora dos limites de controle poderá ser eliminada.

15.3.11 Um gráfico de controle  $\bar{X}$ , com limites de controle 3-sigma e tamanho de subgrupo  $n = 4$ , tem limites de controle  $LSC = 48,75$  e  $LIC = 40,55$ .

- a. Estime o desvio-padrão do processo.
- b. A resposta do item (a) depende de se  $\bar{r}$  ou  $\bar{s}$  foi usado para construir o gráfico de controle de  $\bar{X}$ ?

15.3.12 Um artigo publicado em *Quality & Safety in Health Care* ["Statistical Process Control as a Tool for Research and Healthcare Improvement" (2003, Vol. 12, pp. 458-464)] considerou um número de gráficos de controle em assistência médica. Os dados aproximados a seguir foram usados para construir gráficos  $\bar{X} - S$  para o tempo de resposta (TR, em minutos) para hemogramas completos. O tamanho do subgrupo é  $n = 3$  por turno, e o desvio-padrão da média é 21. Construa o gráfico de  $\bar{X}$  e comente sobre o controle do processo. Se necessário, suponha que as causas atribuídas podem ser encontradas; elimine os pontos suspeitos e reveja os limites de controle.

$t$	1	2	3	4	5	6	7	8	9	10	11	12	13	14
TR	51	73	28	52	65	49	51	50	25	39	40	30	49	31

### Exercícios para a Seção 15.4

15.4.1 Em um processo de fabricação de semicondutores por deposição química a vapor (DQV), a espessura do metal foi medida em 30 pastilhas obtidas ao longo de aproximadamente duas semanas. Os dados são mostrados na tabela seguinte.

- a. Usando todos os dados, calcule os limites de controle iniciais para os gráficos das observações individuais e da amplitude móvel. Construa o gráfico e plote os dados. Determine se o processo está sob controle estatístico. Se não, considere que causas atribuídas possam ser encontradas para eliminar essas amostras e reveja os limites de controle.
- b. Estime a média e o desvio-padrão do processo quando ele estiver sob controle.

Pastilha	$x$	Pastilha	$x$	Pastilha	$x$
1	16,8	11	17,1	21	15,2
2	14,9	12	15,9	22	16,7
3	18,3	13	16,4	23	15,2
4	16,5	14	15,8	24	14,7
5	17,1	15	15,4	25	17,9
6	17,4	16	15,4	26	14,8
7	15,9	17	14,3	27	17,0
8	14,4	18	16,1	28	16,2
9	15,0	19	15,8	29	15,6
10	15,7	20	15,9	30	16,3

15.4.2 A viscosidade de um intermediário químico é medida a cada hora. Vinte amostras, cada uma de tamanho  $n = 1$ , estão na tabela seguinte.

Amostra	Viscosidade	Amostra	Viscosidade	Amostra	Viscosidade
1	495	8	504	15	497
2	491	9	542	16	499
3	501	10	508	17	468
4	501	11	493	18	486
5	512	12	507	19	511
6	540	13	503	20	487
7	492	14	475		

a. Usando todos os dados, calcule os limites de controle iniciais para os gráficos das observações individuais e da amplitude móvel. Determine se o processo está sob controle estatístico. Se não, considere que causas atribuídas possam ser encontradas para eliminar essas amostras e reveja os limites de controle.

b. Estime a média e o desvio-padrão do processo quando ele estiver sob controle.

15.4.3 A tabela a seguir de dados foi analisada em *Quality Engineering* [1991-1992, Vol. 4(1)]. O tamanho médio de partícula da matéria-prima foi obtido a partir de 25 amostras sucessivas.

Observação	Tamanho	Observação	Tamanho	Observação	Tamanho
1	96,1	10	96,8	19	87,4
2	94,4	11	100,9	20	96,1
3	116,2	12	117,7	21	97,1
4	98,8	13	115,6	22	95,7
5	95,0	14	100,5	23	94,2
6	120,3	15	103,1	24	102,4
7	104,8	16	93,1	25	131,9
8	88,4	17	93,7		
9	106,8	18	72,4		

a. Usando todos os dados, calcule os limites de controle iniciais para os gráficos das observações individuais e da amplitude móvel. Construa o gráfico e plote os dados. Determine se o processo está sob controle estatístico. Se não, considere que causas atribuídas possam ser encontradas para eliminar essas amostras e reveja os limites de controle.

b. Estime a média e o desvio-padrão do processo quando ele estiver sob controle.

15.4.4 A técnica de deposição por *laser* pulsado consiste em uma deposição em um filme fino, por meio de um raio *laser* de alta potência. Vinte e cinco filmes foram depositados por essa técnica. A espessura dos filmes obtidos é mostrada na tabela da coluna ao lado.

a. Usando todos os dados, calcule os limites de controle iniciais para os gráficos das observações individuais e da amplitude móvel. Determine se o processo está sob controle estatístico. Se não, considere que causas atribuídas possam ser encontradas para eliminar essas amostras e reveja os limites de controle.

b. Estime a média e o desvio-padrão do processo quando ele estiver sob controle.

Filme	Espessura (em nm)	Filme	Espessura (em nm)	Filme	Espessura (em nm)
1	28	10	35	19	56
2	45	11	47	20	49
3	34	12	50	21	21
4	29	13	32	22	62
5	37	14	40	23	34
6	52	15	46	24	31
7	29	16	59	25	98
8	51	17	20		
9	23	18	33		

15.4.5 Um artigo publicado no *Journal of the Operational Research Society* ["A Quality Control Approach for Monitoring Inventory Stock Levels" (1993, Vol. 44, pp. 1115-1127)] reportou um gráfico de controle para monitorar a acurácia de um sistema de gerenciamento de registro. A acurácia do registro no tempo  $t$ ,  $AC(t)$ , é definida como a diferença entre o registro gravado e o real (em valor absoluto) dividido pelo registro gravado. Consequentemente,  $AC(t)$  varia entre 0 e 1, em que quanto menor o valor, melhor. Dados extraídos são mostrados na tabela a seguir:

$t$	1	2	3	4	5	6	7	8
$AC(t)$	0,190	0,050	0,095	0,055	0,090	0,200	0,030	0,105
$t$	9	10	11	12	13	14	15	16
$AC(t)$	0,115	0,103	0,121	0,089	0,180	0,122	0,098	0,173
$t$	17	18	19	20	21	22		
$AC(t)$	0,298	0,075	0,083	0,115	0,147	0,079		

a. Calcule os gráficos individuais e de amplitude móvel para esses dados.

b. Comente sobre o controle do processo. Se necessário, suponha que causas atribuídas podem ser encontradas, elimine pontos suspeitos e reveja os limites de controle.

