16 Previsão de Séries Temporais

Objetivos do Aprendizado

Neste capítulo, você aprenderá:

- Sobre diferentes modelos de previsão de séries temporais: médias móveis, ajuste exponencial, a tendência linear, a tendência quadrática, a tendência exponencial, bem como o modelo autorregressivo e o modelo dos mínimos quadrados para dados sazonais
- A escolher o modelo mais apropriado para previsão de séries temporais

UTILIZANDO A ESTATÍSTICA

@ The Principled

Você é analista financeiro da The Principled, uma empresa de grande porte que presta serviços financeiros. Você precisa fazer previsões de receitas para três empresas, com o objetivo de melhor avaliar as oportunidades de investimentos para seus clientes. Para auxiliar na previsão, você coletou dados de séries temporais relativos a três empresas: Cabot Corporation, The Coca-Cola Company e Wal-Mart Stores, Inc. Cada uma das séries temporais apresenta características exclusivas, devido a diferentes tipos de atividades de negócios e padrões de crescimento vivenciados por essas três empresas. Você entende que pode utilizar vários tipos diferentes de modelos de previsão. De que modo você decide qual tipo de modelo de previsão é o melhor para cada uma das empresas? De que modo você utiliza as informações obtidas dos modelos de previsão para avaliar as oportunidades de investimentos para seus clientes?

Nos Capítulos 13 a 15, você utilizou a análise da regressão como uma ferramenta para a construção de modelos e para fins de previsão. Neste capítulo, a análise de regressão e outras metodologias estatísticas são aplicadas em dados de séries temporais. Uma **série temporal** é um conjunto de dados numéricos coletados ao longo do tempo. Devido a diferenças nas características dos dados correspondentes a várias empresas, como é o caso das três empresas descritas no cenário Utilizando a Estatística, você precisa considerar várias abordagens diferentes para prever dados de séries temporais.

Este capítulo começa com uma introdução à importância da previsão no mundo dos negócios (veja a Seção 16.1) e uma descrição dos componentes de modelos de séries temporais (veja a Seção 16.2). A cobertura de modelos de previsão tem início com dados de séries temporais anuais. A Seção 16.3 apresenta os métodos de médias móveis e ajuste exponencial para o ajuste de séries. A análise de séries temporais anuais prossegue com o uso de previsões e ajuste de tendência dos mínimos quadrados na Seção 16.4, e a modelagem autorregressiva na Seção 16.5. A Seção 16.6 discorre sobre o modo de escolher entre métodos alternativos para realizar previsões. A Seção 16.7 desenvolve modelos para séries temporais mensais e trimestrais.

16.1 A Importância da Previsão nos Negócios

Previsões são feitas com base no monitoramento de alterações que ocorrem ao longo do tempo e a projeção dessas alterações no futuro. A previsão é habitualmente utilizada tanto para setores com fins lucrativos quanto para setores sem fins lucrativos da economia. Por exemplo, executivos de marketing de uma grande empresa de varejo fazem previsões sobre demanda de produtos, receitas de vendas, preferências do consumidor, estoques e assim por diante, para que possam tomar decisões com relação a promoções e ao planejamento estratégico. Executivos da esfera governamental fazem previsões sobre desemprego, inflação, produção industrial e receitas provenientes do imposto de renda para que possam formular políticas. E, ainda, os administradores de uma faculdade ou universidade realizam previsões sobre o número de alunos matriculados para planejar a construção de dormitórios e outras instalações acadêmicas, planejar a contratação de pessoal e fazer avaliações de outras necessidades.

Existem dois métodos comuns para previsão: *qualitativo* e *quantitativo*. **Métodos qualitativos de previsão** são especialmente importantes quando os dados históricos não estão disponíveis. Métodos qualitativos de previsão são considerados altamente subjetivos e arbitrários.

Métodos quantitativos de previsão fazem uso de dados históricos. O objetivo desses métodos é utilizar dados do passado para prever valores futuros. Os métodos quantitativos de previsão são subdivididos em dois tipos: *séries temporais* e *causais*. **Métodos de previsão de séries temporais** envolvem a projeção de valores futuros com base inteiramente em valores do passado e do presente de uma determinada variável. Por exemplo, os preços diários de fechamento de uma determinada ação na Bolsa de Nova York constituem uma série temporal. Outros exemplos de séries temporais no âmbito dos negócios ou na economia são as publicações mensais do Índice de Preços ao Consumidor (IPC), as informações trimestrais sobre o Produto Interno Bruto (PIB) e as receitas anuais das vendas de uma determinada empresa.

Métodos causais de previsão envolvem a determinação de fatores que se relacionam à variável que você está tentando prever. Neles se incluem a análise de regressão múltipla, com variáveis observadas no passado, a modelagem econométrica, análises dos principais indicadores, índices de difusão e outros barômetros econômicos que estão além do escopo deste livro (veja as referências 2-4). A principal ênfase deste capítulo está nos métodos de previsão de séries temporais.

16.2 Fatores Componentes dos Modelos de Séries Temporais

A previsão com base em séries temporais pressupõe que os fatores que influenciaram atividades no passado e no presente continuarão a fazê-lo, aproximadamente do mesmo modo, no futuro. A previsão com base em séries temporais busca identificar e isolar esses fatores componentes no intuito de realizar previsões (prognósticos). Tipicamente, os quatro fatores a seguir são examinados em modelos de séries temporais:

- Tendência
- Efeito cíclico
- Efeito irregular ou aleatório
- Efeito sazonal

Uma **tendência** é um movimento geral, ascendente ou descendente, de longo prazo, em uma determinada série temporal. A tendência não é o único fator componente que pode influenciar dados em uma série temporal. O **efeito cíclico** demonstra oscilações ou movimentos ascendentes ou descendentes ao longo de toda a série. Movimentos cíclicos variam em termos de extensão, geralmente durando de 2 a 10 anos. Eles diferem em termos de intensidade e estão frequentemente correlacionados a um ciclo econômico. Em alguns períodos de tempo, os valores se mostram mais elevados do que seria previsto por uma linha de tendência (ou seja, eles estão no pico de um ciclo ou próximo dele). Em outros períodos de tempo, os valores estão mais baixos do que seria previsto por uma linha de tendência (ou seja, eles estarão na parte mais baixa de um ciclo, ou próximo a ela). Quaisquer dados que não sigam a tendência modificada pelo componente cíclico são considerados parte do **efeito aleatório** ou **irregular**. Quando você tem em mãos dados mensais ou dados trimestrais, um outro componente, conhecido como **efeito sazonal**, é considerado, juntamente com a tendência, o efeito cíclico e o efeito irregular.

Seu primeiro passo em uma análise de séries temporais é desenhar o gráfico para os dados e observar quaisquer padrões que ocorram ao longo do tempo. Você deve determinar se existe um movimento de longo prazo, ascendente ou descendente, na série (ou seja, uma tendência). Caso não exista nenhuma tendência evidente no longo prazo, de caráter ascendente ou descendente, você pode então utilizar médias móveis ou o ajuste exponencial para ajustar a série e proporcionar uma impressão geral de longo prazo (veja a Seção 16.3). Caso esteja presente uma tendência, você pode considerar vários métodos de previsão de séries temporais (veja as Seções 16.4-16.5 para prognósticos sobre dados anuais e a Seção 16.7 para prognósticos sobre séries temporais mensais ou trimestrais).

16.3 Ajustando uma Série Temporal Anual

Uma das empresas de interesse no cenário que trata da empresa The Principled é a Cabot Corporation. Com matriz em Boston, a Cabot Corporation é uma empresa globalizada, especializada na fabricação e distribuição de material químico. A empresa opera 39 unidades de produção e 8 unidades de pesquisa e desenvolvimento em 19 países, e tem o seu capital em ações negociado na Bolsa de Nova York, com a sigla CBT. As receitas no ano fiscal de 2008 foram de aproximadamente $ 3,2 bilhões. A Tabela 16.1 apresenta o total de receitas, em milhões de dólares, para o período de 1982 a 2008 (dados no arquivo **Cabot**). A Figura 16.1 apresenta o gráfico de séries temporais.

TABELA 16.1

Receitas (em Milhões de Dólares) para a Cabot Corporation, de 1982 a 2008

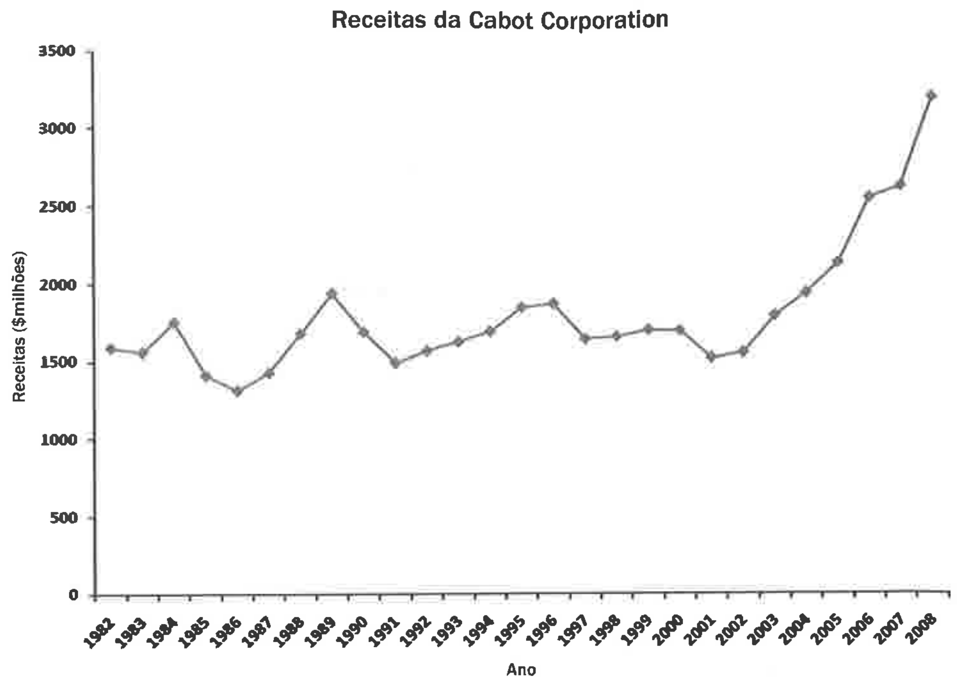
Ano	Receita	Ano	Receita	Ano	Receita
1982	1.588	1991	1.488	2000	1.698
1983	1.558	1992	1.562	2001	1.523
1984	1.753	1993	1.619	2002	1.557
1985	1.408	1994	1.687	2003	1.795
1986	1.310	1995	1.841	2004	1.934
1987	1.424	1996	1.865	2005	2.125
1988	1.677	1997	1.637	2006	2.543
1989	1.937	1998	1.653	2007	2.616
1990	1.685	1999	1.699	2008	3.191

Fonte: *Dados extraídos de* Moody's Handbook of Common Stocks, *1992*; Mergent's Handbook of Common Stocks, *2008; e www.cabot-corp.com*.

Quando você examina os dados anuais, a sua impressão visual sobre a tendência de longo prazo na série é ofuscada pelo montante de variação de ano para ano. De modo geral, você não consegue avaliar se existe na série qualquer tendência de longo prazo, ascendente ou descendente. Para obter melhor impressão geral sobre o padrão de movimento nos dados ao longo do tempo, você pode utilizar o método das *médias móveis* ou o método do *ajuste exponencial*.

FIGURA 16.1
Gráfico para as receitas brutas correntes (em milhões de dólares) para a Cabot Corporation (1982-2008)

Crie essa planilha de gráfico utilizando as instruções na Seção GE2.7.



Médias Móveis

Calcule médias móveis utilizando as instruções na Seção GE16.3.

Médias móveis para um período escolhido de extensão L consistem em uma série de médias aritméticas, cada uma delas calculada ao longo do tempo, para uma sequência de L valores observados. Médias móveis, representadas pelo símbolo $MM(L)$, podem ser significativamente afetadas pelo valor escolhido para L, que deve ser um número inteiro que corresponda ou seja um múltiplo da duração média estimada para um ciclo na série temporal.

Para fins de ilustração, suponha que você deseje calcular médias móveis para 5 anos, a partir de uma série que contenha $n = 11$ anos. Uma vez que $L = 5$, as médias móveis para 5 anos consistem em uma série de médias aritméticas calculadas pelo cálculo da média para sequências consecutivas de 5 valores. Você calcula a primeira média móvel para 5 anos fazendo a soma dos valores correspondentes aos 5 primeiros anos na série e dividindo esse valor por 5:

$$MM(5) = \frac{Y_1 + Y_2 + Y_3 + Y_4 + Y_5}{5}$$

Você calcula a segunda média móvel para 5 anos somando os valores correspondentes aos anos 2 a 6 dentro da série e, em seguida, dividindo esse valor por 5:

$$MM(5) = \frac{Y_2 + Y_3 + Y_4 + Y_5 + Y_6}{5}$$

Você dá continuidade a esse processo até ter calculado a última dessas médias móveis para 5 anos pela soma dos valores correspondentes aos últimos 5 anos na série (ou seja, os anos de 7 a 11) e, depois, dividindo esse resultado por 5:

$$MM(5) = \frac{Y_7 + Y_8 + Y_9 + Y_{10} + Y_{11}}{5}$$

Quando você está lidando com dados de séries temporais anuais, L deve corresponder a um número *ímpar* de anos. Ao seguir essa regra, você não consegue calcular nenhuma média móvel para os primeiros $(L - 1)/2$ anos ou para os últimos $(L - 1)/2$ anos da série. Assim, para uma média móvel de 5 anos, você não consegue realizar cálculos para os 2 primeiros anos ou para os 2 últimos anos da série.

Ao elaborar um gráfico para médias móveis, você insere no gráfico cada um dos valores calculados em relação ao ano do meio para a sequência de anos utilizada para esse cálculo. Se $n = 11$ e $L = 5$, a primeira média móvel é centralizada no terceiro ano; a segunda média móvel é centralizada no quarto ano; e a última média móvel é centralizada no nono ano. O Exemplo 16.1 ilustra o cálculo das médias móveis para 5 anos.

EXEMPLO 16.1

Calculando Médias Móveis para 5 Anos

Os dados a seguir representam o total das receitas (em milhões de dólares) para uma lanchonete, ao longo de um período de 11 anos, de 1999 a 2009:

4,0 5,0 7,0 6,0 8,0 9,0 5,0 2,0 3,5 5,5 6,5

Calcule as médias móveis de 5 anos para essa série temporal anual.

SOLUÇÃO Para fazer o cálculo de médias móveis para 5 anos, você calcula, inicialmente, o total da movimentação para 5 anos e, em seguida, divide esse total por 5. A primeira das médias móveis para 5 anos é

$$MM(5) = \frac{Y_1 + Y_2 + Y_3 + Y_4 + Y_5}{5} = \frac{4,0 + 5,0 + 7,0 + 6,0 + 8,0}{5} = \frac{30,0}{5} = 6,0$$

A média móvel é então centralizada no valor do meio – o terceiro ano dessa série temporal. Para calcular a segunda das médias móveis para 5 anos, você calcula o total da movimentação, partindo do segundo até o sexto ano, e divide esse valor por 5:

$$MM(5) = \frac{Y_2 + Y_3 + Y_4 + Y_5 + Y_6}{5} = \frac{5,0 + 7,0 + 6,0 + 8,0 + 9,0}{5} = \frac{35,0}{5} = 7,0$$

Essa média móvel é centralizada no novo valor do meio – o quarto ano da série temporal. As médias móveis remanescentes são

$$MM(5) = \frac{Y_3 + Y_4 + Y_5 + Y_6 + Y_7}{5} = \frac{7,0 + 6,0 + 8,0 + 9,0 + 5,0}{5} = \frac{35,0}{5} = 7,0$$

$$MM(5) = \frac{Y_4 + Y_5 + Y_6 + Y_7 + Y_8}{5} = \frac{6,0 + 8,0 + 9,0 + 5,0 + 2,0}{5} = \frac{30,0}{5} = 6,0$$

$$MM(5) = \frac{Y_5 + Y_6 + Y_7 + Y_8 + Y_9}{5} = \frac{8,0 + 9,0 + 5,0 + 2,0 + 3,5}{5} = \frac{27,5}{5} = 5,5$$

$$MM(5) = \frac{Y_6 + Y_7 + Y_8 + Y_9 + Y_{10}}{5} = \frac{9,0 + 5,0 + 2,0 + 3,5 + 5,5}{5} = \frac{25,0}{5} = 5,0$$

$$MM(5) = \frac{Y_7 + Y_8 + Y_9 + Y_{10} + Y_{11}}{5} = \frac{5,0 + 2,0 + 3,5 + 5,5 + 6,5}{5} = \frac{22,5}{5} = 4,5$$

Essas médias móveis são então centralizadas em seus respectivos valores do meio – o quinto, o sexto, o sétimo, o oitavo e o nono ano na série temporal. Pelo fato de estarem sendo utilizadas médias móveis para 5 anos, você não consegue calcular uma média móvel para os 2 primeiros ou para os 2 últimos valores na série temporal.

Na prática, você deve evitar cálculos demasiadamente elaborados utilizando fórmulas de planilhas que utilizem a função **MÉDIA** para calcular médias móveis. A Figura 16.2 apresenta os dados correspondentes às receitas anuais da Cabot Corporation relativos ao período de 27 anos, de 1982 a 2008, os cálculos para as médias móveis de 3 e de 7 anos e um gráfico para os dados originais e as médias móveis.

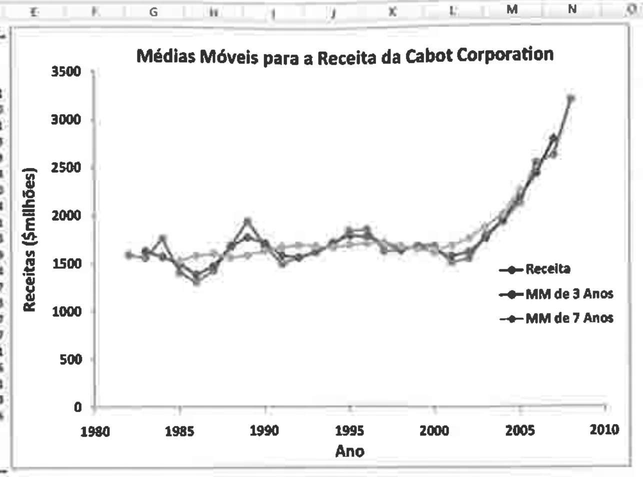
Na Figura 16.2, não existe uma média móvel de 3 anos para o primeiro e o último ano, e não existe uma média móvel de 7 anos para os 3 primeiros anos e para os 3 últimos anos. As médias móveis para 7 anos ajustam bem melhor a série do que as médias móveis para 3 anos, uma vez que o período é mais longo. Infelizmente, quanto mais extenso o período, menor o número de médias móveis que você consegue calcular. Portanto, a seleção de médias móveis com períodos superiores a 7 anos geralmente não é recomendável, já que fica faltando uma quantidade demasiadamente grande de valores para médias móveis no início e no final da série. Isso faz com que seja mais difícil obter uma impressão geral da série como um todo.

A seleção de L, a extensão do período utilizado para a construção das médias, é extremamente subjetiva. Caso oscilações cíclicas estejam presentes nos dados, escolha um valor inteiro para L que corresponda (ou que seja um múltiplo em relação) à extensão estimada para um ciclo na série. Para dados de séries temporais anuais que não apresentem oscilações cíclicas evidentes, a maioria das pessoas escolhe 3 anos, 5 anos ou 7 anos como valor para L, dependendo do montante de ajuste desejado e do volume de dados disponível.

FIGURA 16.2
Planilha com o gráfico sobreposto para as médias móveis de 3 e 7 anos relativas às receitas da Cabot Corporation

Crie essa planilha com o gráfico superposto utilizando as instruções na Seção GE16.3.

	A	B	C	D
1	Ano	Receitas	MM de 3 Anos	MM de 7 Anos
2	1982	1588	#N/D	#N/D
3	1983	1558	1633,0	#N/D
4	1984	1753	1573,0	#N/D
5	1985	1408	1490,3	1531,1
6	1986	1310	1380,7	1581,0
7	1987	1424	1470,3	1599,1
8	1988	1677	1679,3	1561,3
9	1989	1937	1766,3	1583,3
10	1990	1685	1703,3	1627,4
11	1991	1488	1578,3	1665,0
12	1992	1562	1556,3	1688,4
13	1993	1619	1622,7	1678,1
14	1994	1687	1715,7	1671,3
15	1995	1841	1797,7	1694,9
16	1996	1865	1781,0	1714,4
17	1997	1637	1718,3	1725,7
18	1998	1653	1663,0	1702,3
19	1999	1699	1683,3	1661,7
20	2000	1698	1640,0	1651,7
21	2001	1523	1592,7	1694,1
22	2002	1557	1625,0	1761,6
23	2003	1795	1762,0	1882,1
24	2004	1934	1951,3	2013,3
25	2005	2125	2200,7	2251,6
26	2006	2543	2428,0	#N/D
27	2007	2616	2783,3	#N/D
28	2008	3191	#N/D	#N/D



Ajuste Exponencial

Calcule valores exponencialmente ajustados utilizando as instruções da Seção GE16.3.

Ajuste exponencial consiste em uma série de médias móveis *exponencialmente ponderadas*. Os pesos atribuídos aos valores se modificam de modo tal que o valor mais recente recebe o maior peso, o valor anterior recebe o segundo maior peso e assim sucessivamente, com o primeiro valor recebendo o menor peso. Ao longo de toda a série, cada um dos valores exponencialmente ajustado depende de todos os valores anteriores, o que representa uma outra vantagem do ajuste exponencial em relação ao método das médias móveis. O ajuste exponencial também permite que você calcule prognósticos de curto prazo (um período à frente no futuro) quando a presença e o tipo de tendência de longo prazo em uma determinada série temporal são difíceis de se determinar.

A equação desenvolvida para ajustar exponencialmente uma série em qualquer período de tempo, i, é baseada em somente três termos – o valor corrente na série temporal, Y_i; o valor exponencialmente ajustado calculado anteriormente, E_{i-1}; e um peso atribuído, ou coeficiente de ajuste, W. Você utiliza a Equação (16.1) para fazer o ajuste exponencial de uma determinada série temporal.

CALCULANDO UM VALOR EXPONENCIALMENTE AJUSTADO NO PERÍODO DE TEMPO i

$$E_1 = Y_1 \qquad \textbf{(16.1)}$$

$$E_i = WY_i + (1 - W)E_{i-1} \quad i = 2, 3, 4, \ldots$$

em que

E_i = valor da série exponencialmente ajustada, calculado no período de tempo i

E_{i-1} = valor da série exponencialmente ajustada, já calculado no período de tempo $i - 1$

Y_i = valor observado da série temporal no período i

W = peso atribuído subjetivamente ou coeficiente de ajuste (em que $0 < W > 1$). Embora W possa estar próximo de 1,0, praticamente em todas as aplicações de natureza econômica, $W \leq 0{,}5$.

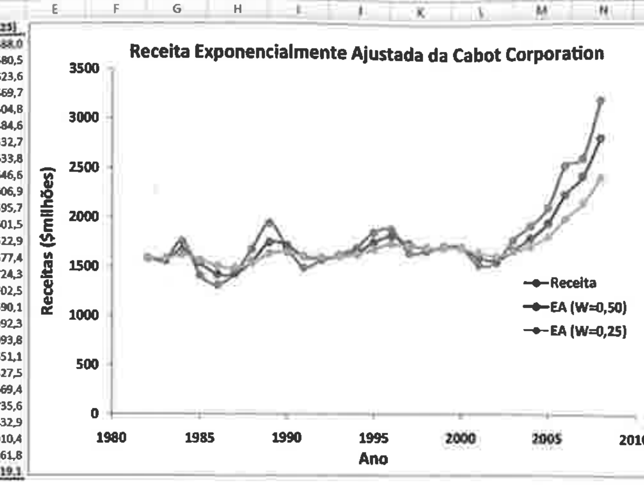
A escolha do peso ou coeficiente de ajuste (ou seja, W) que você atribui à série temporal é fator crítico. Infelizmente, essa seleção é relativamente subjetiva. Caso seu objetivo seja somente ajustar uma série, por meio da eliminação de variações cíclicas e irregulares indesejáveis, você deve selecionar um valor baixo para W (próximo de 0). Caso seu objetivo seja um prognóstico, você deve escolher um valor alto para W (próximo de 0,5). No primeiro caso, as tendências gerais de longo prazo para a série serão mais aparentes; no último caso, direcionamentos futuros de curto prazo podem ser previstos de maneira mais adequada.

A Figura 16.3 ilustra uma planilha que apresenta valores exponencialmente ajustados, obtidos com a utilização do Microsoft Excel (com os coeficientes de ajuste $W = 0{,}50$ e $W = 0{,}25$), para as receitas anuais da Cabot Corporation, ao longo do período de 27 anos, de 1982 a 2008, juntamente com um gráfico para os dados originais e para as duas séries temporais exponencialmente ajustadas.

FIGURA 16.3
Planilha com o gráfico sobreposto para a série exponencialmente ajustada ($W = 0{,}50$ e $W = 0{,}25$), das receitas da Cabot Corporation

Crie essa planilha com o gráfico superposto utilizando as instruções na Seção GE16.3.

	A	B	C	D
1	Ano	Receitas	EA(W=0,50)	EA(W=0,25)
2	1982	1588	1588,0	1588,0
3	1983	1558	1573,0	1580,5
4	1984	1753	1663,0	1623,6
5	1985	1408	1535,5	1569,7
6	1986	1310	1422,8	1504,8
7	1987	1424	1423,4	1484,6
8	1988	1677	1550,2	1532,7
9	1989	1937	1743,6	1633,8
10	1990	1685	1714,3	1646,6
11	1991	1488	1601,1	1606,9
12	1992	1562	1581,6	1595,7
13	1993	1619	1600,3	1601,5
14	1994	1687	1643,6	1622,9
15	1995	1841	1742,3	1677,4
16	1996	1865	1803,7	1724,3
17	1997	1637	1720,3	1702,5
18	1998	1653	1686,7	1690,1
19	1999	1699	1692,8	1692,3
20	2000	1698	1695,4	1693,8
21	2001	1523	1609,2	1651,1
22	2002	1557	1583,1	1627,5
23	2003	1795	1689,1	1669,4
24	2004	1934	1811,5	1735,6
25	2005	2125	1968,3	1832,9
26	2006	2543	2255,6	2010,4
27	2007	2616	2435,8	2161,8
28	2008	3191	2813,4	2419,1



Para ilustrar esses cálculos para o ajuste exponencial com um coeficiente de ajuste de $W = 0{,}25$, você começa lançando mão do valor inicial $Y_{1982} = 1.588$ como primeiro valor ajustado ($E_{1982} = 1.588$). Em seguida, utilizando o valor da série temporal correspondente a 1983 ($Y_{1983} = 1.588$), você ajusta a série para 1983, calculando:

$$E_{1983} = WY_{1983} + (1 - W)E_{1982}$$
$$= (0{,}25)(1.558) + (0{,}75)(1.588) = 1.580{,}5$$

Para ajustar a série para 1984:

$$E_{1984} = WY_{1984} + (1 - W)E_{1983}$$
$$= (0{,}25)(1.753) + (0{,}75)(1.580{,}5) = 1.623{,}6$$

Para ajustar a série para 1985:

$$E_{1985} = WY_{1985} + (1 - W)E_{1984}$$
$$= (0{,}25)(1.408) + (0{,}75)(1.623{,}6) = 1.569{,}7$$

Você dá continuidade ao processo até ter calculado os valores exponencialmente ajustados para todos os 27 anos na série, conforme ilustrado na Figura 16.4.

Para utilizar o ajuste exponencial para fins de previsão, você adota o valor ajustado no período de tempo corrente como uma projeção para o valor no período de tempo seguinte ($\hat{Y}_{i+1}$).

FAZENDO A PREVISÃO PARA O PERÍODO DE TEMPO $i + 1$

$$\hat{Y}_{i+1} = E_i \qquad \textbf{(16.2)}$$

Para prever a receita da Cabot Corporation durante 2009, utilizando um coeficiente de ajuste de $W = 0{,}25$, você utiliza o valor ajustado para 2008 como estimativa. A Figura 16.3 mostra que esse valor corresponde a $2.419,1 milhões. (Quão próxima da realidade está essa previsão? Verifique as receitas da Cabot Corporation no endereço **www.cabot-corp.com** para descobrir.) Quando estiver disponível o valor para 2009, você pode utilizar a Equação (16.1) para realizar um prognóstico para 2010, por meio do cálculo do valor ajustado para 2009, da seguinte maneira:

$$\text{Valor corrente ajustado} = (W)(\text{Valor corrente}) + (1 - W)(\text{Valor ajustado anterior})$$

$$E_{2009} = WY_{2009} + (1 - W)E_{2008}$$

Ou, em termos de previsão, você calcula o seguinte:

$$\text{Nova previsão} = (W)(\text{Valor corrente}) + (1 - W)(\text{Previsão corrente})$$

$$\hat{Y}_{2010} = WY_{2009} + (1 - W)\hat{Y}_{2009}$$

Problemas para a Seção 16.3

APRENDENDO O BÁSICO

16.1 Se você está utilizando o ajuste exponencial para prever uma série temporal anual para receitas, qual é a sua previsão para o ano subsequente se o valor ajustado para o ano corrente for $32,4 milhões?

16.2 Considere uma média móvel de 9 anos utilizada para ajustar uma série temporal que tenha sido primeiramente registrada em 1955.

a. Que ano serve como o primeiro valor centralizado na série ajustada?

b. Quantos anos correspondentes a valores observados nas séries são perdidos ao se calcular todas as médias móveis para 9 anos?

16.3 Você está utilizando o ajuste exponencial em uma série temporal anual que corresponde ao total de receitas (em milhões de dólares). Você decide utilizar um coeficiente de ajuste $W = 0{,}20$, e o valor exponencialmente ajustado para 2009 é $E_{2009} = (0{,}20)\,(12{,}1) + (0{,}80)(9{,}4)$.

a. Qual é o valor ajustado dessa série em 2009?

b. Qual é o valor ajustado dessa série em 2010, caso o valor observado da série para esse ano seja de $11,5 milhões?

APLICANDO OS CONCEITOS

AUTO-AVALIAÇÃO **16.4** Os dados a seguir (armazenados no arquivo **Frequência Cinema**) representam a frequência anual a cinemas (em bilhões) de 2001 a 2008:

Ano	Frequência
2001	1,44
2002	1,60
2003	1,52
2004	1,48
2005	1,38
2006	1,40
2007	1,40
2008	1,36

Fonte: *Dados extraídos de Movie Picture Association of America, www.mpaa.org.*

a. Elabore um gráfico para a série temporal.

b. Ajuste uma média móvel de 3 anos a seus dados e coloque os resultados em um gráfico.

c. Utilizando um coeficiente de ajuste $W = 0{,}50$, ajuste exponencialmente a série e faça um gráfico com os resultados.

d. Repita (c), utilizando $W = 0{,}25$.

e. Compare os resultados de (c) e (d).

16.5 Os dados a seguir, contidos no arquivo **NASCAR**, fornecem o número de acidentes na NASCAR Sprint Cup (Copa NASCAR de Velocidade) desde 2001 até 2008.

a. Elabore um gráfico para a série temporal.

b. Ajuste uma média móvel de 3 anos a seus dados e insira os resultados em um gráfico.

c. Utilizando um coeficiente de ajuste $W = 0{,}50$, ajuste exponencialmente a série e elabore um gráfico com os resultados.

d. Repita (c), utilizando $W = 0{,}25$.

e. Compare os resultados de (c) e (d).

Ano	Acidentes
2001	200
2002	186
2003	235
2004	204
2005	253
2006	237
2007	240
2008	211

Fonte: *Dados extraídos de C. Graves, "On-Track Incidents Decrease in Sprint Cup",* USA Today, *16 de dezembro de 2008, p. 1C.*

16.6 O mercado de ações da bolsa eletrônica Nasdaq inclui empresas de pequeno e médio portes, muitas das quais fazem parte do setor de alta tecnologia. Devido à natureza dessas empresas, o índice Nasdaq tende a ser mais volátil do que a Média Industrial Dow Jones ou o Índice S&P 500. Os valores de fechamento diário para o índice Nasdaq durante os primeiros 70 dias de transações em 2009 estão armazenados no arquivo **Nasdaq Diário** (dados extraídos de **finance.yahoo.com**):

a. Elabore um gráfico para a série temporal.

b. Ajuste uma média móvel de 3 anos a seus dados e insira os resultados em um gráfico.

c. Utilizando um coeficiente de ajuste $W = 0{,}50$, ajuste exponencialmente a série e faça um gráfico com os resultados.

d. Repita (c), utilizando $W = 0{,}25$.

e. A que conclusões você pode chegar, no que concerne à presença ou à ausência de tendências durante os primeiros 70 dias de transações em 2009?

16.7 Os dados a seguir (contidos no arquivo **Tesouro**) representam a taxa de juros de títulos do Tesouro Nacional com vencimento em três meses, nos Estados Unidos, de 1991 a 2008:

Ano	Taxa de Juros	Ano	Taxa de Juros
1991	5,38	2000	5,82
1992	3,43	2001	3,40
1993	3,00	2002	1,61
1994	4,25	2003	1,01
1995	5,49	2004	1,37
1996	5,01	2005	3,15
1997	5,06	2006	4,73
1998	4,78	2007	4,36
1999	4,64	2008	1,37

Fonte: *Board of Governors of the Federal Reserve System, www.federalreserve.gov.*

a. Elabore um gráfico para os dados.

b. Faça a adequação de uma média móvel de 3 anos para os seus dados e elabore um gráfico com os resultados.

c. Utilizando o coeficiente de ajuste $W = 0{,}50$, ajuste exponencialmente a série e elabore um gráfico com os resultados.

d. Qual é a sua previsão exponencialmente ajustada para 2009?

e. Repita (c) e (d), utilizando um coeficiente de ajuste de $W = 0{,}25$.

f. Compare os resultados de (d) e (e).

16.8 Os dados a seguir (contidos no arquivo **Eletricidade**) representam a média correspondente aos preços para o consumo de energia elétrica residencial, em custo por quilowatt/hora, nos meses do inverno norte-americano, outubro-março, de 1994-1995 a 2008-2009.

a. Elabore um gráfico para os dados.

b. Faça a adequação de uma média móvel de 3 anos para os seus dados e elabore um gráfico com os resultados.

c. Utilizando o coeficiente de ajuste $W = 0{,}50$, ajuste exponencialmente a série e elabore um gráfico para os resultados.

d. Qual é a sua previsão exponencialmente ajustada para 2009-2010?

e. Repita (c) e (d), utilizando um coeficiente de ajuste de $W = 0{,}25$.

f. Compare os resultados de (d) e (e).

Ano	Custo	Ano	Custo
1994–1995	8,16	2002–2003	8,20
1995–1996	8,10	2003–2004	8,49
1996–1997	8,17	2004–2005	8,78
1997–1998	8,12	2005–2006	9,67
1998–1999	7,94	2006–2007	10,14
1999–2000	7,98	2007–2008	10,49
2000–2001	8,11	2008–2009	11,32
2001–2002	8,37		

Fonte: *Energy Information Administration, Department of Energy, www.eia.doe.gov.*

16.4 Previsão e Ajuste da Tendência dos Mínimos Quadrados

Tendência é o fator componente de uma série temporal mais frequentemente estudado para realizar projeções de médio prazo e de longo prazo. Para obter uma impressão visual das movimentações gerais de longo prazo em uma série temporal, você constrói um gráfico de séries temporais. Caso uma tendência de linha reta se ajuste adequadamente aos dados, você pode utilizar o modelo de tendência linear [veja a Equação (16.3) e a Seção 13.2]. Se os dados da série temporal indicarem algum movimento quadrático de longo prazo, ascendente ou descendente, você pode utilizar o modelo de tendência quadrática [veja a Equação (16.4) e a Seção 15.1]. Quando os dados da série temporal crescem a uma taxa tal que a diferença percentual, de valor para valor, é constante, você pode utilizar um modelo de tendência exponencial [veja a Equação (16.5)].

O Modelo de Tendência Linear

O modelo de tendência linear

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

é o modelo de previsão mais simples. A Equação (16.3) define a equação para previsão de tendência linear.

EQUAÇÃO PARA PREVISÃO DE TENDÊNCIA LINEAR

$$\hat{Y}_i = b_0 + b_1 X_i \qquad \textbf{(16.3)}$$

Lembre-se de que em uma análise de regressão linear você utiliza o método dos mínimos quadrados para calcular a inclinação da amostra, b_1, e o intercepto de Y, b_0. Depois disso, você faz a substituição dos valores correspondentes a X, na Equação (16.3), para prever Y.

Ao utilizar o método dos mínimos quadrados para ajustar tendências em uma série temporal, você inicialmente simplifica a interpretação dos coeficientes atribuindo valores codificados à variável X (tempo). Você atribui números inteiros consecutivos, iniciando com 0, como os valores codificados para os períodos de tempo. Por exemplo, para dados de séries temporais que tenham sido registrados anualmente ao longo de 14 anos, você atribui ao primeiro ano um valor de código correspondente a 0; ao segundo ano um valor de código correspondente a 1; ao terceiro ano um valor de código correspondente a 2, e assim sucessivamente, concluindo com a atribuição de 13 ao 14º ano.