15.4.6 Um artigo publicado em *Quality & Safety in Health Care* ["Statistical Process Control as a Tool for Research and Healthcare Improvement" (2003, Vol. 12, pp. 458-464)] considerou um número de gráficos de controle em assistência médica. Um gráfico de  $X$  foi construído para a quantidade de resíduos infecciosos descartados (em libras) todo dia. O artigo menciona que resíduo infeccioso classificado imprópriamente (que realmente não seja perigoso) aumenta substancialmente os custos de hospitais a cada ano. As seguintes tabelas mostram dados aproximados para os resíduos médios diários a cada mês, antes e depois de mudanças no processo, respectivamente. A mudança no processo incluiu uma campanha educativa para fornecer uma definição operacional para resíduo infeccioso.

#### Antes da Mudança no Processo

Mês	1	2	3	4	5	6	7	8	9
Resíduo	6,9	6,8	6,9	6,7	6,9	7,5	7	7,4	7
Mês	13	14	15	16	17	18	19	20	21
Resíduo	7,5	7,4	6,5	6,9	7,0	7,2	7,8	6,3	6,7

#### Depois da Mudança no Processo

Mês	1	2	3	4	5	6	7	8	9	10
Resíduo	5,0	4,8	4,4	4,3	4,6	4,3	4,5	3,5	4,0	4,1
Mês	11	12	13	14	15	16	17	18	19	20
Resíduo	3,8	5,0	4,6	4,0	5,0	4,9	4,9	5,0	6,0	4,5
Mês	21	22	23	24	25	26	27	28	29	30
Resíduo	4,0	5,0	4,5	4,6	4,6	3,8	5,3	4,5	4,4	3,8

a. Trabalhe separadamente os dados de antes e de depois da mudança no processo e construa gráficos individuais e de amplitude móvel para cada conjunto de dados. Admita que causas atribuídas podem ser encontradas e elimine observações suspeitas. Se necessário, reveja os limites de controle.

b. Comente sobre o controle de cada gráfico e diferenças entre os gráficos. A mudança no processo foi efetiva?

## Exercícios para a Seção 15.5

15.5.1 Suponha que a característica da qualidade seja normalmente distribuída, com especificações de 20 a 32 unidades.

a. Que valor de  $\sigma$  é necessário para que  $PCR$  seja igual a 1,5?

b. Qual o valor da média do processo que minimiza a fração defeituosa? Essa escolha para a média depende do valor de  $\sigma$ ?

15.5.2 Suponha que a característica da qualidade seja normalmente distribuída com especificações em  $100 \pm 20$ . O desvio-padrão do processo é 6.

a. Suponha que a média do processo seja 100. Quais são os limites naturais de tolerância? Qual é a fração defeituosa? Calcule  $PCR$  e  $PCR_k$  e interprete essas razões.

b. Suponha que a média do processo seja 106. Quais são os limites naturais de tolerância? Qual é a fração defeituosa? Calcule  $PCR$  e  $PCR_k$  e interprete essas razões.

15.5.3 Um processo normalmente distribuído usa 66,7 % da banda de especificação. Ele está centralizado na dimensão nominal, localizada na metade do caminho entre os limites superior e inferior de especificação.

a. Estime  $PCR$  e  $PCR_k$ . Interprete essas razões.

b. Qual o nível produzido de fração defeituosa (*fallout*)?

15.5.4 Reconsidere o Exercício 15.3.2, em que os limites de especificação são  $14,50 \pm 0,50$ .

a. Que conclusões podem ser tiradas sobre a habilidade do processo para operar dentro desses limites? Estime a percentagem de itens defeituosos que serão produzidos.

b. Estime  $PCR$  e  $PCR_k$ . Interprete essas razões.

15.5.5 Reconsidere o Exercício 15.3.1. Suponha que a característica da qualidade seja normalmente distribuída, com especificação em  $220 \pm 50$ . Qual é a proporção fora de especificação? Estime e interprete  $PCR$  e  $PCR_k$ .

15.5.6 Reconsidere as medidas do teor de cobre do Exercício 15.3.6. Dado que as especificações são  $6,0 \pm 1,0$ , estime  $PCR$  e  $PCR_k$  e interprete essas razões.

15.5.7 Suponha que a característica de qualidade seja normalmente distribuída com as especificações em  $120 \pm 20$ . O desvio-padrão do processo é 6,5.

a. Suponha que a média do processo seja 120. Quais são os limites de tolerância naturais? Qual é a fração defeituosa? Calcule  $PCR$  e  $PCR_k$  e interprete essas razões.

b. Suponha que a média do processo se desloque do centro por 1,5 desvio-padrão, em direção ao limite superior de especificação. Recalcule as grandezas do item (a).

c. Compare os resultados dos itens (a) e (b) e comente qualquer diferença.

15.5.8 Reconsidere as medidas da viscosidade do Experimento 15.4.2. As especificações são  $500 \pm 25$ . Calcule as estimativas das razões das capacidades de processo,  $PCR$  e  $PCR_k$ , para esse processo e forneça uma interpretação.

15.5.9 A média de um processo é centralizada entre os limites de especificação e  $PCR = 1,33$ . Considere que a média do processo aumente por  $1,5\sigma$ .

a. Calcule o  $PCR$  e o  $PCR_k$  para o processo deslocado.

b. Calcule a fração defeituosa estimada do processo deslocado e compare seu resultado àqueles da Tabela 15.4. Admita uma distribuição normal para a medida.

15.5.10 O  $PCR$  para uma medida é 1,5 e os limites de controle para um gráfico  $\bar{X}$  com  $n = 4$  são 24,6 e 32,6.

a. Estime o desvio-padrão do processo  $\sigma$ .

b. Considere que os limites de especificação sejam especificados ao redor da média do processo. Calcule os limites de especificação.

## Exercícios para a Seção 15.6

15.6.1 Um exemplo antigo de CEP foi descrito em *Industrial Quality Control* ["The Introduction of Quality Control at Colonial Radio Corporation" (1944, Vol. 1(1), pp. 4-9)]. A seguir, são dadas as frações defeituosas das montagens eixo e arruela, durante o mês de abril em amostras de  $n = 1500$  cada:

a. Faça um gráfico  $P$  para esse processo. Esse processo está sob controle estatístico?

b. Suponha que, em vez de  $n = 1500$ ,  $n = 100$ . Use os dados fornecidos para fazer um gráfico  $P$  para esse processo. Reveja os limites de controle, se necessário.

c. Compare seus limites de controle para os gráficos  $P$  dos itens (a) e (b). Explique por que eles diferem. Também, explique por que sua afirmação sobre controle estatístico difere para os dois tamanhos de  $n$ .

Amostra	Fração Defeituosa	Amostra	Fração Defeituosa	Amostra	Fração Defeituosa
1	0,11	8	0,03	15	0,04
2	0,06	9	0,02	16	0,04
3	0,1	10	0,03	17	0,04
4	0,11	11	0,03	18	0,03
5	0,14	12	0,03	19	0,06
6	0,11	13	0,04	20	0,06
7	0,14	14	0,07		

15.6.2 Os números a seguir representam o número de defeitos por 1000 pés em um fio recoberto com borracha: 1, 1, 3, 7, 8, 10, 5, 13, 0, 19, 24, 6, 9, 11, 15, 8, 3, 6, 7, 4, 9, 20, 11, 7, 18, 10, 6, 4, 0, 9, 7, 3, 1, 8, 12. Os dados são provenientes de um processo controlado?

15.6.3 Os números a seguir representam o número de defeitos de soldagem observados em 24 amostras de cinco circuitos impressos: 7, 6, 8, 10, 24, 6, 5, 4, 8, 11, 15, 8, 4, 16, 11, 12, 8, 6, 5, 9, 7, 14, 8, 21.

- Usando todos os dados, calcule os limites de controle iniciais para um gráfico de controle  $U$ , construa o gráfico e plote os dados.
- Podemos concluir que o processo está sob controle, usando um gráfico  $U$ ? Se não, considere que causas atribuídas possam ser encontradas, liste pontos e reveja os limites de controle.

15.6.4 Os seguintes dados representam o número de erros de ortografia para cada 1000 palavras em um *site* de notícias na internet ao longo de 20 semanas.

Semana	Número de Erros de Ortografia	Semana	Número de Erros de Ortografia
1	3	11	1
2	6	12	6
3	0	13	9
4	5	14	8
5	9	15	6
6	5	16	4
7	2	17	13
8	2	18	3
9	3	19	0
10	2	20	7

- Que gráfico de controle é mais apropriado para esses dados?
- Usando todos os dados, calcule os limites de controle para o gráfico do item (a), construa o gráfico e plote os dados.
- Determine se o processo está sob controle estatístico. Se não, considere que causas atribuídas possam ser encontradas e que os pontos fora de controle possam ser eliminados. Reveja os limites de controle.

15.6.5 Um artigo publicado em *Epilepsy Research* ["Statistical Process Control (SPC): A Simple Objective Method for Monitoring Seizure Frequency and Evaluating Effectiveness of Drug Interventions in Refractory Childhood Epilepsy" (2010, Vol. 91, pp. 205-213)] usou gráficos de controle para monitorar semanalmente mudanças convulsivas em pacientes com epilepsia infantil contumaz.

A tabela seguinte mostra dados representativos de observações semanais de frequência convulsiva (FC).

Semana	1	2	3	4	5	6	7	8	9	10
FC	13	10	17	10	18	14	10	12	16	13
Semana	11	12	13	14	15	16	17	18	19	20
FC	14	11	8	11	10	3	2	13	15	21
Semana	21	22	23	24	25					
FC	15	12	14	18	12					

- Que tipo de gráfico de controle é mais apropriado para esses dados? Construa esse gráfico.
- Comente a respeito do controle do processo.
- Se necessário, considere que as causas atribuídas possam ser encontradas, elimine os pontos suspeitos e reveja os limites de controle.
- Na publicação, os FCs semanais foram aproximados como normalmente distribuídos, e um gráfico com os pontos individuais foi construído. Construa esse gráfico e compare-o ao gráfico de atributos que você construiu no item (a).

15.6.6 Um artigo publicado em *Graefe's Archive for Clinical and Experimental Ophthalmology* ["Statistical Process Control Charts for Ophthalmology" (2011, Vol. 249, pp. 1103-1105)] considerou o número mensal de casos de cirurgia de catarata. Os dados são mostrados na tabela seguinte.

Janeiro	Fevereiro	Março	Abril	Maior	Junho	Julho
61	88	80	68	80	70	60
Agosto	Setembro	Outubro	Novembro	Dezembro		
56	72	118	106	60		

- Que tipo de gráfico de controle é mais apropriado para esses dados? Construa esse gráfico.
- Comente a respeito do controle do processo.
- Se necessário, considere que as causas atribuídas possam ser encontradas, elimine os pontos suspeitos e reveja os limites de controle.
- Na publicação, os dados foram aproximados como normalmente distribuídos e um gráfico com os pontos individuais foi construído. Construa esse gráfico e compare-o ao gráfico de atributos que você construiu no item (a). Por que gráficos para medidas individuais podem ser razoáveis?

## Exercícios para a Seção 15.7

15.7.1 Um gráfico  $\bar{X}$  usa amostras de tamanho 1. A linha central está em 100 e os limites, 3-sigma, superior e inferior de controle estão em 112 e 88, respectivamente.

- Qual é o  $\sigma$  do processo?
- Suponha que a média do processo se desloque para 96. Encontre a probabilidade de esse deslocamento ser detectado na próxima amostra.
- Encontre o CMC para detectar o deslocamento no item (b).

15.7.2 Um gráfico  $\bar{X}$  usa amostras de tamanho 4. A linha central está em 100 e os limites, 3-sigma, superior e inferior de controle estão em 106 e 94, respectivamente.

- Qual é o  $\sigma$  do processo?
- Suponha que a média do processo se desloque para 96. Encontre a probabilidade de esse deslocamento ser detectado na próxima amostra.
- Encontre o CMC para detectar o deslocamento no item (b).

15.7.3 Considere o gráfico de controle  $\bar{X}$  com  $LSC = 0,0635$ ,  $LIC = 0,0624$  e  $n = 5$ . Suponha que a média se desloque para 0,0625.

- Qual é a probabilidade de esse deslocamento ser detectado na próxima amostra?
- Qual é o CMC depois do deslocamento?

15.7.4 Considere o gráfico de controle  $\bar{X}$  com  $LSC = 14,708$ ,  $LIC = 14,312$  e  $n = 5$ . Suponha que a média se desloque para 14,6.

a. Qual é a probabilidade de esse deslocamento ser detectado na próxima amostra?

b. Qual é o CMC depois do deslocamento?