No cenário que trata da The Principled, no início deste capítulo, uma das empresas de interesse é a Coca-Cola Company. Fundada em 1886 e com matriz em Atlanta, Georgia, a Coca-Cola fabrica, distribui e comercializa bebidas não alcoólicas e xaropes por todo o mundo. A empresa comercializa suas bebidas não alcoólicas com os rótulos de marca Coca-Cola, Diet Coke, Sprite, Fanta e Diet Sprite. As receitas em 2008 ultrapassaram $31,9 bilhões (The Coca-Cola Company, **www.coca-cola.com**).

A Tabela 16.2 apresenta a lista com as receitas brutas (em $ bilhões de dólares) no período de 1995 a 2008 (contidos no arquivo **Coca-Cola**),

TABELA 16.2
Receitas (em Bilhões de Dólares), para The Coca-Cola Company (1995-2008)

Ano	Receita	Ano	Receita
1995	18,0	2002	19,6
1996	18,5	2003	21,0
1997	18,9	2004	21,9
1998	18,8	2005	23,1
1999	19,8	2006	24,1
2000	20,5	2007	28,9
2001	20,1	2008	31,9

Fonte: *Dados extraídos de* Mergent's Handbook of Common Stocks, *2006; e www.coca-cola.com.*

A Figura 16.4 apresenta a planilha com os resultados da regressão para a regressão linear simples que utiliza valores codificados consecutivos de 0 a 13 como a variável X (ano codificado). Esses resultados produzem a seguinte equação para previsão de tendência linear:

$$\hat{Y}_i = 16{,}3143 + 0{,}8429X_i$$

em que $X_i = 0$ representa 1995.

FIGURA 16.4
Planilha com os resultados da regressão para um modelo de tendência linear, para prever receitas (em bilhões de dólares) para a Coca-Cola Company

Crie essa planilha utilizando as instruções nas Seções GE13.2 e GE16.4.

	A	B	C	D	E	F	G
1	Modelo de Tendência Linear para a Receita da Coca-Cola Company						
2							
3	*Estatística de Regressão*						
4	R Múltiplo	0,8646					
5	R-quadrado	0,7475					
6	R-quadrado ajustado	0,7265					
7	Erro-padrão	2,1329					
8	Observações	14					
9							
10	ANOVA						
11		*gl*	*SQ*	*MQ*	*F*	*F de significação*	
12	Regressão	1	161,6179	161,6179	33,5260	0,0001	
13	Resíduo	12	54,5914	4,5493			
14	Total	13	216,2093				
15							
16		*Coeficientes*	*Erro-padrão*	*Stat t*	*Valor-p*	*95% inferiores*	*95% superiores*
17	Interseção	16,3143	1,0816	15,0837	0,0000	13,9577	18,6708
18	Ano Ocultado	0,8429	0,1414	5,9604	0,0001	0,5348	1,1510

Você interpreta os coeficientes da regressão da seguinte maneira:

- O intercepto de Y, $b_0 = 16{,}3143$, corresponde às receitas previstas (em bilhões de dólares) para a The Coca-Cola Company durante o ano de origem, ou ano-base, 1995.
- A inclinação, $b_1 = 0{,}8429$, indica que se prevê que as receitas brutas reais cresçam em \$0,8429 bilhão por ano.

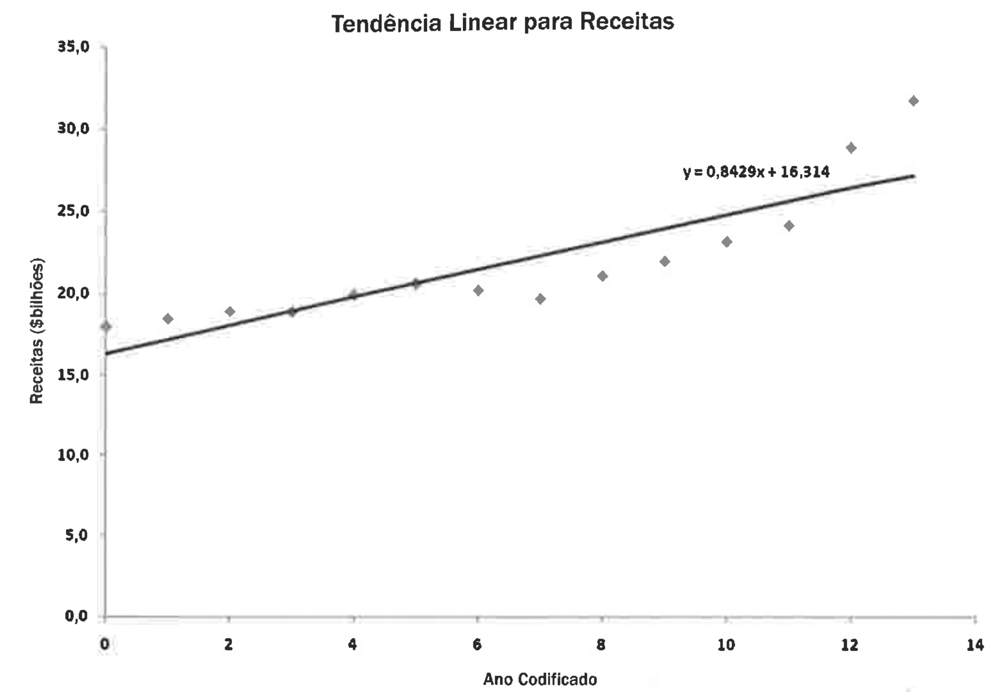
Para projetar a tendência nas receitas da Coca-Cola para o ano de 2009, você substitui X_{15} por 14, o código correspondente a 2009, na equação da previsão de tendência linear:

$$\hat{Y}_i = 16{,}3143 + 0{,}8429(14) = 28{,}1149 \text{ bilhões de dólares}$$

A linha de tendência é colocada em um gráfico na Figura 16.5, juntamente com os valores da série temporal observados. Existe uma forte tendência linear crescente, e o r^2 ajustado é somente 0,7265, indicando que um percentual superior a 72% da variação nas receitas é explicado pela tendência linear ao longo da série temporal. Para investigar se um modelo de tendência diferente poderia proporcionar um melhor ajuste, um modelo de tendência *quadrática* e um modelo de tendência *exponencial* são apresentados a seguir.

FIGURA 16.5
Gráfico da equação para a previsão de tendência linear, para os dados sobre receitas da Coca-Cola Company

Crie essa planilha de gráfico utilizando as instruções para "Acrescentando uma Linha de Previsão e a Equação da Regressão a um Gráfico de Dispersão" na Seção GE13.2.



O Modelo de Tendência Quadrática

Um **modelo de tendência quadrática**

$$\hat{Y}_i = \beta_0 + \beta_1X_i + \beta_2X_i^2 + \varepsilon_i$$

é o modelo não linear mais simples. Utilizando o método dos mínimos quadrados descrito na Seção 15.1, você pode desenvolver uma equação para previsão de tendência quadrática, conforme apresentado na Equação (16.4).

EQUAÇÃO DE PREVISÃO DE TENDÊNCIA QUADRÁTICA

$$\hat{Y}_i = b_0 + b_1X_i + b_2X_i^2 \qquad \textbf{(16.4)}$$

em que

b_0 = intercepto estimado de Y

b_1 = efeito *linear* estimado em Y

b_2 = efeito *quadrático* estimado em Y

A Figura 16.6 apresenta a planilha com os resultados da regressão para o modelo de tendência quadrática utilizado para prever receitas na Coca-Cola Company.

Na Figura 16.6,

$$\hat{Y}_i = 19{,}2804 - 0{,}6402X_i + 0{,}1141X_i^2$$

em que o ano codificado como 0 é 1995.

Para calcular uma previsão utilizando a equação da tendência quadrática, você faz as substituições para os valores de X codificados apropriados dentro dessa equação. Por exemplo, para prever a tendência nas receitas brutas reais para 2009 (ou seja, $X = 14$),

$$\hat{Y}_i = 19{,}2804 - 0{,}6402(14) + 0{,}1141(14)^2 = 32{,}6812$$

A Figura 16.7 apresenta, sob a forma de gráfico, a equação para previsão de tendência quadrática, juntamente com a série temporal para os dados reais. Esse modelo de tendência quadrática proporciona um melhor ajuste (r^2 ajustado = 0,9087) para a série temporal do que o modelo de tendência linear. A estatística do teste t_{ESTAT} para a contribuição do termo quadrático em relação ao modelo é 4,9971 (valor-p = 0,0004).

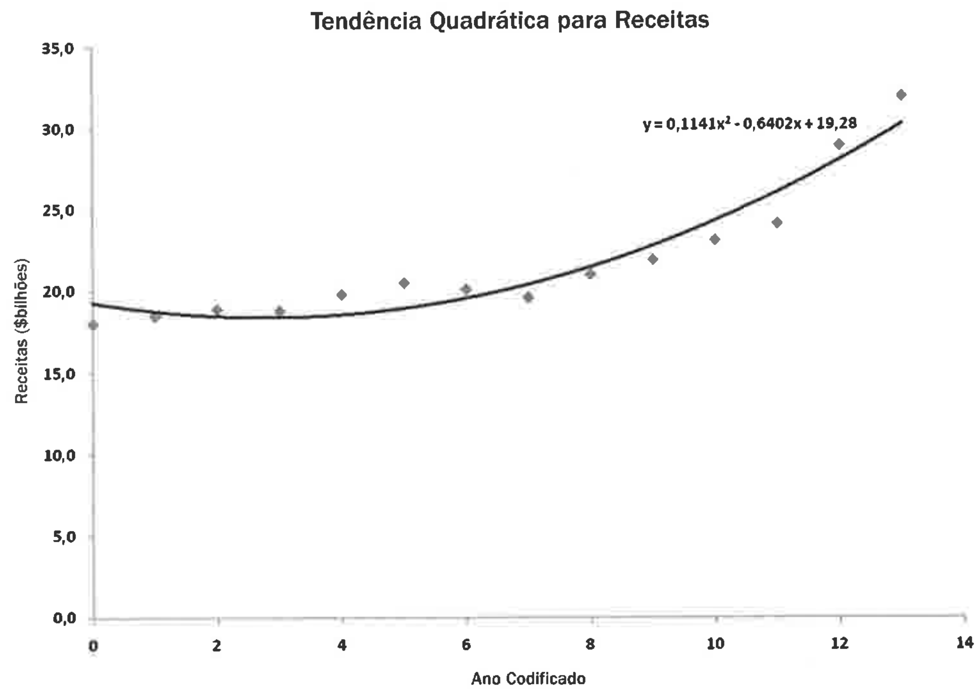
FIGURA 16.6 Planilha com os resultados da regressão para o modelo de tendência quadrática, para prever receitas da Coca-Cola Company

Crie essa planilha de gráfico utilizando as instruções na Seção GE16.4, juntamente com as instruções nas Seções GE15.1 e GE14.1.

	A	B	C	D	E	F	G
1	Modelo de Tendência Quadrática para a Receita da Coca-Cola Company						
2							
3	*Estatística de Regressão*						
4	R Múltiplo	0,9606					
5	R-quadrado	0,9228					
6	R-quadrado ajustado	0,9087					
7	Erro-padrão	1,2319					
8	Observações	14					
9							
10	ANOVA						
11		*gl*	*SQ*	*MQ*	*F*	*F de significação*	
12	Regressão	2	199,5151	99,7576	65,7316	0,0000	
13	Resíduo	11	16,6942	1,5177			
14	Total	13	216,2093				
15							
16		*Coeficientes*	*Erro-padrão*	*Stat t*	*Valor-p*	*95% inferiores*	*95% superiores*
17	Interseção	19,2804	0,8617	22,3742	0,0000	17,0794	21,4813
18	Ano Codificado	-0,6402	0,3078	-2,0798	0,0617	-2,5368	1,2565
19	Ano Codificado ao Quadrado	0,1141	0,0228	4,9971	0,0004	0,5634	0,7916

FIGURA 16.7 Gráfico da equação para a previsão de tendência quadrática, para os dados sobre receitas da Coca-Cola Company

Crie essa planilha de gráfico utilizando as instruções nas Seções GE13.2 e GE16.4.



O Modelo de Tendência Exponencial

Quando uma série temporal cresce a uma taxa tal que a diferença percentual de um valor para outro é constante, está presente uma tendência exponencial. A Equação (16.7) define o **modelo de tendência exponencial**.

MODELO DE TENDÊNCIA EXPONENCIAL

$$Y_i = \beta_0 \beta_1^{X_i} \varepsilon_i \quad \textbf{(16.5)}$$

em que

β_0 = intercepto de Y

$(\beta_1 - 1) \times 100\%$ é a taxa de crescimento anual composta (em %)

O modelo da Equação (16.5) não está no formato de um modelo de regressão linear. Para transformar esse modelo não linear em um modelo linear, você utiliza uma transformação logarítmica de base 10.[1] A adoção do logaritmo em cada um dos lados da Equação (16.5) resulta na Equação (16.6).

[1]Como alternativa, você pode utilizar logaritmos de base e. Para mais informações sobre logaritmos, veja a Seção A.3 no Apêndice A.

MODELO DE TENDÊNCIA EXPONENCIAL TRANSFORMADO

$$\begin{aligned} \log(Y_i) &= \log(\beta_0 \beta_1^{X_i} \varepsilon_i) \quad \textbf{(16.6)} \\ &= \log(\beta_0) + \log(\beta_1^{X_i}) + \log(\varepsilon_i) \\ &= \log(\beta_0) + X_i \log(\beta_1) + \log(\varepsilon_i) \end{aligned}$$

A Equação (16.6) é um modelo linear que você pode estimar utilizando o método dos mínimos quadrados, com $\log(Y_i)$ como a variável dependente e X_i como a variável independente. Isso resulta na Equação (16.7).

EQUAÇÃO PARA PREVISÃO DE TENDÊNCIA EXPONENCIAL

$$\log(\hat{Y}_i) = b_0 + b_1 X_i \quad \textbf{(16.7a)}$$

em que

b_0 = estimativa de $\log(\beta_0)$ e, por conseguinte $10^{b_0} = \hat{\beta}_0$

b_1 = estimativa de $\log(\beta_1)$ e, por conseguinte $10^{b_1} = \hat{\beta}_1$

portanto,

$$\hat{Y}_i = \hat{\beta}_0 \hat{\beta}_1^{X_i} \quad \textbf{(16.7b)}$$

em que

$(\hat{\beta}_1 - 1) \times 100\%$ corresponde à taxa de crescimento anual composta estimada (em %)

A Figura 16.8 ilustra uma planilha de resultados da regressão para um modelo de tendência exponencial relativo às receitas da Coca-Cola Company.

FIGURA 16.8 Planilha com os resultados da regressão para um modelo exponencial, para prever receitas da Coca-Cola Company

Crie essa planilha de gráfico utilizando as instruções nas Seções GE13.2, GE13.5, GE15.2 e GE16.4.

	A	B	C	D	E	F	G
1	Modelo de Tendência Quadrática para a Receita da Coca-Cola Company						
2							
3	*Estatística de Regressão*						
4	R Múltiplo	0,8968					
5	R-quadrado	0,8043					
6	R-quadrado ajustado	0,7880					
7	Erro-padrão	0,0339					
8	Observações	14					
9							
10	ANOVA						
11		*gl*	*SQ*	*MQ*	*F*	*F de significação*	
12	Regressão	1	0,0566	0,0566	49,3150	0,0000	
13	Resíduo	12	0,0138	0,0011			
14	Total	13	0,0704				
15							
16		*Coeficientes*	*Erro-padrão*	*Stat t*	*Valor-p*	*95% inferiores*	*95% superiores*
17	Interseção	1,2296	0,0172	71,5699	0,0000	1,1922	1,2670
18	Ano Codificado	0,0158	0,0022	7,0225	0,0000	0,0109	0,0207

Utilizando a Equação (16.7a) e os resultados gerados na Figura 16.8,

$$\log(\hat{Y}_i) = 1{,}2296 + 0{,}0158 X_i$$

em que o ano codificado como 0 é 1995.

Você calcula os valores para $\hat{\beta}_0$ e $\hat{\beta}_1$ adotando o antilogaritmo dos coeficientes da regressão (b_0 e b_1):

$$\hat{\beta}_0 = \text{antilog}\,(b_0) = \text{antilog}\,(1{,}2296) = 10^{1{,}2296} = 16{,}9668$$

$$\hat{\beta}_1 = \text{antilog}\,(b_1) = \text{antilog}\,(0{,}0158) = 10^{0{,}0158} = 1{,}0371$$

Por conseguinte, utilizando a Equação (16.7b), a equação para previsão de tendência exponencial é

$$\hat{Y}_i = (16{,}9668)(1{,}0371)^{X_i}$$

em que o ano codificado como 0 é 1995.

O intercepto de Y, $\hat{\beta}_0 = 16{,}9668$ bilhões de dólares, corresponde à previsão para as receitas correspondentes ao ano-base 1995. O valor $(\hat{\beta}_1 - 1) \times 100\% = 3{,}71\%$ representa a taxa composta de crescimento anual para as receitas na Coca-Cola Company.

Para fins de prognósticos, você faz as substituições referentes aos valores de X codificados apropriados dentro da Equação (16.7a) ou dentro da Equação (16.7b). Por exemplo, para prever as receitas para 2009 (ou seja, $X = 14$), utilizando a Equação (16.7a):

$$\log(\hat{Y}_i) = 1{,}2296 + 0{,}0158(14) = 1{,}4508$$

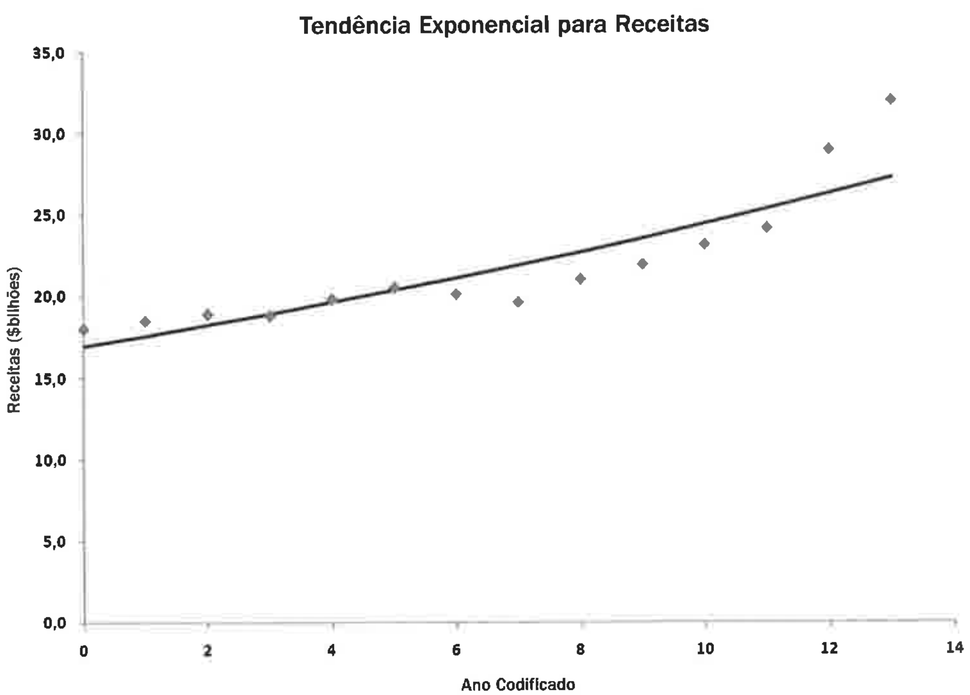
$$\hat{Y}_i = \text{antilog}(1{,}4508) = 10^{1{,}4508} = 28{,}2358 \text{ bilhões de dólares}$$

A Figura 16.9 corresponde a um gráfico da equação para a previsão de tendência exponencial, juntamente com os dados da série temporal. O r^2 ajustado correspondente ao modelo de tendência exponencial (0,7880) é mais alto do que o r^2 ajustado correspondente ao modelo de tendência linear (0,7265), embora seja mais baixo do que o modelo quadrático (0,9087).