15.7.5 Um gráfico  $\bar{X}$  usa um subgrupo de tamanho 3. A linha central está em 200 e os limites superior e inferior de controle estão em 212 e 188, respectivamente.

a. Encontre a probabilidade de esse deslocamento ser detectado na próxima amostra.

b. Encontre o CMC para detectar o deslocamento no item (a).

15.7.6 Considere o gráfico de controle  $\bar{X}$  com  $LSC = 17,40$ ,  $LIC = 12,79$  e  $n = 3$ . Suponha que a média se desloque para 13.

a. Qual é a probabilidade de esse deslocamento ser detectado na próxima amostra?

b. Qual é o CMC depois do deslocamento?

15.7.7 Considere o gráfico  $P$  com tamanho de subgrupo  $n = 50$  e linha central em 0,12.

a. Calcule o  $LSC$  e o  $LIC$ .

b. Suponha que a proporção verdadeira de defeitos varie de 0,12 a 0,18. Qual é o CMC depois do deslocamento? Admita que as proporções amostrais sejam distribuídas de forma aproximadamente normal.

c. Refaça os itens (a) e (b) com  $n = 100$  e comente a diferença no CMC. O aumento do tamanho da amostra varia substancialmente o CMC?

15.7.8 Considere o gráfico  $U$  para placas de circuito impresso no Exemplo 15.3.3. A linha central = 1,6,  $LSC = 3,3$  e  $n = 5$ .

a. Calcule o  $LSC$  e  $LIC$ .

b. Suponha que a proporção verdadeira de defeitos se desloque de 1,6 a 2,4. Qual é o CMC depois do deslocamento? Admita que as proporções amostrais sejam distribuídas de forma aproximadamente normal.

c. Refaça o item (b) se os defeitos médios verdadeiros por unidade se deslocarem de 1,6 a 2,0 e comente a diferença no CMC.

## Exercícios para a Seção 15.8

15.8.1 Os seguintes dados foram considerados no artigo publicado em *Quality Engineering* ["Parabolic Control Limits for the Exponentially Weighted Moving Average Control Charts in Quality Engineering" (1992, Vol. 4(4), pp. 487-495)]. Em uma planta química, os dados para uma característica (viscosidade) da qualidade foram obtidos para o final de cada batelada de 12 horas. Os resultados de 15 medidas consecutivas são mostrados na tabela seguinte.

Batelada	Viscosidade	Batelada	Viscosidade
1	13,3	9	14,6
2	14,5	10	14,1
3	15,3	11	14,3
4	15,3	12	16,1
5	14,3	13	13,1
6	14,8	14	15,5
7	15,2	15	12,6
8	14,9		

a. Estabeleça um gráfico de controle MMPE para esse processo, com  $\lambda = 0,2$ . Considere o alvo desejado do processo como 14,1. O processo parece estar sob controle?

b. Suponha que as cinco próximas observações sejam 14,6; 15,3; 15,7; 16,1 e 16,8. Aplique a MMPE do item (a) a essas novas observações. Existe qualquer evidência de que o processo tenha se deslocado para a condição de fora de controle?

15.8.2 A pureza de um produto químico é medida a cada 2 horas. Os resultados de 20 medidas consecutivas são apresentados a seguir.

Amostra	Pureza	Amostra	Pureza
1	89,11	11	88,55
2	90,59	12	90,43
3	91,03	13	91,04
4	89,46	14	88,17

Amostra	Pureza	Amostra	Pureza
5	89,78	15	91,23
6	90,05	16	90,92
7	90,63	17	88,86
8	90,75	18	90,87
9	89,65	19	90,73
10	90,15	20	89,78

Use  $\sigma = 0,8$  e considere que o valor-alvo do processo seja 90.

a. Construa um gráfico de controle MMPE para esse processo com  $\lambda = 0,2$ . O processo parece estar sob controle?

b. Construa um gráfico de controle MMPE para esse processo com  $\lambda = 0,5$ . Compare seus resultados com o item (a).

c. Suponha que as cinco próximas observações sejam 90,75; 90,00; 91,15; 90,95 e 90,86. Aplique a MMPE dos itens (a) e (b) a essas novas observações. Existe qualquer evidência de que o processo tenha se deslocado para a condição de fora de controle?

15.8.3 Considere um gráfico de controle MMPE. O valor-alvo para o processo é  $\mu_0 = 50$  e  $\sigma = 2$ . Use a Tabela 15.8.

a. Se o tamanho da amostra fosse  $n = 1$ , você preferiria um gráfico MMPE com  $\lambda = 0,1$  e  $L = 2,81$  ou  $\lambda = 0,5$  e  $L = 3,07$  para detectar um deslocamento na média do processo para  $\mu = 52$  em média? Por quê?

b. Se o tamanho da amostra aumentar para  $n = 4$ , que gráfico do item (a) você prefere? Por quê?

c. Se um gráfico MMPE com  $\lambda = 0,1$  e  $L = 2,81$  for usado, que tamanho de amostra será necessário para detectar um deslocamento para  $\mu = 52$  em aproximadamente três amostras em média?

15.8.4 A concentração de um produto químico é medida considerando quatro amostras provenientes de cada batelada de material. A concentração média dessas medidas é mostrada para as 20 bateladas na seguinte tabela:

Batelada	Concentração	Batelada	Concentração
1	104,5	11	95,4
2	99,9	12	94,5
3	106,7	13	104,5
4	105,2	14	99,7
5	94,8	15	97,7
6	94,6	16	97
7	104,4	17	95,8
8	99,4	18	97,4
9	100,3	19	99
10	100,3	20	102,6

Use  $\sigma = 8$  e considere que o valor-alvo desejado para o processo seja 100.

- Construa um gráfico de controle MMPE para esse processo, com  $\lambda = 0,2$ . O processo parece estar sob controle?
- Construa um gráfico de controle MMPE para esse processo, com  $\lambda = 0,5$ . Compare seus resultados com aqueles do item (a).
- Se a concentração se deslocar para 104, você preferiria o gráfico do item (a) ou do (b)? Explique.

**15.8.5** A frequência cardíaca (contagens/minuto) é medida a cada 30 minutos. Os resultados de 20 medidas consecutivas são mostrados a seguir:

Amostra	Frequência Cardíaca	Amostra	Frequência Cardíaca
1	68	11	79
2	71	12	79
3	67	13	78
4	69	14	78
5	71	15	78
6	70	16	79
7	69	17	79
8	67	18	82
9	70	19	82
10	70	20	81

Use  $\mu = 70$  e  $\sigma = 3$ .

- Construa um gráfico de controle MMPE com  $\lambda = 0,1$ . Use  $L = 2,81$ . O processo parece estar sob controle?
- Construa um gráfico de controle MMPE com  $\lambda = 0,5$ . Use  $L = 3,07$ . Compare seus resultados com os do item (a).
- Se a frequência cardíaca se deslocar para 76, aproxime as CMCs para os gráficos dos itens (a) e (b).

**15.8.6** O número (em milhares) de pacientes com influenza que visitam semanalmente hospitais é mostrado na seguinte tabela.

Amostra	Número de Pacientes	Amostra	Número de Pacientes
1	162,27	13	159,989
2	157,47	14	159,09
3	157,065	15	162,699
4	160,45	16	163,89
5	157,993	17	164,247
6	162,27	18	162,7
7	160,652	19	164,859
8	159,09	20	163,65
9	157,442	21	165,99
10	160,78	22	163,22
11	159,138	23	164,338
12	161,08	24	164,83

Use  $\mu = 160$  e  $\sigma = 2$ .

- Construa um gráfico de controle MMPE com  $\lambda = 0,1$ . Use  $L = 2,81$ . O processo parece estar sob controle?
- Construa um gráfico de controle MMPE com  $\lambda = 0,5$ . Use  $L = 3,07$ . Compare seus resultados com os do item (a).

## Exercícios para a Seção 15.10

**15.10.1** Suponha que o custo de um reparo maior sem a garantia estendida no Exemplo 15.8 seja mudada para \$ 1000. Determine a decisão selecionada, baseando-se nos critérios de minimax, de mais provável e de custo esperado.

**15.10.2** Reconsidere a decisão da garantia estendida do Exemplo 15.8. Suponha que as probabilidades dos estados de maior reparo, de menor reparo e de nenhum reparo sejam mudadas para 0,2; 0,4 e 0,4, respectivamente. Determine a decisão estabelecida, baseando-se no critério minimax, de mais provável e de custo esperado.

**15.10.3** Analise o Exemplo 15.9 baseado no critério de mais provável e determine as ações que são selecionadas em cada nó de decisão. Alguma ação difere daquelas selecionadas no exemplo?

**15.10.4** Analise o Exemplo 15.9 baseado no critério de lucro esperado e determine as ações que são selecionadas em cada nó de decisão. Alguma ação difere daquelas selecionadas no exemplo?

## Exercícios Suplementares para o Capítulo 15

15.S7 O diâmetro dos pinos de fusíveis, usados em uma aplicação de motor de avião, é uma característica importante da qualidade. Vinte e cinco amostras, de três pinos cada uma, são mostradas a seguir:

Número da Amostra	Diâmetro		
1	64,030	64,002	64,019
2	63,995	63,992	64,001
3	63,988	64,024	64,021
4	64,002	63,996	63,993
5	63,992	64,007	64,015
6	64,009	63,994	63,997
7	63,995	64,006	63,994
8	63,985	64,003	63,993
9	64,008	63,995	64,009
10	63,998	74,000	63,990
11	63,994	63,998	63,994
12	64,004	64,000	64,007
13	63,983	64,002	63,998
14	64,006	63,967	63,994
15	64,012	64,014	63,998
16	64,000	63,984	64,005
17	63,994	64,012	63,986
18	64,006	64,010	64,018
19	63,984	64,002	64,003
20	64,000	64,010	64,013
21	63,988	64,001	64,009
22	64,004	63,999	63,990
23	64,010	63,989	63,990
24	64,015	64,008	63,993
25	63,982	63,984	63,995

- Estabeleça os gráficos  $\bar{X}$  e  $R$  para esse processo. Se necessário, reveja os limites de modo que nenhuma observação esteja fora de controle.
- Estime a média e o desvio-padrão do processo.
- Suponha que as especificações do processo sejam  $64 \pm 0,02$ . Calcule uma estimativa de  $PCR$ . O processo encontra um nível mínimo de capacidade de  $PCR \geq 1,33$ ?
- Calcule uma estimativa de  $PCR_k$ . Use essa razão para tirar conclusões sobre a capacidade de processo.
- De modo a tornar esse um processo 6-sigma, a variância  $\sigma^2$  teria de ser diminuída de tal forma que  $PCR_k = 2,0$ . Qual deve ser o valor da nova variância?
- Suponha que a média se desloque para 64,01. Qual é a probabilidade de esse deslocamento ser detectado na próxima amostra? Qual é o CMC depois desse deslocamento?

15.S8 Refaça o Exercício 15.S7 com os gráficos  $\bar{X}$  e  $S$ .

15.S9 Garrafas de plástico para detergentes líquidos de lavanderia são formadas pela moldagem de sopra. Vinte amostras de  $n = 100$

garrafas são inspecionadas em uma ordem de tempo de produção, sendo a fração defeituosa reportada em cada amostra. Os dados são mostrados a seguir:

Amostra	Fração Defeituosa	Amostra	Fração Defeituosa
1	0,12	11	0,10
2	0,15	12	0,07
3	0,18	13	0,12
4	0,10	14	0,08
5	0,12	15	0,09
6	0,11	16	0,15
7	0,05	17	0,10
8	0,09	18	0,06
9	0,13	19	0,12
10	0,13	20	0,13

- Estabeleça um gráfico  $P$  para esse processo. O processo está sob controle estatístico?
- Suponha que, em vez de  $n = 100$ ,  $n = 200$ . Use os dados fornecidos para estabelecer um gráfico  $P$  para esse processo. Reveja os limites de controle, se necessário.
- Compare seus limites de controle para os gráficos  $P$  nos itens (a) e (b). Explique por que eles diferem. Também, explique por que sua verificação sobre o controle estatístico difere para os dois tamanhos de  $n$ .

15.S10 Os dados seguintes, provenientes da página na internet, do Departamento Americano de Energia (<http://www.eia.doe.gov>), reportaram o consumo total anual de energia renovável (quadrilhões de BTU) dos Estados Unidos, de 1973 a 2015.

Ano	Consumo Total de Energia Renovável (Quadrilhão de BTU)	Ano	Consumo Total de Energia Renovável (Quadrilhão de BTU)
1973	4,41	1995	6,56
1974	4,74	1996	7,01
1975	4,69	1997	7,01
1976	4,73	1998	6,49
1977	4,21	1999	6,51
1978	5,01	2000	6,10
1979	5,12	2001	5,16
1980	5,43	2002	5,73
1981	5,41	2003	5,94
1982	5,98	2004	6,07
1983	6,50	2005	6,23
1984	6,44	2006	6,64
1985	6,08	2007	6,52
1986	6,11	2008	7,17
1987	5,62	2009	7,60
1988	5,46	2010	8,03

(continua)

(continuação)

Ano	Consumo Total de Energia Renovável (Quadrilhão de BTU)	Ano	Consumo Total de Energia Renovável (Quadrilhão de BTU)
1989	6,23	2011	9,00
1990	6,04	2012	8,71
1991	6,07	2013	9,27
1992	5,82	2014	9,56
1993	6,08	2015	9,58
1994	5,99		

a. Usando todos os dados, encontre os limites para um gráfico de controle de medidas individuais, construa o gráfico e plote os dados.

b. Os dados parecem ter sido gerados a partir de um processo sob controle? Comente qualquer padrão de comportamento do gráfico.