FIGURA 16.9

Gráfico da equação para a previsão da tendência exponencial, para as receitas da Coca-Cola Company

Crie essa planilha de gráfico utilizando as instruções na Seção GE16.4.



Seleção de Modelos Utilizando a Primeira Diferença, a Segunda Diferença e Diferenças Percentuais

Você utilizou o modelo linear, o modelo quadrático e o modelo exponencial para prever receitas brutas reais para a Coca-Cola Company. Como você consegue determinar qual desses modelos é o mais apropriado? Além de analisar visualmente gráficos de séries temporais e comparar valores para r^2 ajustados, você pode calcular e examinar a primeira diferença, a segunda diferença e diferenças percentuais. Os elementos de identificação para os modelos de tendência linear, tendência quadrática e tendência exponencial são os seguintes:

- Se um modelo de tendência linear proporciona um ajuste perfeito em relação a uma série temporal, então as primeiras diferenças são constantes. Consequentemente,

$$(Y_2 - Y_1) = (Y_3 - Y_2) = \ldots = (Y_n - Y_{n-1})$$

- Se um modelo de tendência quadrática proporciona um ajuste perfeito em relação a uma série temporal, então as segundas diferenças são constantes. Consequentemente,

$$[(Y_3 - Y_2) - (Y_2 - Y_1)] = [(Y_4 - Y_3) - (Y_3 - Y_2)] = \ldots = [(Y_n - Y_{n-1}) - (Y_{n-1} - Y_{n-2})]$$

- Se um modelo de tendência exponencial proporciona um ajuste perfeito em relação a uma série temporal, então as diferenças percentuais entre valores consecutivos são constantes. Consequentemente,

$$\frac{Y_2 - Y_1}{Y_1} \times 100\% = \frac{Y_3 - Y_2}{Y_2} \times 100\% = \cdots = \frac{Y_n - Y_{n-1}}{Y_{n-1}} \times 100\%$$

Embora não deva esperar um modelo que se ajuste perfeitamente a qualquer conjunto específico de dados de séries temporais, você pode considerar as primeiras diferenças, as segundas diferenças e as diferenças percentuais para uma determinada série como norteadoras na escolha de um modelo apropriado. Os Exemplos 16.2, 16.3 e 16.4 ilustram modelos de tendência linear, quadrática e exponencial que proporcionam ajustes perfeitos (ou praticamente perfeitos) em relação a seus respectivos conjuntos de dados.

EXEMPLO 16.2
Um Modelo de Tendência Linear com um Ajuste Perfeito

A série temporal a seguir representa o número de passageiros por ano (em milhões) para a ABC Airlines:

	Ano									
	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
Passageiros	30,0	33,0	36,0	39,0	42,0	45,0	48,0	51,0	54,0	57,0

Utilizando as primeiras diferenças, mostre que o modelo de tendência linear proporciona um ajuste perfeito em relação a esses dados.

SOLUÇÃO A tabela a seguir mostra a solução:

	Ano									
	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
Passageiros	30,0	33,0	36,0	39,0	42,0	45,0	48,0	51,0	54,0	57,0
Primeiras diferenças		3,0	3,0	3,0	3,0	3,0	3,0	3,0	3,0	3,0

As diferenças entre valores consecutivos na série são as mesmas ao longo de toda a série. Assim, a ABC Airlines exibe um padrão de crescimento linear. O número de passageiros cresce 3 milhões a cada ano.

EXEMPLO 16.3
Um Modelo de Tendência Quadrática com um Ajuste Perfeito

A série temporal a seguir representa o número de passageiros por ano (em milhões) da XYZ Airlines:

	Ano									
	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
Passageiros	30,0	31,0	33,5	37,5	43,0	50,0	58,5	68,5	80,0	93,0

Utilizando as segundas diferenças, mostre que o modelo de tendência quadrática proporciona um ajuste perfeito em relação a esses dados.

SOLUÇÃO A tabela a seguir mostra a solução:

	Ano									
	2000	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
Passageiros	30,0	31,0	33,5	37,5	43,0	50,0	58,5	68,5	80,0	93,0
Primeiras diferenças		1,0	2,5	4,0	5,5	7,0	8,5	10,0	11,5	13,0
Segundas diferenças			1,5	1,5	1,5	1,5	1,5	1,5	1,5	1,5

As segundas diferenças entre pares consecutivos de valores dentro da série são as mesmas ao longo de toda a série. Assim, a XYZ Airlines exibe um padrão de crescimento quadrático. A taxa de crescimento da empresa está acelerando ao longo do tempo.

EXEMPLO 16.4

Um Modelo de Tendência Exponencial com um Ajuste Perfeito

A série temporal a seguir representa o número de passageiros por ano (em milhões) para a EXP Airlines:

	Ano									
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Passageiros	30,0	31,5	33,1	34,8	36,5	38,3	40,2	42,2	44,3	46,5

Utilizando as diferenças percentuais, mostre que o modelo de tendência exponencial proporciona um ajuste quase perfeito em relação a esses dados.

SOLUÇÃO A tabela a seguir mostra a solução:

	Ano									
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Passageiros	30,0	31,5	33,1	34,8	36,5	38,3	40,2	42,2	44,3	46,5
Primeiras diferenças		1,5	1,6	1,7	1,7	1,8	1,9	2,0	2,1	2,2
Diferenças percentuais		5,0	5,1	5,1	4,9	4,9	5,0	5,0	5,0	5,0

As diferenças percentuais entre valores consecutivos na série são aproximadamente as mesmas ao longo de toda a série. Assim, a EXP Airlines exibe um padrão de crescimento exponencial. Sua taxa de crescimento corresponde a aproximadamente 5% ao ano.

A Figura 16.10 apresenta as primeiras diferenças, as segundas diferenças e as diferenças percentuais para os dados sobre receitas da Coca-Cola Company. Nem as primeiras diferenças, nem as segundas diferenças, nem as diferenças percentuais são constantes ao longo da série. Portanto, outros modelos (incluindo aqueles considerados na Seção 16.5) podem ser mais apropriados.

FIGURA 16.10

Planilha que compara as primeiras diferenças, as segundas diferenças e as diferenças percentuais em relação às receitas (em bilhões de dólares) para a Coca-Cola Company

Crie essa planilha utilizando as instruções na Seção GE16.4.

	A	B	C	D	E
1	Ano	Receita	Primeiras Diferenças	Segundas Diferenças	Diferenças Percentuais
2	1995	18,0	#N/D	#N/D	#N/D
3	1996	18,5	0,5	#N/D	2,78%
4	1997	18,9	0,4	-0,1	2,16%
5	1998	18,8	-0,1	-0,5	-0,53%
6	1999	19,8	1,0	1,1	5,32%
7	2000	20,5	0,7	-0,3	3,54%
8	2001	20,1	-0,4	-1,1	-1,95%
9	2002	19,6	-0,5	-0,1	-2,49%
10	2003	21,0	1,4	1,9	7,14%
11	2004	21,9	0,9	-0,5	4,29%
12	2005	23,1	1,2	0,3	5,48%
13	2006	24,1	1,0	-0,2	4,33%
14	2007	28,9	4,8	3,8	19,92%
15	2008	31,9	3,0	-1,8	10,38%

Problemas para a Seção 16.4

APRENDENDO O BÁSICO

16.9 Caso esteja sendo utilizado o método dos mínimos quadrados para ajustar as tendências em uma série temporal anual que contém 25 observações anuais consecutivas,

a. qual valor de código você atribui a X para o primeiro ano na série?

b. qual valor de código você atribui a X para o quinto ano na série?

c. qual valor de código você atribui a X para o ano mais recente que tenha sido registrado na série?

d. qual valor de código você atribui a X caso queira projetar a tendência e fazer uma previsão de 5 anos à frente do último valor observado?

16.10 A equação para previsão de tendência linear para uma série temporal anual que contém 22 valores (de 1985 a 2009) correspondente ao total de receitas (em milhões de dólares) é:

$$\hat{Y}_i = 4{,}0 + 1{,}5X_i$$

a. Interprete o intercepto de Y, b_0.

b. Interprete a inclinação, b_1.

c. Qual é o valor de tendência ajustado para o quinto ano?

d. Qual é o valor de tendência ajustado para o ano mais recente?

e. Qual é a previsão para a tendência projetada de 3 anos depois do último valor?

16.11 A equação para previsão da tendência linear de uma série temporal anual que contém 42 valores (de 1968 a 2009) em vendas líquidas (em bilhões de dólares) é:

$$\hat{Y}_i = 1{,}2 + 0{,}5X_i$$

a. Interprete o intercepto de Y, b_0.

b. Interprete a inclinação, b_1.

c. Qual é o valor de tendência ajustado para o 10º ano?

d. Qual é o valor de tendência ajustado para o ano mais recente?

e. Qual é a previsão para a tendência projetada de 2 anos depois do último valor?

APLICANDO OS CONCEITOS

AUTO-AVALIAÇÃO **16.12** A Bed Bath & Beyond Inc. é uma cadeia de lojas de varejo de âmbito nacional que comercializa uma ampla variedade de mercadorias, incluindo, principalmente, produtos domésticos e artigos de decoração, bem como alimentos, presentes e itens de cuidados de saúde e beleza. Os dados a seguir (armazenados no arquivo **Bed & Bath**) mostram o número de lojas abertas ao final do ano fiscal, desde 1993 até 2009:

Ano	Lojas Abertas	Ano	Lojas Abertas
1993	38	2002	396
1994	45	2003	519
1995	61	2004	629
1996	80	2005	721
1997	108	2006	809
1998	141	2007	888
1999	186	2008	971
2000	241	2009	1,037
2001	311		

Fonte: *Dados extraídos de Bed Bath & Beyond Annual Report, 2005, 2007, 2009.*

a. Elabore um gráfico com os dados.

b. Desenvolva uma equação para previsão de tendência linear e faça um gráfico com os resultados.

c. Desenvolva uma equação para previsão de tendência quadrática e faça um gráfico com os resultados.

d. Desenvolva uma equação para previsão de tendência exponencial e faça um gráfico com os resultados.

e. Utilizando as equações para previsão nos itens (b) a (d), quais são suas previsões anuais para o número de lojas abertas nos anos 2010 e 2011?

f. De que modo você explica as diferenças entre as três previsões em (c)? Qual das previsões você acredita que deveria utilizar? Por quê?

16.13 O Produto Interno Bruto (PIB) é o principal indicador da atividade econômica geral de uma nação. Ele consiste nos gastos pessoais com consumo, investimentos brutos internos, exportações líquidas de bens e serviços e gastos governamentais com consumo. O PIB (em bilhões de dólares atuais) dos Estados Unidos, de 1980 a 2005, está no arquivo **PIB**)

Fonte: *Dados extraídos de Bureau of Economic Analysis, U.S. Department of Commerce, **www.bea.gov**.*

a. Elabore um gráfico com os dados.

b. Desenvolva uma equação de previsão de tendência linear e faça um gráfico para a linha de tendência.

c. Quais são as suas previsões para 2009 e 2010?

d. Que conclusões você consegue tirar no que concerne à tendência no PIB?

16.14 Os dados no arquivo **RecFed** representam as receitas federais de 1978 a 2008, em bilhões de dólares correntes, geradas pelo imposto de renda de pessoa física e de pessoa jurídica, pela seguridade social, pelo imposto sobre o consumo, pelo imposto sobre imóveis e doações, pelas tarifas aduaneiras e pelos depósitos junto ao banco central norte-americano (o Federal Reserve).

Fonte: *Dados extraídos de Tax Policy Center: **www.taxpolicycenter.org**.*

a. Elabore um gráfico com a série de dados.

b. Desenvolva uma equação para a previsão de tendência linear e faça um gráfico para a linha de tendência.

c. Quais são as suas previsões de receitas federais para 2009 e 2010?

d. Que conclusões você consegue tirar no que concerne à tendência nas receitas federais?

16.15 Os dados no arquivo **Estratégica** representam o montante de petróleo, em bilhões de barris, mantidos como reserva estratégica de petróleo dos EUA, de 1981 a 2008.

Fonte: *Dados extraídos de Energy Information Administration, U.S. Department of Energy, **www.eia.doe.gov**.*

a. Elabore um gráfico com os dados.

b. Desenvolva uma equação para a previsão de tendência linear e faça um gráfico para a linha de tendência.

c. Desenvolva uma equação para previsão de tendência quadrática e faça um gráfico com os resultados.

d. Desenvolva uma equação para previsão de tendência exponencial e faça um gráfico com os resultados.

e. Qual modelo é o mais apropriado?

f. Utilizando o modelo mais apropriado, faça a previsão do número de barris, em bilhões de unidades, para 2009. Verifique a precisão de sua previsão localizando o verdadeiro valor para 2009 na Internet ou em uma biblioteca.

16.16 Os dados ilustrados na tabela a seguir (e contidos no arquivo **Energia Solar**) representam as quantidades anuais de energia solar instalada (em megawatts) nos Estados Unidos, desde 2000 até 2008.

Ano	Quantidade de Energia Solar Instalada
2000	18
2001	27
2002	44
2003	68
2004	83
2005	100
2006	140
2007	210
2008	250

Fonte: *Dados extraídos de P. Davidson, "Glut of Rooftop Solar Systems Sinks Price",* USA Today, *31 de janeiro de 2009, p. 1B.*

a. Faça um gráfico com os dados.
b. Desenvolva uma equação para previsão de tendência linear e faça um gráfico para a linha de tendência.
c. Desenvolva uma equação para previsão de tendência quadrática e faça um gráfico com os resultados.
d. Desenvolva uma equação para previsão de tendência exponencial e faça um gráfico com os resultados.
e. Utilizando os modelos em (b) a (d), quais são as suas previsões para a quantidade de energia solar instalada (em megawatts) nos Estados Unidos, nos anos de 2009 e 2010?

16.17 Os dados na tabela a seguir representam os valores de fechamento para a Média Industrial Dow Jones (DJIA), de 1979 a 2007 (veja **DJIA**):

Ano	DJIA	Ano	DJIA	Ano	DJIA
1979	838,7	1989	2.753,2	1999	11.497,1
1980	964,0	1990	2.633,7	2000	10.788,0
1981	875,0	1991	3.168,8	2001	10.021,5
1982	1.046,5	1992	3.301,1	2002	8.341,6
1983	1.258,6	1993	3.754,1	2003	10.453,9
1984	1.211,6	1994	3.834,4	2004	10.788,0
1985	1.546,7	1995	5.117,1	2005	10.717,5
1986	1.896,0	1996	6.448,3	2006	12.463,2
1987	1.938,8	1997	7.908,3	2007	13.264,8
1988	2.168,6	1998	9.181,4		

Fonte: *Dados extraídos de* ***finance.yahoo.com***.

a. Elabore um gráfico com os dados.
b. Desenvolva uma equação para previsão de tendência linear e faça um gráfico para a linha de tendência.
c. Desenvolva uma equação para previsão de tendência quadrática e faça um gráfico com os resultados.
d. Desenvolva uma equação de previsão de tendência exponencial e faça um gráfico com os resultados.
e. Qual modelo é o mais apropriado?
f. Utilizando o modelo mais apropriado, faça a previsão para o valor de fechamento da Média Industrial Dow Jones em 2008. O valor verdadeiro foi 8.776,4. Qual a precisão de seu prognóstico? Caso não tenha sido suficientemente preciso, o que ocorreu?
g. Acrescente ao conjunto de dados o valor relativo a 2008 e refaça os itens de (a) a (e).
h. Utilizando o modelo mais apropriado em (g), faça a previsão para os valores de fechamento para 2009 e 2010.

16.18 A General Electric (GE) é uma das maiores empresas multinacionais do mundo; desenvolve, produz e comercializa uma ampla gama de produtos, incluindo aparelhos de imagem para diagnóstico médico, motores de aeronaves, produtos para iluminação e produtos químicos. Por intermédio de sua afiliada, a NBC Universal, a GE produz e distribui sistemas de rede de televisão e filmes para cinema. Os dados no arquivo **GE** representam o preço de ações em bolsa, em 1º de janeiro, para o período de 22 anos, de 1987 a 2008.

Fonte: *Dados extraídos de* ***finance.yahoo.com***.

a. Elabore um gráfico com os dados.
b. Desenvolva uma equação para previsão de tendência linear e faça um gráfico para a linha de tendência.
c. Desenvolva uma equação para previsão de tendência quadrática e faça um gráfico com os resultados.
d. Desenvolva uma equação para previsão de tendência exponencial e faça um gráfico com os resultados.
e. Qual modelo é o mais apropriado?
f. Utilizando o modelo mais apropriado, faça a previsão para o preço de ações em bolsa em 1º de janeiro de 2009. O valor real foi $16,58. Quão precisa foi a sua previsão? Caso não tenha sido suficientemente precisa, o que ocorreu?
g. Acrescente o valor de 2009 ao conjunto de dados e refaça os itens (a) a (e).
h. Utilizando o modelo mais apropriado em (g), faça a previsão para o preço da ação para 1º de janeiro de 2010 e 1º de janeiro de 2011.

16.19 Embora não deva esperar um modelo perfeitamente ajustado em relação a um determinado conjunto de dados de séries temporais, você pode considerar as primeiras diferenças, as segundas diferenças e as diferenças percentuais, correspondentes a uma determinada série, como norteadores na escolha de um modelo apropriado. Para este problema, utilize cada uma das séries temporais apresentadas na tabela a seguir e armazenadas no arquivo **Modelost1**:

	Ano				
	1999	**2000**	**2001**	**2002**	**2003**
Série temporal I	10,0	15,1	24,0	36,7	53,8
Série temporal II	30,0	33,1	36,4	39,9	43,9
Série temporal III	60,0	67,9	76,1	84,0	92,2

	Ano				
	2004	**2005**	**2006**	**2007**	**2008**
Série temporal I	74,8	100,0	129,2	162,4	199,0
Série temporal II	48,2	53,2	58,2	64,5	70,7
Série temporal III	100,0	108,0	115,8	124,1	132,0

a. Determine o modelo mais apropriado.
b. Desenvolva a equação para a previsão.
c. Faça a previsão do valor para o ano de 2009.