15.S11 Um artigo em *Quality Engineering* ["Is the Process Capable? Tables and Graphs in Assessing  $C_{pm}$ " (1992, Vol. 4(4), pp. 563-576)] considerou dados de fabricação. Especificações para o diâmetro externo dos orifícios foram de  $60,3265 \pm 0,001$  mm. Uma amostra aleatória com tamanho  $n = 20$  foi tomada, e os dados são mostrados na tabela seguinte.

Orifício	$x$	Orifício	$x$	Orifício	$x$
1	60,3262	8	60,3267	15	60,3265
2	60,3262	9	60,3263	16	60,3266
3	60,3262	10	60,3269	17	60,3265
4	60,3266	11	60,3262	18	60,3268
5	60,3263	12	60,3262	19	60,3262
6	60,3260	13	60,3269	20	60,3266
7	60,3262	14	60,3261		

a. Construa um gráfico para medidas individuais. Reveja os limites de controle, se necessário.

b. Compare seu gráfico do item (a) àquele que usa somente o último dígito (o menos significativo) de cada diâmetro como a medida. Explique sua conclusão.

c. Estime  $\mu$  e  $\sigma$  a partir da média móvel do gráfico revisado e use esse valor para estimar  $PCR$  e  $PCR_x$ ; interprete essas razões.

15.S12 Suponha que um gráfico de controle  $\bar{X}$ , com limites de 2-sigma, seja usado para controlar um processo. Encontre a probabilidade de que um sinal de falso alarme de processo fora de controle seja produzido na próxima amostra. Compare essa com a probabilidade correspondente para o gráfico com limites de 3-sigma e discuta. Comente quando você preferiria usar os limites de 2-sigma em vez dos limites de 3-sigma.

15.S13 O seguinte conjunto de dados foi considerado em *Quality Engineering* ["Analytic Examination of Variance Components" (1994-1995, Vol. 7(2), pp. 315-336)]. Uma característica de qualidade para briquetes de argamassa de cimento foi monitorada. Amostras de tamanho  $n = 6$  foram retiradas do processo, e 25 amostras do processo são apresentadas na tabela seguinte.

Batelada	$\bar{x}$	$s$	Batelada	$\bar{x}$	$s$
1	572,00	73,25	14	485,67	103,33
2	583,83	79,30	15	746,33	107,88
3	720,50	86,44	16	436,33	98,69
4	368,67	98,62	17	556,83	99,25
5	374,00	92,36	18	390,33	117,35
6	580,33	93,50	19	562,33	75,69
7	388,33	110,23	20	675,00	90,10
8	559,33	74,79	21	416,50	89,27
9	562,00	76,53	22	568,33	61,36
10	729,00	49,80	23	762,67	105,94
11	469,00	40,52	24	786,17	65,05
12	566,67	113,82	25	530,67	99,42
13	578,33	58,03			

a. Usando todos os dados, encontre os limites tentativas para  $\bar{X}$  e  $S$ . O processo está sob controle?

b. Suponha que as especificações sejam  $580 \pm 250$ . Que afirmações você pode fazer acerca da capacidade de processo? Calcule as estimativas das razões apropriadas da capacidade de processo.

c. Para tornar esse processo um "processo 6-sigma", a variância  $\sigma^2$  teria de ser diminuída, de modo que  $PCR_x = 2,0$ . Qual deve ser esse novo valor da variância?

d. Suponha que a média se desloque para 600. Qual é a probabilidade de que esse deslocamento seja detectado na próxima amostra? Qual é o CMC depois do deslocamento?

15.S14 Suponha que um processo esteja sob controle e um gráfico  $\bar{X}$  seja usado, com uma amostra de tamanho igual a quatro, para monitorar o processo. Há um deslocamento repentino de  $1,5\sigma$  na média.

a. Se os limites de controle 3-sigma estiverem em uso no gráfico  $\bar{X}$ , qual será a probabilidade de que esse deslocamento permaneça sem ser detectado por três amostras consecutivas?

b. Se os limites de controle 2-sigma estiverem em uso no gráfico  $\bar{X}$ , qual será a probabilidade de que esse deslocamento permaneça sem ser detectado por três amostras consecutivas?

c. Compare suas respostas dos itens (a) e (b) e explique por que elas diferem. Além disso, que limites você recomendaria usar e por quê?

15.S15 Considere os dados do diâmetro no Exercício 15.S7.

a. Construa um gráfico de controle MMPE, com  $\lambda = 0,2$  e  $L = 3$ . Comente a respeito do controle do processo.

b. Construa um gráfico de controle MMPE, com  $\lambda = 0,5$  e  $L = 3$ , e compare sua conclusão com o item (a).

15.S16 Considere o gráfico de controle para medidas individuais, aplicado a um processo químico contínuo (24 horas), com observações tomadas a cada hora.

a. Se o gráfico tiver limites 3-sigma, quantos alarmes falsos, em média, ocorreriam a cada mês de 30 dias com esse gráfico?

b. Suponha que o gráfico tenha limites 2-sigma. Isso reduz o CMC para detectar um deslocamento de magnitude  $\sigma$  na média?

c. Encontre o CMC sob controle, se os limites 2-sigma forem usados no gráfico. Quantos alarmes falsos ocorrerão a cada mês, com esse gráfico? O desempenho desse CMC sob controle é satisfatório? Explique sua resposta.

15.S17 Considere os dados do orifício no Exercício 15.S11.

- a. Construa um gráfico de controle MMPE, com  $\lambda = 0,2$  e  $L = 3$ . Comente a respeito do controle do processo.
- b. Construa um gráfico de controle MMPE, com  $\lambda = 0,5$  e  $L = 3$ , e compare sua conclusão com o item (a).

15.S18 Considere um gráfico de controle para medidas individuais, com limites 3-sigma. Qual é a probabilidade de que não haja um sinal em três amostras? Em seis amostras? Em dez amostras?