16.20 Um gráfico de séries temporais geralmente ajuda a determinar o modelo apropriado a ser utilizado. Para este problema, utilize cada uma das séries temporais apresentadas na tabela a seguir e armazenadas no arquivo **Modelost2**:

	Ano				
	1999	**2000**	**2001**	**2002**	**2003**
Série temporal I	100,0	115,2	130,1	144,9	160,0
Série temporal II	100,0	115,2	131,7	150,8	174,1

	Ano				
	2004	**2005**	**2006**	**2007**	**2008**
Série temporal I	175,0	189,8	204,9	219,8	235,0
Série temporal II	200,0	230,8	266,1	305,5	351,8

a. Elabore um gráfico para os dados observados (Y) em relação ao período de tempo (X) e elabore um gráfico para o logaritmo dos dados observados (log Y) em relação ao período de tempo (X), para determinar se é mais apropriado um modelo de tendência linear ou um modelo de tendência exponencial. (Dica: Caso o gráfico de log Y em relação a X aparente ser linear, um modelo de tendência exponencial proporciona um ajuste apropriado.)
b. Desenvolva a equação apropriada para previsão.
c. Faça a previsão do valor para o ano de 2009.

16.21 Os dados a seguir, armazenados em **Hotéis**, apresentam as tarifas médias para quartos de hotel, desde 1996 até 2006:

Ano	Tarifa	Ano	Tarifa
1996	70,63	2002	83,54
1997	75,31	2003	82,52
1998	78,62	2004	86,23
1999	81,33	2005	90,88
2000	85,89	2006	97,78
2001	88,27		

Fonte: *Dados extraídos de* USA Today, *Snapshots, 13 de fevereiro de 2008, p. 1A.*

a. Compare as primeiras diferenças, as segundas diferenças e as diferenças percentuais para determinar o modelo mais apropriado.
b. Desenvolva a equação apropriada para previsão.
c. Faça a previsão para a média da tarifa cobrada para quartos de hotel para 2007 e 2008.

16.22 Os dados no arquivo **IPC-U** refletem os valores anuais do IPC (Índice de Preços ao Consumidor) nos Estados Unidos, ao longo do período de 44 anos, de 1965 a 2008, utilizando 1982 a 1984 como período base. Esse índice mede a variação média nos preços, ao longo do tempo, de uma "cesta básica" fixa de bens e serviços adquiridos por todos os consumidores urbanos, incluindo assalariados urbanos (ou seja, trabalhadores em funções administrativas, profissionais especializados, gerentes e técnicos; profissionais autônomos e trabalhadores temporários), desempregados e aposentados.

Fonte: *Dados extraídos de Bureau of Labor Statistics, U.S. Department of Labor,* ***www.bls.gov***.

a. Coloque os dados em um gráfico.
b. Descreva a movimentação nessa série temporal ao longo do período de 44 anos.
c. Desenvolva uma equação de previsão de tendência linear e faça um gráfico com os resultados.
d. Desenvolva uma equação de previsão de tendência quadrática e faça um gráfico com os resultados.
e. Desenvolva uma equação de previsão de tendência exponencial e faça um gráfico com os resultados.
f. Qual modelo é o mais apropriado?
g. Utilizando o modelo mais apropriado, faça a previsão para o IPC de 2009 e de 2010.

16.5 Modelagem Autorregressiva para Ajustes de Tendência e Previsões

Frequentemente, os valores de uma série temporal anuais estão fortemente correlacionados aos valores que os antecedem e àqueles que os sucedem. Esse tipo de correlação é conhecido como *autocorrelação*. **Modelagem autorregressiva**[2] é uma técnica utilizada para prever séries temporais com autocorrelação. Uma **autocorrelação de primeira ordem** refere-se à associação entre valores consecutivos em uma série temporal. Uma **autocorrelação de segunda ordem** refere-se à relação entre valores consecutivos com um intervalo de dois períodos. Uma **autocorrelação de *p*-ésima ordem** refere-se à correlação entre valores em uma série temporal que se encontram a um intervalo de p períodos. Você pode levar em conta a autocorrelação inerente aos dados utilizando métodos de modelagem autorregressiva.

As Equações (16.8), (16.9) e (16.10) definem modelos autorregressivos de primeira ordem, de segunda ordem e de p-ésima ordem.

[2] O modelo de ajuste exponencial, descrito na Seção 16.3, e os modelos autorregressivos descritos nesta seção são casos especiais de médias móveis autorregressivas integradas, desenvolvidas por Box e Jenkins (referência 2).

MODELO AUTORREGRESSIVO DE PRIMEIRA ORDEM

$$Y_i = A_0 + A_1Y_{i-1} + \delta_i \quad \textbf{(16.8)}$$

MODELO AUTORREGRESSIVO DE SEGUNDA ORDEM

$$Y_i = A_0 + A_1Y_{i-1} + A_2Y_{i-2} + \delta_i \quad \textbf{(16.9)}$$

MODELO AUTORREGRESSIVO DE p-ÉSIMA ORDEM

$$Y_i = A_0 + A_1Y_{i-1} + A_2Y_{i-2} + \cdots + A_pY_{i-p} + \delta_i \quad \textbf{(16.10)}$$

em que

Y_i = o valor observado da série no período i

Y_{i-1} = o valor observado da série no período $i - 1$

Y_{i-2} = o valor observado da série no período $i - 2$

Y_{i-p} = o valor observado da série no período $i - p$

$A_0, A_1, A_2,..., A_p$ = parâmetros autorregressivos a serem estimados a partir da análise da regressão dos mínimos quadrados

δ_i = um componente de erro aleatório não autocorrelacionado (com média aritmética = 0 e variância constante)

O **modelo autorregressivo de primeira ordem** [Equação (16.8)] é semelhante, no formato, ao modelo de regressão linear simples [Equação (13.1)]. O **modelo autorregressivo de segunda ordem** [Equação (16.9)] é semelhante ao modelo de regressão múltipla, com duas variáveis independentes [Equação (14.2)]. O **modelo autorregressivo de *p*-ésima ordem** [Equação (16.10)] é semelhante ao modelo de regressão múltipla [Equação (14.1)]. Nos modelos de regressão, os parâmetros da regressão são representados pelos símbolos $\beta_0, \beta_1,..., \beta_k$, com as estimativas correspondentes representadas por $b_0, b_1,..., b_k$. Nos modelos autorregressivos, os parâmetros são representados pelos símbolos $A_0, A_1,..., A_p$, com as estimativas correspondentes representadas por $a_0, a_1,..., a_p$.

A seleção de um modelo autorregressivo apropriado pode ser complicada. Você precisa ponderar as vantagens que são decorrentes da simplicidade em contraposição à preocupação de deixar de levar em conta uma autocorrelação importante nos dados. Você deve também se preocupar com a seleção de um modelo de ordem mais elevada que requeira a estimativa de inúmeros parâmetros desnecessários — especialmente se n, o número de valores na série, for pequeno. A razão para essa preocupação é que, ao se calcular uma estimativa para A_p, p dentre n valores de dados serão perdidos ao se comparar cada um dos valores de dados a um outro valor de dado, p períodos mais cedo.

Os Exemplos 16.5 e 16.6 ilustram essa perda de valores de dados.

EXEMPLO 16.5

Esquema de Comparação para um Modelo Autorregressivo de Primeira Ordem

Considere a série a seguir, com n = 7 valores anuais consecutivos:

	Ano						
	1	2	3	4	5	6	7
Série	31	34	37	35	36	43	40

Mostre as comparações necessárias para um modelo autorregressivo de primeira ordem.

Ano	**Modelo Autorregressivo de Primeira Ordem**
i	(Y_i vs. Y_{i-1})
1	31 ↔ . . .
2	34 ↔ 31
3	37 ↔ 34
4	35 ↔ 37
5	36 ↔ 35
6	43 ↔ 36
7	40 ↔ 43

SOLUÇÃO Uma vez que não existe nenhum valor registrado antes de Y_1, esse valor é perdido para efeito de análise da regressão. Portanto, o modelo autorregressivo de primeira ordem é baseado em seis pares de valores.

EXEMPLO 16.6

Esquema de Comparação para um Modelo Autorregressivo de Segunda Ordem

Considere a seguinte série com n = 7 valores anuais consecutivos:

	Ano						
	1	2	3	4	5	6	7
Série	31	34	37	35	36	43	40

Mostre as comparações necessárias para um modelo autorregressivo de segunda ordem.

Ano	**Modelo Autorregressivo de Segunda Ordem**
i	(Y_i vs. Y_{i-1} e Y_i vs. Y_{i-2})
1	31 ↔ . . . e 31 ↔ . . .
2	34 ↔ 31 e 34 ↔ . . .
3	37 ↔ 34 e 37 ↔ 31
4	35 ↔ 37 e 35 ↔ 34
5	36 ↔ 35 e 36 ↔ 37
6	43 ↔ 36 e 43 ↔ 35
7	40 ↔ 43 e 40 ↔ 36

SOLUÇÃO Uma vez que não existe nenhum valor registrado antes de Y_1, dois valores são perdidos para efeito de análise da regressão. Portanto, o modelo autorregressivo de segunda ordem é baseado em cinco pares de valores.

Depois de selecionar um modelo e utilizar o método dos mínimos quadrados para calcular estimativas para os parâmetros, você precisa determinar a conveniência do modelo. Você pode selecionar um determinado modelo autorregressivo de p-ésima ordem, baseado em experiências anteriores com dados semelhantes, ou, como ponto de partida, você pode escolher um modelo com diversos parâmetros e, em seguida, eliminar os parâmetros de ordem mais elevada que não contribuam significativamente para o modelo. Nessa última abordagem, você utiliza um teste t para a significância de A_p, o parâmetro autorregressivo de ordem mais elevada, no modelo corrente que está sendo considerado. A hipótese nula e a hipótese alternativa são

$$H_0: A_p = 0$$

$$H_1: A_p \neq 0$$

A Equação (16.11) define a estatística do teste.

TESTE t PARA A SIGNIFICÂNCIA DO PARÂMETRO AUTORREGRESSIVO DE ORDEM MAIS ELEVADA, A_P

$$t_{ESTAT} = \frac{a_p - A_p}{S_{a_p}} \qquad \textbf{(16.11)}$$

em que

A_p = valor estipulado na hipótese para o parâmetro de ordem mais elevada, A_p, no modelo autorregressivo

a_p = estimativa do parâmetro de ordem mais elevada, A_p, no modelo autorregressivo

S_{a_p} = desvio-padrão de a_p

A estatística do teste t_{ESTAT} segue uma distribuição t com $n - 2p - 1$ graus de liberdade.[3]

[3]Além dos graus de liberdade perdidos para cada um dos p parâmetros da população que você está estimando, p graus de liberdade adicionais são perdidos pelo fato de existirem p comparações a menos a serem feitas dentre os n valores originais na série temporal.

Para um determinado nível de significância, α, você rejeita a hipótese nula caso a estatística do teste t_{ESTAT} calculada seja maior do que o valor crítico da cauda superior da distribuição t, ou caso a estatística do teste t_{ESTAT} calculada seja menor do que o valor crítico da cauda inferior da distribuição t. Por conseguinte, a regra de decisão é

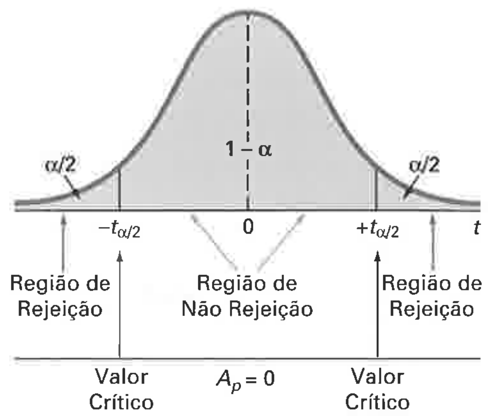
$$\text{Rejeitar } H_0, \text{ se } t_{ESTAT} < -t_{\alpha/2} \text{ ou se } t > t_{\alpha/2};$$

$$\text{caso contrário, não rejeitar } H_0.$$

A Figura 16.11 ilustra a regra de decisão e as regiões de rejeição e de não rejeição.

Caso não rejeite a hipótese nula de que $A_p = 0$, você conclui que o modelo selecionado contém uma quantidade demasiadamente grande de parâmetros estimados. Você descarta, então, o termo de ordem mais elevada e estima um modelo autorregressivo de ordem $p - 1$, utilizando o método dos mínimos quadrados. Depois disso, você repete o teste da hipótese de que o novo parâmetro de ordem mais elevada é igual a 0. Esse procedimento de teste e modelagem continua

FIGURA 16.11
Regiões de rejeição de um teste bicaudal para a significância do parâmetro autorregressivo de ordem mais elevada, A_p



até que você rejeite H_0. Quando isso ocorre, você conclui que o parâmetro de ordem mais elevada remanescente é significativo, e pode utilizar esse modelo para fins de previsão.

A Equação (16.12) define a equação autorregressiva ajustada de p-ésima ordem.

EQUAÇÃO AUTORREGRESSIVA DE p-ÉSIMA ORDEM AJUSTADA

$$\hat{Y}_i = a_0 + a_1 Y_{i-1} + a_2 Y_{i-2} + \cdots + a_p Y_{i-p} \qquad (16.12)$$

em que

$\hat{Y}_i$ = valores ajustados da série no período i

Y_{i-1} = valor observado da série no período $i - 1$

Y_{i-2} = valor observado da série no período $i - 2$

Y_{i-p} = valor observado da série no período $i - p$

$a_0, a_1, a_2, \ldots, a_p$ = estimativas da regressão dos parâmetros $A_0, A_1, A_2, \ldots, A_p$

Você utiliza a Equação (16.13) para prever j anos no futuro, a partir do n-ésimo período de tempo corrente.

EQUAÇÃO PARA PREVISÃO AUTORREGRESSIVA DE p-ÉSIMA ORDEM

$$\hat{Y}_{n+j} = a_0 + a_1 \hat{Y}_{n+j-1} + a_2 \hat{Y}_{n+j-2} + \cdots + a_p \hat{Y}_{n+j-p} \qquad (16.13)$$

em que

$a_0, a_1, a_2, \ldots, a_p$ = estimativas da regressão para os parâmetros $A_0, A_1, A_2, \ldots, A_p$

j = número de anos no futuro

$\hat{Y}_{n+j-p}$ = previsão para Y_{n+j-p} a partir do período de tempo corrente, para $j - p > 0$

$\hat{Y}_{n+j-p}$ = valor observado de Y_{n+j-p} para $j - p \leq 0$

Assim, para fazer previsões para j anos no futuro, utilizando um modelo autorregressivo de terceira ordem, você precisa somente dos $p = 3$ valores mais recentes (Y_n, Y_{n-1} e Y_{n-2}) e das estimativas da regressão a_0, a_1, a_2 e a_3.

Para fazer previsões 1 ano à frente, a Equação (16.13) passa a ser

$$\hat{Y}_{n+1} = a_0 + a_1 Y_n + a_2 Y_{n-1} + a_3 Y_{n-2}$$

Para fazer previsões 2 anos à frente, a Equação (16.13) passa a ser

$$\hat{Y}_{n+2} = a_0 + a_1 \hat{Y}_{n+1} + a_2 Y_n + a_3 Y_{n-1}$$

Para fazer previsões 3 anos à frente, a Equação (16.13) passa a ser

$$\hat{Y}_{n+3} = a_0 + a_1 \hat{Y}_{n+2} + a_2 \hat{Y}_{n+1} + a_3 Y_n$$

e assim sucessivamente.

A modelagem autorregressiva é uma técnica de previsão bastante eficaz para séries temporais que apresentem autocorrelação. Embora ligeiramente mais complicada do que outros métodos, a abordagem passo a passo apresentada a seguir orienta você ao longo da análise:

1. Escolha um valor para p, o parâmetro de ordem mais elevada no modelo autorregressivo a ser avaliado, observando que o teste t para a significância está baseado em $n - 2p - 1$ graus de liberdade.
2. Crie um conjunto de p variáveis "indicadoras do passado" tais que a primeira variável esteja um período de tempo no passado, a segunda variável esteja dois períodos de tempo no passado e assim sucessivamente, e a última variável de previsão esteja p períodos de tempo no passado (veja a Figura 16.12).
3. Realize uma análise dos mínimos quadrados para o modelo de regressão múltipla que contenha todas as p variáveis de previsão do passado (utilizando o Excel).
4. Realize o teste para a significância de A_p, o parâmetro autorregressivo de ordem mais elevada no modelo.
   a. Caso você não rejeite a hipótese nula, descarte a p-ésima variável e repita as etapas 3 e 4. O teste para a significância do novo parâmetro de ordem mais elevada é baseado em uma distribuição t cujos graus de liberdade são revisados de modo tal a que correspondam com o novo número de previsores.
   b. Caso você rejeite a hipótese nula, selecione o modelo autorregressivo com todos os previsores p para fins de ajuste [veja a Equação (16.12)] e de previsão [veja a Equação (16.13)].

Para demonstrar a técnica de modelagem autorregressiva, retorne à série temporal que trata das receitas da Coca-Cola Company, ao longo do período de 14 anos, de 1995 a 2008. A Figura 16.12 exibe uma planilha que organiza os dados para os modelos autorregressivos de primeira ordem, de segunda ordem e de terceira ordem. A planilha contém as variáveis de previsão do passado, Período Passado 1, Período Passado 2 e Período Passado 3 nas colunas C, D e E. Utilize todos os indicadores de previsão do passado para ajustar o modelo autorregressivo de terceira ordem. Utilize somente Período Passado 1 e Período Passado 2 para ajustar o modelo autorregressivo de segunda ordem, e utilize somente Período Passado 1 para ajustar os modelos autorregressivos de primeira ordem. Por conseguinte, dentre os $n = 14$ valores, $p = 1$, 2 ou 3 valores são perdidos durante as comparações necessárias para desenvolver os modelos autorregressivos de primeira ordem, de segunda ordem e de terceira ordem.

FIGURA 16.12
Dados da planilha para desenvolver modelos autorregressivos de primeira ordem, de segunda ordem e de terceira ordem, para as receitas da Coca-Cola Company (1995-2008)

Crie essa planilha utilizando as instruções na Seção GE16.5.

	A	B	C	D	E
1	Ano	Receita Real	Período Passado 1	Período Passado 2	Período Passado 3
2	1995	18,0	#N/D	#N/D	#N/D
3	1996	18,5	18,0	#N/D	#N/D
4	1997	18,9	18,5	18,0	#N/D
5	1998	18,8	18,9	18,5	18,0
6	1999	19,8	18,8	18,9	18,5
7	2000	20,5	19,8	18,8	18,9
8	2001	20,1	20,5	19,8	18,8
9	2002	19,6	20,1	20,5	19,8
10	2003	21,0	19,6	20,1	20,5
11	2004	21,9	21,0	19,6	20,1
12	2005	23,1	21,9	21,0	19,6
13	2006	24,1	23,1	21,9	21,0
14	2007	28,9	24,1	23,1	21,9
15	2008	31,9	28,9	24,1	23,1

A seleção de um modelo autorregressivo que melhor se ajuste a séries temporais anuais começa com o modelo autorregressivo de terceira ordem, ilustrado na Figura 16.13.

Com base na Figura 16.13, a equação autorregressiva de terceira ordem ajustada é:

$$\hat{Y}_i = -15{,}1024 + 0{,}8849 Y_{i-1} + 0{,}3294 Y_{i-2} + 0{,}5987 Y_{i-3}$$

em que o primeiro ano na série é 1998.

FIGURA 16.13
Planilha com os resultados da regressão para o modelo autorregressivo de terceira ordem, para receitas da Coca-Cola Company

Crie essa planilha utilizando as instruções na Seção GE16.5.

	A	B	C	D	E	F	G
1	Modelo Autorregressivo de Terceira Ordem						
2							
3	*Estatística de Regressão*						
4	R Múltiplo	0,9714					
5	R-quadrado	0,9437					
6	R-quadrado ajustado	0,9196					
7	Erro-padrão	1,1814					
8	Observações	11					
9							
10	ANOVA						
11		*gl*	*SQ*	*MQ*	*F*	*F de significação*	
12	Regressão	3	163,7904	54,5968	39,1192	0,0001	
13	Resíduo	7	9,7696	1,3957			
14	Total	10	173,5600				
15							
16		*Coeficientes*	*Erro-padrão*	*Stat t*	*Valor-p*	*95% inferiores*	*95% superiores*
17	Interseção	-15,1024	6,3920	-2,3627	0,0501	-30,3171	0,0123
18	Variável X 1	0,8849	0,3571	2,4778	0,0423	0,0404	1,7293
19	Variável X 2	0,3294	0,7476	0,4406	0,6728	-1,4384	2,0973
20	Variável X 3	0,5987	0,7187	0,8330	0,4324	-1,1008	2,2982

Em seguida, você testa a significância de A_3, o parâmetro de ordem mais elevada. A estimativa do parâmetro de ordem mais elevada, a_3, para o modelo autorregressivo de terceira ordem ajustado é igual a $-0{,}5987$, com um erro-padrão de 0,7187.

Para testar a hipótese nula:

$$H_0: A_3 = 0$$

contra a hipótese alternativa:

$$H_1: A_3 \neq 0$$

utilizando a Equação (16.11) e os resultados da planilha apresentados na Figura 16.13,

$$t_{ESTAT} = \frac{a_3 - A_3}{S_{a_3}} = \frac{0{,}5987 - 0}{0{,}7187} = 0{,}833$$

Utilizando um nível de significância de 0,05, o teste t bicaudal com 7 graus de liberdade apresenta valores críticos de $\pm 2{,}3646$. Tendo em vista que $-2{,}3646 < t_{ESTAT} = -0{,}833 < +2{,}3646$, ou tendo em vista que o valor-p = $0{,}4324 > 0{,}05$, você não rejeita H_0. Você conclui que o parâmetro de terceira ordem do modelo autorregressivo não é significativo, e pode ser excluído.

Depois disso, você ajusta um modelo autorregressivo de segunda ordem (veja a Figura 16.14).

FIGURA 16.14
Planilha com os resultados da regressão para o modelo autorregressivo de segunda ordem, para os dados sobre receitas da Coca-Cola Company

Crie essa planilha utilizando as instruções na Seção GE16.5.

	A	B	C	D	E	F	G
1	Modelo Autorregressivo de Segunda Ordem						
2							
3	*Estatística de Regressão*						
4	R Múltiplo	0,9692					
5	R-quadrado	0,9394					
6	R-quadrado ajustado	0,9259					
7	Erro-padrão	1,1218					
8	Observações	12					
9							
10	ANOVA						
11		*gl*	*SQ*	*MQ*	*F*	*F de significação*	
12	Regressão	2	175,4704	87,7352	69,7153	0,0000	
13	Resíduo	9	11,3263	1,2585			
14	Total	11	186,7967				
15							
16		*Coeficientes*	*Erro-padrão*	*Stat t*	*Valor-p*	*95% inferiores*	*95% superiores*
17	Interseção	-10,7579	4,6459	-2,3156	0,0458	-13,0201	-8,4958
18	Variável X 1	0,9903	0,3246	3,0505	0,0138	-9,5193	11,5000
19	Variável X 2	0,5934	0,5112	1,1607	0,2756	-0,1410	1,3278

A equação autorregressiva de segunda ordem ajustada é

$$\hat{Y}_i = -10{,}7579 + 0{,}9903 Y_{i-1} + 0{,}5934\, Y_{i-2}$$

em que o primeiro ano da série é 1997.

Com base na Figura 16.14, a estimativa para o parâmetro de ordem mais elevada é igual a $a_2 = -0{,}5934$, com um erro-padrão de 0,5112.

Para testar a hipótese nula:

$$H_0: A_2 = 0$$

contra a hipótese alternativa:

$$H_1: A_2 \neq 0$$

utilizando a Equação (16.11),

$$t_{ESTAT} = \frac{a_2 - A_2}{S_{a_2}} = \frac{0{,}5934 - 0}{0{,}5112} = 1{,}1607$$

Utilizando o nível de significância de 0,05, o teste t bicaudal com 9 graus de liberdade possui valores críticos correspondentes a $\pm 2{,}2622$. Uma vez que $-2{,}2622 < t_{ESTAT} = -1{,}1607 < 2{,}2622$, ou tendo em vista que o valor-p = $0{,}2756 > 0{,}05$, você não rejeita H_0. Você conclui que o parâmetro de segunda ordem do modelo autorregressivo não é significativamente importante e deve ser excluído do modelo. Assim, você ajusta um modelo autorregressivo de primeira ordem (veja a Figura 16.15).

FIGURA 16.15
Planilha com os resultados da regressão para o modelo autorregressivo de primeira ordem, para os dados sobre receitas da Coca-Cola Company

Crie essa planilha utilizando as instruções na Seção GE16.5.

	A	B	C	D	E	F	G
1	Modelo Autorregressivo de Primeira Ordem						
2							
3	*Estatística de Regressão*						
4	R Múltiplo	0,9665					
5	R-quadrado	0,9341					
6	R-quadrado ajustado	0,9281					
7	Erro-padrão	1,0969					
8	Observações	14					
9							
10	ANOVA						
11		*gl*	*SQ*	*MQ*	*F*	*F de significação*	
12	Regressão	1	187,4826	187,4826	155,8307	0,0000	
13	Resíduo	11	13,2343	1,2031			
14	Total	12	200,7169				
15							
16		*Coeficientes*	*Erro-padrão*	*Stat t*	*Valor-p*	*95% inferiores*	*95% superiores*
17	Interseção	-5,8384	2,2574	-2,5863	0,0253	-10,8070	-0,8698
18	Ano Codificado	1,3287	0,1064	12,4832	0,0000	1,0944	1,5630

Com base na Figura 16.15, a equação autorregressiva de primeira ordem ajustada é

$$\hat{Y}_i = -5{,}8384 + 1{,}3287 Y_{i-1}$$

em que o primeiro ano da série é 1996.

Com base na Figura 16.15, a estimativa para o parâmetro de ordem mais elevada é $a_1 = 1{,}3287$, com um erro-padrão de 0,1064.

Para testar a hipótese nula:

$$H_0: A_1 = 0$$

contra a hipótese alternativa:

$$H_1: A_1 \neq 0$$

utilizando a Equação (16.11),

$$t_{ESTAT} = \frac{a_1 - A_1}{S_{a_1}} = \frac{1{,}3287 - 0}{0{,}1064} = 12{,}4832$$

Utilizando o nível de significância de 0,05, o teste t bicaudal com 11 graus de liberdade apresenta valores críticos de $\pm 2{,}201$. Uma vez que $t_{ESTAT} = 12{,}4832 > 2{,}201$, ou tendo em vista que o valor-p = $0{,}0000 < 0{,}05$, você rejeita H_0. Você conclui que o parâmetro de primeira ordem do modelo autorregressivo é significativo e deve permanecer no modelo.

A técnica de construção de modelos levou à seleção do modelo autorregressivo de primeira ordem como o mais apropriado para os dados apresentados. Utilizando as estimativas $a_0 = -5{,}8384$

e $a_1 = 1{,}3287$, assim como o valor de dado mais recente, $Y_{14} = 31{,}9$, as previsões para receitas da Coca-Cola Company para 2009 e 2010, partindo da Equação (16.13), são

$$\hat{Y}_{n+j} = -5{,}8384 + 1{,}3287\,\hat{Y}_{n+j-1}$$

Portanto,

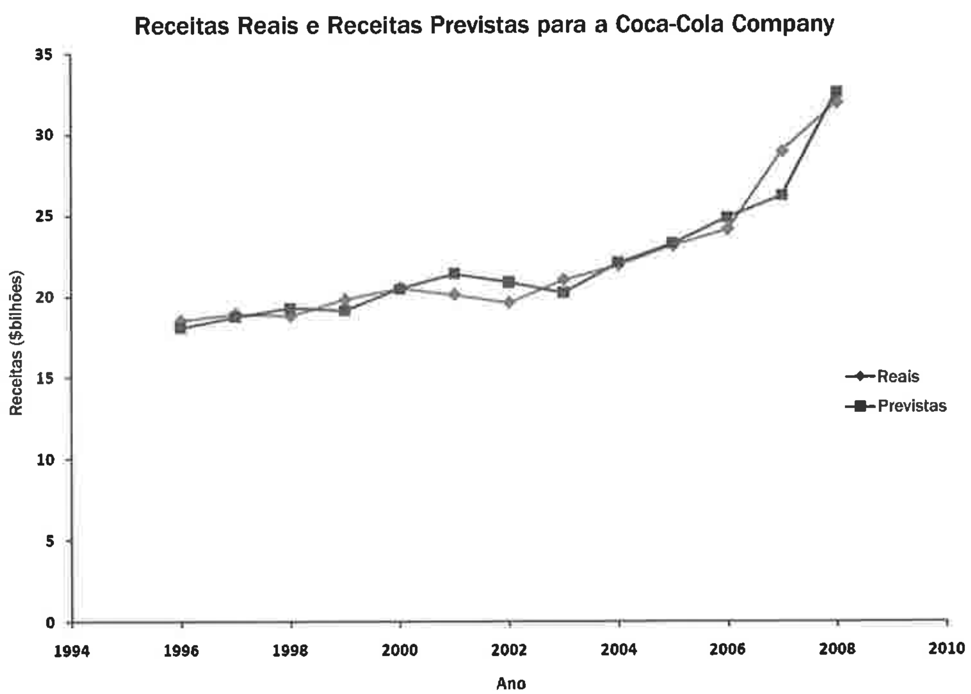
$$2009\text{: 1 ano à frente } \hat{Y}_{15} = -5{,}8384 + 1{,}3287(31{,}9) = 36{,}5471 \text{ bilhões de dólares}$$

$$2010\text{: 2 anos à frente } \hat{Y}_{16} = -5{,}8384 + 1{,}3287\,(36{,}5471) = 42{,}7217 \text{ bilhões de dólares}$$

A Figura 16.16 ilustra os valores reais e os valores previstos para Y oriundos do modelo autorregressivo de primeira ordem.

FIGURA 16.16
Gráfico para receitas reais e receitas previstas, a partir de um modelo autorregressivo de primeira ordem, para a Coca-Cola Company

Crie essa planilha utilizando as instruções na Seção GE2.7.



Problemas para a Seção 16.5

APRENDENDO O BÁSICO

16.23 Foi apresentada a você uma série temporal anual com 40 observações consecutivas, e lhe foi solicitado que ajuste um modelo autorregressivo de quinta ordem.

a. Quantas comparações são perdidas no desenvolvimento do modelo autorregressivo?
b. Quantos parâmetros você precisa estimar?
c. Dos 40 valores originais, de quais você precisa para realizar previsões?
d. Expresse o modelo autorregressivo de quinta ordem.
e. Escreva uma equação para indicar como você realizaria previsões para j anos no futuro.

16.24 Um modelo autorregressivo de terceira ordem é ajustado a uma série temporal anual contendo 17 valores e apresenta os seguintes parâmetros estimados e erros-padrão:

$$a_0 = 4{,}50 \quad a_1 = 1{,}80 \quad a_2 = 0{,}80 \quad a_3 = 0{,}24$$
$$S_{a_1} = 0{,}50 \quad S_{a_2} = 0{,}30 \quad S_{a_3} = 0{,}10$$

No nível de significância de 0,05, teste a adequação do modelo ajustado.

16.25 Reporte-se ao Problema 16.24. Os três valores mais recentes são:

$$Y_{15} = 23 \quad Y_{16} = 28 \quad Y_{17} = 34$$

Faça a previsão para os valores correspondentes ao ano subsequente e ao ano posterior a ele.

16.26 Reporte-se ao Problema 16.24. Suponha que, ao testar a adequação do modelo ajustado, os erros-padrão sejam:

$$S_{a_1} = 0{,}45 \quad S_{a_2} = 0{,}35 \quad S_{a_3} = 0{,}15$$

a. Que conclusões você pode tirar?
b. Discuta sobre o modo de proceder caso a previsão ainda seja o seu principal objetivo.

APLICANDO OS CONCEITOS

16.27 Reporte-se aos dados apresentados no Problema 16.15, que representam o montante de petróleo (em bilhões de barris) mantido como reserva estratégica dos EUA, de 1981 a 2008 (dados armazenados no arquivo **Estratégica**).

a. Faça o ajuste de um modelo autorregressivo de terceira ordem em relação ao montante de petróleo e teste a significância do

parâmetro autorregressivo de terceira ordem. (Utilize $\alpha = 0{,}05$.)
b. Caso necessário, faça o ajuste de um modelo autorregressivo de segunda ordem em relação ao montante de petróleo e teste a significância do parâmetro autorregressivo de segunda ordem. (Utilize $\alpha = 0{,}05$.)
c. Caso necessário, faça o ajuste de um modelo autorregressivo de primeira ordem em relação ao montante de petróleo e teste a significância do parâmetro autorregressivo de primeira ordem. (Utilize $\alpha = 0{,}05$.)
d. Caso seja apropriado, apresente previsões para a quantidade de barris mantida em 2009.

AUTO-AVALIAÇÃO **16.28** Reporte-se aos dados apresentados no Problema 16.12, que correspondem à quantidade de lojas da Bed Bath & Beyond abertas no período entre 1993 e 2009 (dados no arquivo **Bed & Bath**).

a. Faça o ajuste de um modelo autorregressivo de terceira ordem em relação ao número de lojas abertas e teste a significância do parâmetro autorregressivo de terceira ordem. (Utilize $\alpha = 0{,}05$.)
b. Caso necessário, faça o ajuste de um modelo autorregressivo de segunda ordem em relação ao número de lojas abertas e teste a significância do parâmetro autorregressivo de segunda ordem. (Utilize $\alpha = 0{,}05$.)
c. Caso necessário, faça o ajuste de um modelo autorregressivo de primeira ordem em relação ao número de lojas abertas e teste a significância do parâmetro autorregressivo de primeira ordem. (Utilize $\alpha = 0{,}05$.)
d. Caso seja apropriado, apresente a previsão para o número de lojas abertas em 2010 e 2011.

16.29 Reporte-se aos dados apresentados no Problema 16.17, que representam os valores de fechamento da DJIA (Média Industrial Dow Jones), de 1979 a 2007 (dados no arquivo **DJIA**).

a. Faça o ajuste de um modelo autorregressivo de terceira ordem para a Média Industrial Dow Jones e teste a significância do parâmetro autorregressivo de terceira ordem. (Utilize $\alpha = 0{,}05$.)
b. Caso necessário, faça o ajuste de um modelo autorregressivo de segunda ordem para a Média Industrial Dow Jones e teste a significância do parâmetro autorregressivo de segunda ordem. (Utilize $\alpha = 0{,}05$.)
c. Caso necessário, faça o ajuste de um modelo autorregressivo de primeira ordem para a Média Industrial Dow Jones e teste a significância do parâmetro autorregressivo de primeira ordem. (Utilize $\alpha = 0{,}05$.)
d. Faça previsões da Média Industrial Dow Jones para 2008. O valor real foi 8.776,4. Quão precisa foi sua previsão? Caso não tenha sido suficientemente precisa, o que aconteceu?
e. Acrescente o valor correspondente a 2008 ao conjunto de dados, e refaça os itens de (a) a (c).
f. Utilizando o modelo mais apropriado em (e), faça a previsão para os valores de fechamento para 2009 e 2010.

16.30 Reporte-se aos dados apresentados no Problema 16.18 (dados no arquivo **GE**), que representam os preços das ações da GE em bolsa de 1987 a 2008.

a. Faça o ajuste de um modelo autorregressivo de terceira ordem em relação aos preços das ações e teste a significância do parâmetro autorregressivo de terceira ordem. (Utilize $\alpha = 0{,}05$.)
b. Caso necessário, faça o ajuste de um modelo autorregressivo de segunda ordem com relação aos preços das ações e teste a significância do parâmetro autorregressivo de segunda ordem. (Utilize $\alpha = 0{,}05$.)
c. Caso necessário, faça o ajuste de um modelo autorregressivo de primeira ordem com relação aos preços das ações e teste a significância do parâmetro autorregressivo de primeira ordem. (Utilize $\alpha = 0{,}05$.)
d. Faça a previsão para o preço das ações em 1º de janeiro de 2009. O valor real foi $16,58. Quão precisa foi sua previsão? Caso não tenha sido suficientemente precisa, o que aconteceu?
e. Acrescente o valor correspondente a 2008 ao conjunto de dados e refaça os itens de (a) a (c).
f. Utilizando o modelo mais apropriado em (e), faça a previsão para os preços de ações para 1º de janeiro de 2010 e 1º de janeiro de 2011.

16.31 Reporte-se aos dados apresentados no Problema 16.16 (e contidos no arquivo **Energia Solar**) que representam a quantidade anual de energia solar instalada (em megawatts) nos Estados Unidos, desde 2000 até 2008.

a. Faça o ajuste de um modelo autorregressivo de terceira ordem para a quantidade anual de energia solar instalada e teste a significância do parâmetro autorregressivo de terceira ordem. (Utilize $\alpha = 0{,}05$.)
b. Caso necessário, faça o ajuste de um modelo autorregressivo de segunda ordem para a quantidade anual de energia solar instalada e teste a significância do parâmetro autorregressivo de segunda ordem. (Utilize $\alpha = 0{,}05$.)
c. Caso necessário, faça o ajuste de um modelo autorregressivo de primeira ordem para a quantidade anual de energia solar instalada e teste a significância do parâmetro autorregressivo de primeira ordem. (Utilize $\alpha = 0{,}05$.)
d. Faça a previsão para a quantidade anual de energia solar instalada (em megawatts) nos Estados Unidos em 2009 e 2010.