15.S19 Os dados seguintes foram considerados em *Quality Progress* ["Digidot Plots for Process Surveillance" (1990, Vol. 23(5), pp. 66-68)]. São mostradas, a seguir, medidas da espessura (em milésimo de polegada) do centro, provenientes de 25 lentes de contato incluídas na amostra de um processo de produção, em intervalos regulares.

Amostra	$x$	Amostra	$x$
1	0,3978	14	0,3999
2	0,4019	15	0,4062
3	0,4031	16	0,4048
4	0,4044	17	0,4071
5	0,3984	18	0,4015
6	0,3972	19	0,3991
7	0,3981	20	0,4021
8	0,3947	21	0,4009
9	0,4012	22	0,3988
10	0,4043	23	0,3994
11	0,4051	24	0,4016
12	0,4016	25	0,4010
13	0,3994		

Construa um gráfico de controle MMPE, com  $\mu_0 = 0,4$ ,  $\lambda = 0,5$  e  $L = 3$ , e compare suas conclusões com o gráfico de medidas individuais com limites 3-sigma.

15.S20 Um processo está controlado por um gráfico  $P$ , usando amostras de tamanho 100. A linha central no gráfico é 0,05.

- a. Qual é a probabilidade de que o gráfico de controle detecte um deslocamento para 0,08 na primeira amostra seguinte ao deslocamento?
- b. Qual é a probabilidade de o gráfico de controle não detectar um deslocamento para 0,08 na primeira amostra seguinte ao deslocamento, mas detectá-lo na segunda amostra?
- c. Suponha que, em vez de um deslocamento para 0,08, a média se desloque para 0,10. Repita os itens (a) e (b).
- d. Compare suas respostas entre um deslocamento para 0,08 e um deslocamento para 0,10. Explique por que elas diferem. Explique também por que um deslocamento para 0,10 é mais fácil de detectar.

15.S21 Considere o gráfico de controle para alternativa, com limites 3-sigma.

- a. Suponha que um deslocamento de magnitude  $\sigma$  ocorra na média do processo. Verifique se o CMC para detectar o deslocamento é  $CMC = 43,9$ .
- b. Encontre o CMC para detectar um deslocamento de magnitude  $2\sigma$  na média do processo.
- c. Encontre o CMC para detectar um deslocamento de magnitude  $3\sigma$  na média do processo.
- d. Compare suas respostas nos itens (a), (b) e (c) e explique por que o CMC para detecção está diminuindo à medida que a magnitude de deslocamento aumenta.

15.S22 Considere um gráfico de controle  $\bar{X}$  com  $LSC = 32,802$ ,  $LIC = 24,642$  e  $n = 5$ . Suponha que a média se desloque para 30.

- a. Qual é a probabilidade de que esse deslocamento seja detectado na próxima amostra?
- b. Qual é o CMC para detectar o deslocamento?

15.S23 A profundidade de uma fechadura é uma característica importante da qualidade da peça. Amostras de tamanho  $n = 5$  são retiradas do processo a cada quatro horas, sendo 20 amostras resumidas na tabela seguinte.

Amostra	$\bar{x}$	$r$	Amostra	$\bar{x}$	$r$
1	139,7	1,1	11	138,4	0,8
2	139,8	1,4	12	138,5	0,9
3	140,0	1,3	13	137,9	1,2
4	140,1	1,6	14	138,5	1,1
5	139,8	0,9	15	140,8	1,0
6	139,9	1,0	16	140,5	1,3
7	139,7	1,4	17	139,4	1,4
8	140,2	1,2	18	139,9	1,0
9	139,3	1,1	19	137,5	1,5
10	140,7	1,0	20	139,2	1,3

- a. Usando todos os dados, encontre os limites de controle iniciais para os gráficos  $\bar{X}$  e  $R$ . O processo está sob controle?
- b. Use os limites de controle iniciais do item (a) para identificar pontos fora de controle. Se necessário, reveja seus limites de controle e estime então o desvio-padrão do processo.
- c. Suponha que as especificações sejam  $140 \pm 2$ . Usando os resultados do item (b), que afirmações você pode fazer acerca da capacidade do processo? Calcule as estimativas das razões apropriadas da capacidade de processo.
- d. De modo a tornar esse um "processo 6-sigma", a variância  $\sigma^2$  teria de ser diminuída para que  $PCR_k = 2,0$ . Qual deveria ser o novo valor dessa variância?
- e. Suponha que a média se desloque para 139,7. Qual é a probabilidade de que esse deslocamento seja detectado na próxima amostra? Qual é o CMC depois do deslocamento?

15.S24 A tabela seguinte apresenta o número de defeitos observados nas 15 amostras das unidades de transmissão em uma fábrica de automóveis. Cada lote contém cinco unidades de transmissão.

Amostra	Nº de Defeitos	Amostra	Nº de Defeitos
1	8	11	6
2	10	12	10
3	24	13	11
4	6	14	17
5	5	15	9
6	21		
7	10		
8	7		
9	9		
10	15		

- a. Usando todos os dados, calcule os limites de controle iniciais para um gráfico de controle de  $U$ , construa o gráfico e plote os dados.
- b. Determine se o processo está sob controle estatístico. Se não, considere que as causas especiais podem ser encontradas e os pontos fora de controle podem ser eliminados. Reveja os limites de controle.

15.S25 Suponha que o número médio de defeitos em uma unidade seja 8. Se o número médio de defeitos em uma unidade se deslocar para 16, qual é a probabilidade de que ele seja detectado pelo gráfico  $U$  na primeira amostra seguinte ao deslocamento,

- Se o tamanho da amostra for  $n = 4$ ?
- Se o tamanho da amostra for  $n = 10$ ?

Use a aproximação normal para  $U$ .

15.S26 O número de visitas (em milhões) em um *site* da internet é registrado a cada dia. A seguinte tabela apresenta as amostras para 25 dias consecutivos.

Amostra	Número de Visitas	Amostra	Número de Visitas
1	10,12	16	9,66
2	9,92	17	10,42
3	9,76	18	11,30
4	9,35	19	12,53
5	9,60	20	10,76
6	8,60	21	11,92
7	10,46	22	13,24
8	10,58	23	10,64
9	9,95	24	11,31
10	9,50	25	11,26
11	11,26	26	11,79
12	10,02	27	10,53
13	10,95	28	11,82
14	8,99	29	11,47
15	9,50	30	11,76

- Calcule a estimação-padrão do processo.
- Estabeleça um gráfico de controle MMPE para esse processo, considerando o alvo como 10, com  $\lambda = 0,4$ . O processo parece estar sob controle?