16.6 Escolhendo um Modelo de Previsão Apropriado

Nas Seções 16.4 e 16.5, você estudou seis métodos de previsão de séries temporais para previsões de curto prazo, de médio prazo e de longo prazo: o modelo de tendência linear, o modelo de tendência quadrática e o modelo de tendência exponencial na Seção 16.4; e os modelos autorregressivos de primeira ordem, de segunda ordem e de p-ésima ordem na Seção 16.5. Existe um modelo que seja o *melhor*? Dentre esses modelos, qual deles você deve selecionar para fins de previsão? As diretrizes a seguir são fornecidas para determinar a adequação de um modelo de previsão específico. Essas diretrizes são baseadas no julgamento em relação a quão bem o modelo se ajuste aos dados e adotam o pressuposto de que futuros movimentos na série temporal podem ser projetados com base em um estudo dos dados do passado:

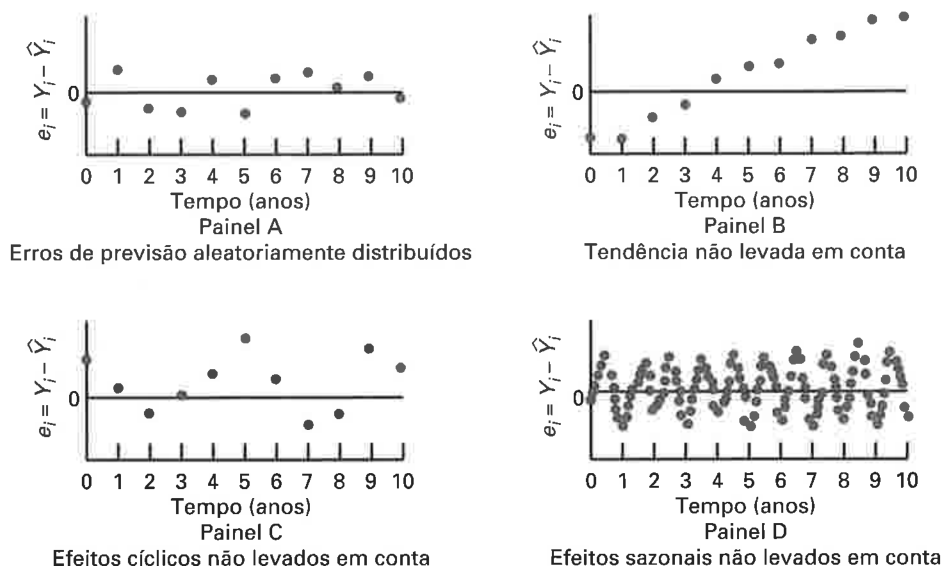
- Realize uma análise nos resíduos.
- Meça a magnitude dos resíduos, por meio das diferenças ao quadrado.
- Meça a magnitude dos resíduos, por meio das diferenças absolutas.
- Utilize o princípio da parcimônia.

Segue uma discussão sobre essas diretrizes.

Realizando uma Análise de Resíduos

Lembre-se, com base nas Seções 13.5 e 14.3, que resíduos são as diferenças entre os valores observados e os valores previstos. Depois de ajustar um determinado modelo a uma dada série temporal, você desenha um gráfico dos resíduos ao longo dos n períodos de tempo. Conforme ilustrado no Painel A da Figura 16.17, caso o modelo em questão se ajuste adequadamente, os resíduos representam o componente irregular da série temporal. Portanto, eles devem estar aleatoriamente distribuídos ao longo da série. Entretanto, conforme ilustrado nos três painéis remanescentes da Figura 16.17, caso o modelo em questão não se ajuste adequadamente, os resíduos podem demonstrar algum padrão sistemático, como, por exemplo, deixar de levar em conta alguma tendência (Painel B), deixar de levar em conta alguma variação cíclica (Painel C) ou, com relação a dados mensais ou trimestrais, deixar de levar em conta variações sazonais (Painel D).

FIGURA 16.17
Análise de resíduos para estudar padrões nos erros



Medindo a Magnitude do Erro Residual por Meio das Diferenças ao Quadrado ou das Diferenças Absolutas

Se, depois de realizar uma análise de resíduos, você ainda acreditar que dois ou mais modelos parecem se ajustar adequadamente aos dados, você pode utilizar outros métodos para a seleção do modelo. Inúmeras medidas baseadas no erro residual encontram-se disponíveis (veja as referências 1 e 4).

Com base no princípio dos mínimos quadrados, uma medida que você já utilizou na análise de regressão (veja a Seção 13.3) é o erro-padrão da estimativa (S_{YX}). Para um determinado modelo, essa medida é baseada na soma das diferenças ao quadrado entre os valores reais e os valores previstos em uma determinada série temporal. Caso um modelo ajuste perfeitamente os dados da série temporal, o erro-padrão da estimativa é, então, igual a zero. Caso um determinado modelo ajuste precariamente os dados da série temporal, o S_{YX} é, então, grande. Consequentemente, ao comparar a adequação de dois ou mais modelos de previsão, você pode selecionar o modelo com o S_{YX} mais baixo como o mais apropriado.

Entretanto, uma grande desvantagem em utilizar o S_{YX} quando se comparam modelos de previsão é que, sempre que existe uma grande diferença entre um valor Y_i individual e $\hat{Y}_i$, o valor de S_{YX} fica superdimensionado em razão do processo de elevação ao quadrado. Por essa razão, muitos estatísticos preferem o **desvio médio absoluto (*DMA*)**. A Equação (16.14) define o *DMA* como a média aritmética das diferenças absolutas entre os valores reais e os valores previstos em uma série temporal.

Calcule o desvio médio absoluto (DMA) utilizando as instruções na Seção GE16.6.

DESVIO MÉDIO ABSOLUTO

$$DMA = \frac{\sum_{i=1}^{n} |Y_i - \hat{Y}_i|}{n} \quad (16.14)$$

Se um modelo ajusta perfeitamente os dados da série temporal, o *DMA* é igual a zero. Se um modelo ajusta precariamente os dados da série temporal, o *DMA* é grande. Ao comparar dois ou mais modelos de previsão, você pode selecionar o modelo com o *DMA* mais baixo como o mais apropriado.

Utilizando o Princípio da Parcimônia

Se, depois de realizar uma análise de resíduos e comparar as medidas de S_{YX} e de *DMA* obtidas, você ainda acreditar que dois ou mais modelos parecem ajustar adequadamente os dados, você pode utilizar então o princípio da parcimônia para a seleção do modelo. Como inicialmente explicado na Seção 15.4, a **parcimônia** orienta você no sentido de selecionar o modelo com a menor quantidade de variáveis independentes que seja capaz de prever adequadamente a variável dependente. Em um sentido mais geral, o princípio da parcimônia orienta você na direção de selecionar o modelo de regressão menos complexo. Dentre os seis modelos de previsão estudados neste capítulo, a maior parte dos estatísticos considera o modelo dos mínimos quadrados linear, o modelo dos mínimos quadrados quadrático e o modelo autorregressivo de primeira ordem mais simples do que os modelos autorregressivos de segunda ordem e de p-ésima ordem e o modelo dos mínimos quadrados exponencial.

Uma Comparação entre Quatro Métodos de Previsão

Considere, mais uma vez, os dados sobre receitas da Coca-Cola Company. Para ilustrar o processo de seleção do modelo, você pode comparar quatro dos métodos de previsão utilizados nas Seções 16.4 e 16.5: o modelo linear, o modelo quadrático, o modelo exponencial e o modelo autorregressivo de primeira ordem. (Não há necessidade de estudar mais profundamente os modelos autorregressivos de segunda ordem ou de terceira ordem para essa série temporal, uma vez que esses modelos não melhoraram significativamente o ajuste, em comparação com o modelo autorregressivo de primeira ordem.)

A Figura 16.18 ilustra os gráficos de resíduos para os quatro modelos. Ao inferir conclusões desses gráficos de resíduos, você deve tomar cuidado, uma vez que existem somente 14 valores.

Na Figura 16.18, observe a estrutura sistemática dos resíduos no modelo linear (Painel A), no modelo quadrático (Painel B) e no modelo exponencial (Painel C). No que diz respeito ao modelo autorregressivo (Painel D), os resíduos aparentam ser mais aleatórios.

Resumindo, com base nas análises de resíduos de todos os quatro modelos de previsão, parece que o modelo autorregressivo de primeira ordem é o mais apropriado, enquanto os modelos linear, exponencial e quadrático são menos apropriados. Para verificar ainda mais minuciosamente esse fato, você pode comparar os quatro modelos quanto à magnitude de seus resíduos. A Figura 16.19 apresenta os valores reais (Y_i), juntamente com os valores previstos ($\hat{Y}_i$), os resíduos (e_i), a soma dos quadrados dos resíduos (*SQR*), o erro-padrão da estimativa (S_{YX}) e o desvio médio absoluto (*DMA*) em relação a cada um dos quatro modelos.

Para essa série temporal, S_{YX} e *DMA* proporcionam resultados similares. Uma comparação entre S_{YX} e *DMA* indica claramente que o modelo linear proporciona o ajuste mais precário. O modelo autorregressivo de primeira ordem proporciona o melhor ajuste. Considerando os resultados da análise de resíduos, juntamente com S_{YX} e *DMA*, o modelo autorregressivo de primeira ordem é a escolha como o melhor modelo.

Depois de selecionar um determinado modelo de previsão, você precisa monitorar continuamente as suas previsões. Caso ocorram erros de grande dimensão entre os valores previstos e os valores reais, a estrutura subjacente da série temporal pode ter se modificado. Lembre-se de que os métodos de previsão apresentados neste capítulo pressupõem que os padrões inerentes ao passado terão continuidade no futuro. Erros de previsão de grande porte são um indicativo de que um determinado pressuposto não mais é verdadeiro.

FIGURA 16.18
Gráficos de resíduos para quatro métodos de previsão

Crie gráficos de resíduos para os quatro métodos de previsão utilizando as instruções na Seção GE16.6.

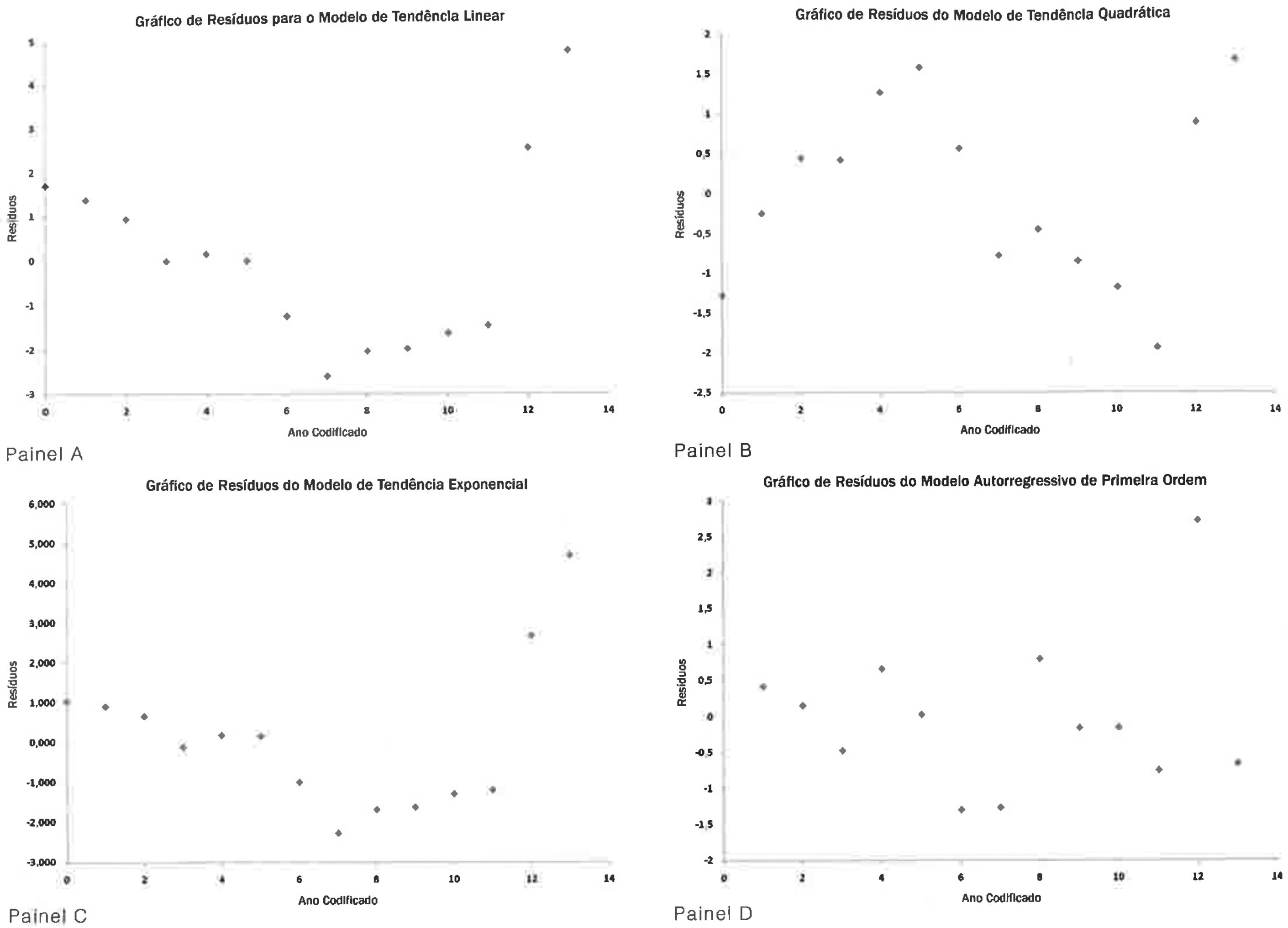


FIGURA 16.19
Planilha que sintetiza e compara quatro métodos para previsões, utilizando S_{YX} e *DMA*

	A	B	D	E	G	H	J	K	M	N
1			Linear		Quadrático		Exponencial		Autorregressivo de Primeira Ordem	
2	Ano	Receita	Previsto	Resíduo	Previsto	Resíduo	Previsto	Resíduo	Previsto	Resíduo
3	1995	18,0	16,3143	1,6857	19,2804	-1,2804	16,9670	1,0330	#N/D	#N/D
4	1996	18,5	17,1571	1,3429	18,7543	-0,2543	17,5946	0,9054	18,0781	0,4219
5	1997	18,9	18,0000	0,9000	18,4563	0,4437	18,2454	0,6546	18,7424	0,1576
6	1998	18,8	18,8429	-0,0429	18,3865	0,4135	18,9203	-0,1203	19,2739	-0,4739
7	1999	19,8	19,6857	0,1143	18,5449	1,2551	19,6201	0,1799	19,1411	0,6589
8	2000	20,5	20,5286	-0,0286	18,9315	1,5685	20,3458	0,1542	20,4697	0,0303
9	2001	20,1	21,3714	-1,2714	19,5462	0,5538	21,0984	-0,9984	21,3998	-1,2998
10	2002	19,6	22,2143	-2,6143	20,3890	-0,7890	21,8788	-2,2788	20,8684	-1,2684
11	2003	21,0	23,0571	-2,0571	21,4600	-0,4600	22,6881	-1,6881	20,2040	0,7960
12	2004	21,9	23,9000	-2,0000	22,7592	-0,8592	23,5273	-1,6273	22,0642	-0,1642
13	2005	23,1	24,7429	-1,6429	24,2865	-1,1865	24,3976	-1,2976	23,2600	-0,1600
14	2006	24,1	25,5857	-1,4857	26,0420	-1,9420	25,3000	-1,2000	24,8544	-0,7544
15	2007	28,9	26,4286	2,4714	28,0257	0,8743	26,2358	2,6642	26,1831	2,7169
16	2008	31,9	27,2714	4,6286	30,2375	1,6625	27,2063	4,6937	32,5608	-0,6608
17			*SQR*	54,5914	*SQR*	16,6942	*SQR*	46,326	*SQR*	13,2343
18			S_{YX}	2,1329	S_{YX}	1,2319	S_{YX}	1,965	S_{YX}	1,0969
19			*DMA*	1,592	*DMA*	0,967	*DMA*	1,393	*DMA*	0,762

Problemas para a Seção 16.6

APRENDENDO O BÁSICO

16.32 Os resíduos a seguir são oriundos de um modelo de tendência linear utilizado para prever vendas:

2,0 −0,5 1,5 1,0 0,0 1,0 −3,0 1,5 −4,5 2,0 0,0 −1,0

a. Calcule S_{YX} e interprete suas descobertas.
b. Calcule o *DMA* e interprete suas descobertas.

16.33 Reporte-se ao Problema 16.32. Suponha que o primeiro resíduo seja igual a 12,0 (em vez de 2,0) e que o último valor seja igual a −11,0 (em vez de −1,0).
a. Calcule S_{YX} e interprete suas descobertas.
b. Calcule o *DMA* e interprete suas descobertas.

APLICANDO OS CONCEITOS

16.34 Reporte-se aos resultados do Problema 16.13 (veja o arquivo **PIB**).
a. Realize uma análise nos resíduos.
b. Calcule o erro-padrão da estimativa (S_{YX}).
c. Calcule o *DMA*.
d. Com base em (a) a (c), você está satisfeito com suas previsões de tendência linear no Problema 16.13? Discuta.

16.35 Reporte-se aos resultados do Problema 16.15, e do Problema 16.27, que tratam do número de barris de petróleo na reserva estratégica de petróleo dos EUA (dados no arquivo **Estratégica**).
a. Realize uma análise nos resíduos para cada um dos modelos.
b. Calcule o erro-padrão da estimativa (S_{YX}) para cada um dos modelos.
c. Calcule o *DMA* para cada um dos modelos.
d. Com base em (a) a (c) e no princípio da parcimônia, qual modelo de previsão você selecionaria? Discuta.

AUTO-AVALIAÇÃO **16.36** Reporte-se aos resultados para o Problema 16.12 e para o Problema 16.28, que tratam do número de lojas da Bed Bath & Beyond abertas (dados no arquivo **Bed & Bath**).
a. Realize uma análise nos resíduos para cada um dos modelos.
b. Calcule o erro-padrão da estimativa (S_{YX}) para cada um dos modelos.
c. Calcule o *DMA* para cada um dos modelos.
d. Com base em (a) a (c) e no princípio da parcimônia, qual modelo de previsão você selecionaria? Discuta.