15.S27 A tabela seguinte mostra o número de *e-mails* que um estudante recebeu a cada hora, das 8 h às 18 h. As amostras foram coletadas durante cinco dias, de segunda a sexta-feira.

Hora	Segunda	Terça	Quarta	Quinta	Sexta
1	2	2	2	3	1
2	2	4	0	1	2
3	2	2	2	1	2
4	4	4	3	3	2
5	1	1	2	2	1
6	1	3	2	2	1
7	3	2	1	1	0
8	2	3	2	3	1
9	1	3	3	2	0
10	2	3	2	3	0

- Use o princípio de subgrupo racional para comentar por que um gráfico de  $\bar{X}$  que plota um ponto a cada hora com um subgrupo de tamanho 5 não é apropriado.
- Construa um gráfico de controle apropriado de atributo. Use todos os dados para encontrar os limites de controle iniciais, construa o gráfico e plote os dados.

c. Use os limites de controle iniciais a partir do item (b) para identificar os pontos fora de controle. Se necessário, reveja os seus limites de controle, considerando que qualquer amostra que fique fora dos limites de controle pode ser eliminada.

15.S28 Um artigo publicado no *Journal of Quality in Clinical Practice* ["The Application of Statistical Process Control Charts to the Detection and Monitoring of Hospital-Acquired Infections" (2001, Vol. 21, pp. 112-117)] reportou o uso de métodos de CEP para monitorar infecções adquiridas em hospitais. Os autores aplicaram os gráficos de Shewhart, de MMPE e outros gráficos para monitorar as infecções de pneumonia *Klebsiella* causadas por ESB (Extended Spectrum *b Lactamase*). O número mensal de infecções de junho de 1994 a maio de 1998 é mostrado na tabela.

	Jan	Feb	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1994						5	0	0	2	2	3	1
1995	1	3	2	6	4	1	2	4	3	2	8	7
1996	10	6	10	11	5	6	3	0	3	3	1	3
1997	0	2	0	4	1	1	4	2	6	7	1	5
1998	3	0	1	0	2							

- Que distribuição deve ser esperada para esses dados? Que tipo de gráfico de controle pode ser apropriado?
- Construa o gráfico que você selecionou no item (a).
- Construa um gráfico MMPE para esses dados com  $\lambda = 0,2$ . O artigo incluiu um gráfico construído de forma similar. O que foi admitido para a distribuição dos dados nesse gráfico? O seu gráfico MMPE pode ter desempenho adequado?

15.S29 Um artigo publicado em *Microelectronics Reliability* ["Advanced Electronic Prognostics Through System Telemetry and Pattern Recognition Methods" (2007, Vol. 47(12), pp. 1865-1873)] apresentou um exemplo de prognóstico eletrônico (uma técnica para detectar falhas de modo a diminuir a indisponibilidade do sistema e o número de reparos não planejados em sistemas de alta confiabilidade e alta disponibilidade). Sinais de voltagem de servidores das empresas foram monitorados ao longo do tempo. As medidas são fornecidas na seguinte tabela.

Observação	Sinal da Voltagem	Observação	Sinal da Voltagem
1	1,498	26	1,510
2	1,494	27	1,521
3	1,500	28	1,507
4	1,495	29	1,493
5	1,502	30	1,499
6	1,509	31	1,509
7	1,480	32	1,491
8	1,490	33	1,478
9	1,486	34	1,495
10	1,510	35	1,482
11	1,495	36	1,488
12	1,481	37	1,480
13	1,529	38	1,519
14	1,479	39	1,486
15	1,483	40	1,517
16	1,505	41	1,517

(continua)

(continuação)

Observação	Sinal da Voltagem	Observação	Sinal da Voltagem
17	1,536	42	1,490
18	1,493	43	1,495
19	1,496	44	1,545
20	1,587	45	1,501
21	1,610	46	1,503
22	1,592	47	1,486
23	1,585	48	1,473
24	1,587	49	1,502
25	1,482	50	1,497

a. Usando todos os dados, calcule os limites de controle iniciais para as observações individuais e os gráficos da amplitude móvel. Construa o gráfico e plote os dados. Determine se o processo está sob controle estatístico. Se não, considere que as causas especiais podem ser encontradas para eliminar essas amostras e reveja os limites de controle.

b. Estime a média e o desvio-padrão do processo para o processo sob controle.

c. O artigo considerou que o sinal é normalmente distribuído com uma média de 1,5 V e um desvio-padrão de 0,02 V. Os seus resultados no item (b) corroboram essa suposição?

**15.S30** Considere o tempo de resposta (TR) para hemogramas completos no Exercício 15.3.12. Suponha que as especificações para TR sejam fixadas em 20 e 80 minutos. Use as estatísticas do resumo do gráfico de controle para calcular o seguinte:

- Estime o desvio-padrão do processo.
- Calcule o  $PCR$  e o  $PCR_x$  para o processo.

**15.S31** Um artigo publicado em *Electric Power Systems Research* ["On the Self-Scheduling of a Power Producer in Uncertain Trading Environments" (2008, Vol. 78(3), pp. 311-317)] considerou uma abordagem de autoagendamento para um fornecedor de energia. A tabela seguinte mostra os preços previstos de energia para um período de 24 horas, de acordo com um cenário de base.

Hora	Preço	Hora	Preço	Hora	Preço
1	38,77	9	48,75	17	52,07
2	37,52	10	51,18	18	51,34
3	37,07	11	51,79	19	52,55
4	35,82	12	55,22	20	53,11
5	35,04	13	53,48	21	50,88
6	35,57	14	51,34	22	52,78
7	36,23	15	45,8	23	42,16
8	38,93	16	48,14	24	42,16

a. Construa gráficos individuais e da amplitude móvel. Determine se os preços de energia flutuam no controle estatístico.

b. A suposição de observações independentes é razoável para esses dados?

**15.S32** Considere os dados de TR do Exercício 15.3.12.

a. Construa um gráfico de controle MMPE com um alvo igual à média estimada do processo,  $k = 0,5$  e  $h = 4$ . O processo parece estar sob controle no alvo?

b. Se a média aumentar por cinco minutos, aproxime o CMC do gráfico.

**15.S33** Considere os dados de resíduos infectados do Exercício 15.4.6. Use os dados somente depois da mudança no processo.

a. Construa um gráfico de controle MMPE com um alvo igual à média estimada do processo,  $k = 0,5$  e  $h = 4$ . O processo parece estar sob controle no alvo?

b. Se a média aumentar por 1,0 lb, aproxime o CMC do gráfico.