16.37 Reporte-se aos resultados do Problema 16.17 e do Problema 16.29, que tratam da Média Industrial Dow Jones (DJIA) (dados no arquivo **DJIA**).
a. Realize uma análise nos resíduos para cada um dos modelos.
b. Calcule o erro-padrão da estimativa (S_{YX}) para cada um dos modelos.
c. Calcule o *DMA* para cada um dos modelos.
d. Com base em (a) a (c) e no princípio da parcimônia, qual modelo de previsão você selecionaria? Discuta.

16.38 Reporte-se aos resultados do Problema 16.18 e do Problema 16.30, que tratam do preço por cota de ações da GE (dados no arquivo **GE**).
a. Realize uma análise de resíduos para cada um dos modelos.
b. Calcule o erro-padrão da estimativa (S_{YX}) para cada um dos modelos.
c. Calcule o *DMA* para cada um dos modelos.
d. Com base em (a) a (c) e no princípio da parcimônia, qual modelo de previsão você selecionaria? Discuta.

16.39 Reporte-se aos resultados do Problema 16.16 e do Problema 16.31, que tratam da quantidade de energia solar instalada (em megawatts) nos Estados Unidos, de 2000 a 2008 (dados no arquivo **Energia Solar**).
a. Realize uma análise de resíduos para cada um dos modelos.
b. Calcule o erro-padrão da estimativa (S_{YX}) para cada um dos modelos.
c. Calcule o *DMA* para cada um dos modelos.
d. Com base em (a) a (c) e no princípio da parcimônia, qual modelo de previsão você selecionaria? Discuta.

16.7 Previsão de Séries Temporais para Dados Sazonais

Até este ponto, este capítulo centrou seu foco na previsão de dados anuais. Entretanto, inúmeras séries temporais são coletadas trimestral ou mensalmente, e outras, ainda, são obtidas semanalmente, diariamente e até mesmo em uma base horária. Quando uma série temporal é coletada trimestral ou mensalmente, você deve considerar o impacto decorrente dos efeitos sazonais. Nesta seção, a construção de modelos de regressão é utilizada para prever dados mensais ou trimestrais.

Uma das empresas de interesse no cenário Utilizando a Estatística é a Wal-Mart Stores, Inc. Em 2009, a Wal-Mart, Inc. operava mais de 8.000 unidades de vendas de varejo em 15 países e tinha um patamar de receitas que excedia \$ 400 bilhões (Wal-Mart Stores, Inc., **investor.walmartstores.com**). As receitas da Wal-Mart são altamente sazonais, e, portanto, você precisa analisar as receitas trimestrais. O ano fiscal da empresa se encerra em 31 de janeiro. Consequentemente, o quarto trimestre de 2009 inclui os meses de novembro e dezembro de 2008 e janeiro de 2009. A Tabela 16.3 apresenta uma lista com as receitas trimestrais, em bilhões de dólares, de 2004 a 2009 (contidas no arquivo **WalMart**). A Figura 16.20 apresenta a série temporal.

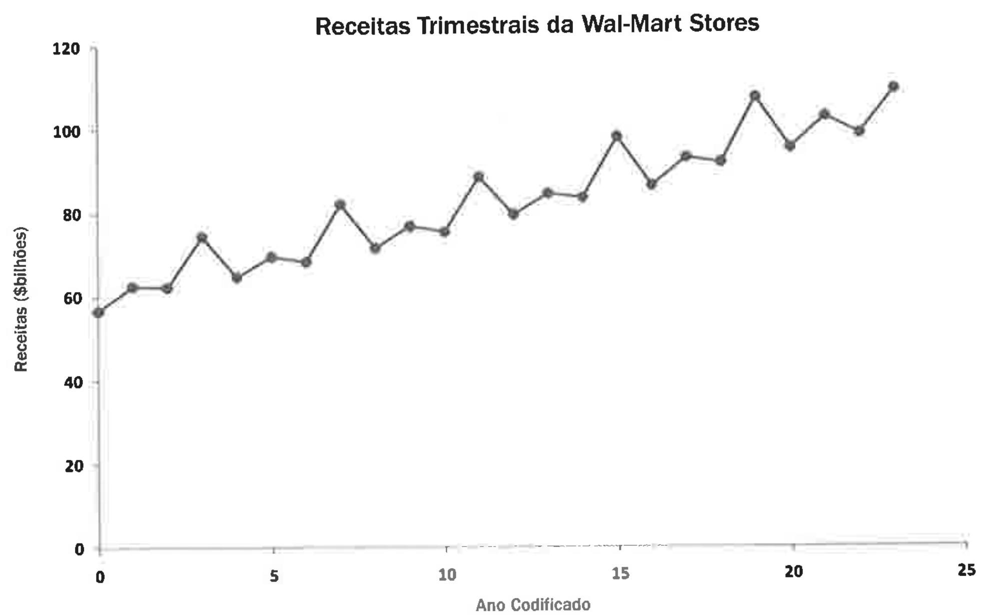
TABELA 16.3
Receitas Trimestrais da Wal-Mart Stores, Inc., em Bilhões de Dólares (2004-2009)

	Ano					
Trimestre	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
1	56,7	64,8	71,6	79,6	86,4	95,3
2	62,6	69,7	76,8	84,5	93,0	102,7
3	62,4	68,5	75,4	83,5	91,9	98,6
4	74,5	82,2	88,6	98,1	107,3	109,1

Fonte: *Dados extraídos de Wal-Mart Stores, Inc.*, ***investor.walmartstores.com***.

FIGURA 16.20
Gráfico para receitas trimestrais da Wal-Mart Stores, Inc., em bilhões de dólares (2004-2009)

Crie essa planilha utilizando as instruções na Seção GE2.7.



Previsão dos Mínimos Quadrados com Dados Mensais ou Trimestrais

Para desenvolver um modelo de regressão dos mínimos quadrados que inclua um componente sazonal, a abordagem do ajuste de tendência dos mínimos quadrados, na Seção 16.4, é combinada a variáveis binárias (*dummy*) (veja a Seção 14.6) para modelar o componente sazonal.

A Equação (16.15) define o modelo de tendência exponencial para dados trimestrais.

MODELO EXPONENCIAL COM DADOS TRIMESTRAIS

$$Y_i = \beta_0 \beta_1^{X_i} \beta_2^{Q_1} \beta_3^{Q_2} \beta_4^{Q_3} \varepsilon_i \qquad \textbf{(16.15)}$$

em que

X_i = valor trimestral codificado, $i = 0, 1, 2,...$

Q_1 = 1 se for o primeiro trimestre, 0 se não for o primeiro trimestre

Q_2 = 1 se for o segundo trimestre, 0 se não for o segundo trimestre

Q_3 = 1 se for o terceiro trimestre, 0 se não for o terceiro trimestre

β_0 = intercepto de Y

$(\beta_1 - 1) \times 100\%$ = taxa de crescimento trimestral composta (em %)

β_2 = multiplicador do primeiro trimestre em relação ao quarto trimestre

β_3 = multiplicador do segundo trimestre em relação ao quarto trimestre

β_4 = multiplicador do terceiro trimestre em relação ao quarto trimestre

ε_i = valor do componente irregular para o período de tempo i

O modelo na Equação (16.15) não está sob a forma de um modelo de regressão linear. Para transformar esse modelo não linear em um modelo linear, você utiliza uma transformação logarítmica de base 10.[4] A aplicação do logaritmo em cada um dos lados da Equação (16.15) resulta na Equação (16.16).

[4]Alternativamente, você pode utilizar logaritmos na base *e*. Para mais informações sobre logaritmos, veja a Seção A.3 do Apêndice A.

MODELO EXPONENCIAL TRANSFORMADO COM DADOS TRIMESTRAIS

$$\begin{aligned}\log(Y_i) &= \log(\beta_0 \beta_1^{X_i} \beta_2^{Q_1} \beta_3^{Q_2} \beta_4^{Q_3} \varepsilon_i) \qquad \textbf{(16.16)}\\ &= \log(\beta_0) + \log(\beta_1^{X_i}) + \log(\beta_2^{Q_1}) + \log(\beta_3^{Q_2}) + \log(\beta_4^{Q_3}) + \log(\varepsilon_i)\\ &= \log(\beta_0) + X_i \log(\beta_1) + Q_1 \log(\beta_2) + Q_2 \log(\beta_3) + Q_3 \log(\beta_4) + \log(\varepsilon_i)\end{aligned}$$

A Equação (16.16) representa um modelo linear que você pode estimar utilizando a regressão dos mínimos quadrados. A realização da análise da regressão utilizando $\log(Y_i)$ como a variável dependente e X_i, Q_1, Q_2 e Q_3 como as variáveis independentes resulta na Equação (16.17).

EQUAÇÃO DE PREVISÃO DE CRESCIMENTO EXPONENCIAL COM DADOS TRIMESTRAIS

$$\log(\hat{Y}_i) = b_0 + b_1 X_i + b_2 Q_1 + b_3 Q_2 + b_4 Q_3 \qquad \textbf{(16.17)}$$

em que

b_0 = estimativa de $\log(\beta_0)$ e por conseguinte $10^{b_0} = \hat{\beta}_0$

b_1 = estimativa de $\log(\beta_1)$ e por conseguinte $10^{b_1} = \hat{\beta}_1$

b_2 = estimativa de $\log(\beta_2)$ e por conseguinte $10^{b_2} = \hat{\beta}_2$

b_3 = estimativa de $\log(\beta_3)$ e por conseguinte $10^{b_3} = \hat{\beta}_3$

b_4 = estimativa de $\log(\beta_4)$ e por conseguinte $10^{b_4} = \hat{\beta}_4$

A Equação (16.18) é utilizada para dados mensais.

MODELO EXPONENCIAL COM DADOS MENSAIS

$$Y_i = \beta_0 \beta_1^{X_i} \beta_2^{M_1} \beta_3^{M_2} \beta_4^{M_3} \beta_5^{M_4} \beta_6^{M_5} \beta_7^{M_6} \beta_8^{M_7} \beta_9^{M_8} \beta_{10}^{M_9} \beta_{11}^{M_{10}} \beta_{12}^{M_{11}} \varepsilon_i \qquad \textbf{(16.18)}$$

em que

X_i = valor mensal codificado, $i = 0, 1, 2,...$

M_1 = 1 se for janeiro, 0 se não for janeiro

M_2 = 1 se for fevereiro, 0 se não for fevereiro

M_3 = 1 se for março, 0 se não for março

⋮

M_{11} = 1 se for novembro, 0 se não for novembro

β_0 = intercepto de Y

$(\beta_1 - 1) \times 100\%$ = taxa de crescimento mensal composta (em %)

β_2 = multiplicador de janeiro em relação a dezembro

β_3 = multiplicador de fevereiro em relação a dezembro

β_4 = multiplicador de março em relação a dezembro

⋮

β_{12} = multiplicador de novembro em relação a dezembro

ε_i = valor do componente irregular para o período de tempo i

O modelo na Equação (16.18) não está no formato de um modelo de regressão linear. Para transformar esse modelo não linear em um modelo linear, você pode utilizar uma transformação logarítmica de base 10. A aplicação do logaritmo em cada um dos lados da Equação (16.18) resulta na Equação (16.19).

MODELO EXPONENCIAL TRANSFORMADO COM DADOS MENSAIS

$$\begin{aligned}\log(Y_i) &= \log(\beta_0\beta_1^{X_i}\beta_2^{M_1}\beta_3^{M_2}\beta_4^{M_3}\beta_5^{M_4}\beta_6^{M_5}\beta_7^{M_6}\beta_8^{M_7}\beta_9^{M_8}\beta_{10}^{M_9}\beta_{11}^{M_{10}}\beta_{12}^{M_{11}}\varepsilon_i) \qquad \textbf{(16.19)}\\ &= \log(\beta_0) + X_i\log(\beta_1) + M_1\log(\beta_2) + M_2\log(\beta_3)\\ &\quad + M_3\log(\beta_4) + M_4\log(\beta_5) + M_5\log(\beta_6) + M_6\log(\beta_7)\\ &\quad + M_7\log(\beta_8) + M_8\log(\beta_9) + M_9\log(\beta_{10}) + M_{10}\log(\beta_{11})\\ &\quad + M_{11}\log(\beta_{12}) + \log(\varepsilon_i)\end{aligned}$$

A Equação (16.19) representa um modelo linear que você pode estimar utilizando o método dos mínimos quadrados. A realização da análise da regressão utilizando $\log(Y_i)$ como a variável dependente e X_i, M_1, M_2,..., e M_{11} como as variáveis independentes resulta na Equação (16.20).

EQUAÇÃO PARA PREVISÃO DE CRESCIMENTO EXPONENCIAL COM DADOS MENSAIS

$$\begin{aligned}\log(\hat{Y}_i) &= b_0 + b_1X_i + b_2M_1 + b_3M_2 + b_4M_3 + b_5M_4 + b_6M_5 + b_7M_6\\ &\quad + b_8M_7 + b_9M_8 + b_{10}M_9 + b_{11}M_{10} + b_{12}M_{11} \qquad \textbf{(16.20)}\end{aligned}$$

em que

$$b_0 = \text{estimativa de } \log(\beta_0) \text{ e por conseguinte } 10^{b_0} = \hat{\beta}_0$$
$$b_1 = \text{estimativa de } \log(\beta_1) \text{ e por conseguinte } 10^{b_1} = \hat{\beta}_1$$
$$b_2 = \text{estimativa de } \log(\beta_2) \text{ e por conseguinte } 10^{b_2} = \hat{\beta}_2$$
$$b_3 = \text{estimativa de } \log(\beta_3) \text{ e por conseguinte } 10^{b_3} = \hat{\beta}_3$$
$$\vdots$$
$$b_{12} = \text{estimativa de } \log(\beta_{12}) \text{ e por conseguinte } 10^{b_{12}} = \hat{\beta}_{12}$$

Q_1, Q_2 e Q_3 são as três variáveis binárias (*dummy*) necessárias para representar os 4 períodos trimestrais, em uma série temporal trimestral. M_1, M_2,... e M_{11} correspondem às 11 variáveis binárias (*dummy*) necessárias para representar os 12 meses em uma série temporal mensal. Na construção do modelo, você utiliza $\log(Y_i)$ no lugar dos valores de Y_i e, depois disso, encontra os coeficientes da regressão com a aplicação do antilogaritmo dos coeficientes da regressão desenvolvidos a partir das Equações (16.17) e (16.20).

Embora, à primeira vista, esses modelos de regressão possam parecer complexos, ao serem efetuados os ajustes ou as previsões em qualquer período de tempo, os valores de todas ou o valor de todas menos uma das outras variáveis binárias (*dummy*) no modelo são considerados iguais a zero, e as equações são drasticamente simplificadas. Ao estabelecer as variáveis binárias (*dummy*) para dados de séries temporais trimestrais, o quarto trimestre corresponde ao período base e tem zero como valor codificado para cada uma das variáveis binárias (*dummy*). Com uma série temporal trimestral, a Equação (16.17) fica reduzida do seguinte modo:

Para qualquer primeiro trimestre: $\log(\hat{Y}_i) = b_0 + b_1X_i + b_2$
Para qualquer segundo trimestre: $\log(\hat{Y}_i) = b_0 + b_1X_i + b_3$
Para qualquer terceiro trimestre: $\log(\hat{Y}_i) = b_0 + b_1X_i + b_4$
Para qualquer quarto trimestre: $\log(\hat{Y}_i) = b_0 + b_1X_i$

Ao serem estabelecidas as variáveis binárias (*dummy*) para cada um dos meses, dezembro serve de ano-base e tem 0 como valor codificado para cada uma das variáveis binárias (*dummy*). Por exemplo, com uma série temporal mensal, a Equação (16.20) é reduzida do seguinte modo:

Para qualquer janeiro: $\log(\hat{Y}_i) = b_0 + b_1X_i + b_2$
Para qualquer fevereiro: $\log(\hat{Y}_i) = b_0 + b_1X_i + b_3$
$\vdots$
Para qualquer novembro: $\log(\hat{Y}_i) = b_0 + b_1X_i + b_{12}$
Para qualquer dezembro: $\log(\hat{Y}_i) = b_0 + b_1X_i$

Para demonstrar o processo de construção de modelos e de previsão dos mínimos quadrados com uma série temporal trimestral, retorne aos dados das receitas da Wal-Mart Stores, Inc. (em bilhões de dólares), originalmente ilustrados na Tabela 16.3. Os dados correspondem a cada um dos trimestres, desde o primeiro trimestre de 2004 até o último trimestre de 2009. A Figura 16.21 ilustra a planilha com os resultados da regressão para o modelo de tendência exponencial trimestral.

FIGURA 16.21
Planilha com os resultados da regressão para ajustes e previsões com os dados de receitas trimestrais da Wal-Mart Stores, Inc.

Crie essa planilha utilizando as instruções na Seção GE14.1.

	A	B	C	D	E	F	G
1	Modelo de Regressão para as Receitas Trimestrais da Wal-Mart Stores						
2							
3	*Estatística de Regressão*						
4	R Múltiplo	0,9943					
5	R-quadrado	0,9887					
6	R-quadrado ajustado	0,9863					
7	Erro-padrão	0,0093					
8	Observações	24					
9							
10	ANOVA						
11		*gl*	*SQ*	*MQ*	*F*	*F de significação*	
12	Regressão	4	0,1432	0,0358	414,6073	0,0000	
13	Resíduo	19	0,0016	0,0001			
14	Total	23	0,1449				
15							
16		*Coeficientes*	*Erro-padrão*	*Stat t*	*Valor-p*	*95% inferiores*	*95% superiores*
17	Interseção	1,8339	0,0052	350,1573	0,0000	1,8229	1,8448
18	Trimestre Codificado	0,0101	0,0003	36,5301	0,0000	0,0096	0,0107
19	Q1	-0,0626	0,0054	-11,5210	0,0000	-0,0739	-0,0512
20	Q2	-0,0401	0,0054	-7,4379	0,0000	-0,0514	-0,0288
21	Q3	-0,0578	0,0054	-10,7502	0,0000	-0,0690	-0,0465

Com base na Figura 16.21, o modelo ajusta os dados extremamente bem. O coeficiente de determinação $r^2 = 0{,}9887$ e o r^2 ajustado $= 0{,}9863$, bem como os resultados do teste F geral, resultam em uma estatística F de 414,6073 (valor-p = 0,000). Observando mais atentamente, no nível de significância de 0,05, cada um dos coeficientes de regressão é extremamente significativo em termos estatísticos e contribui para o modelo. Aplicando os antilogaritmos para todos os coeficientes da regressão, você obtém o seguinte resumo:

Coeficiente de Regressão	$b_i = \log \hat{\beta}_i$	$\hat{\beta}_i = \text{antilog}(b_i) = 10^{b_i}$
b_0: intercepto de Y	1,8339	68,2182
b_1: trimestre codificado	0,0101	1,0235
b_2: primeiro trimestre	−0,0626	0,8658
b_3: segundo trimestre	−0,0401	0,9118
b_4: terceiro trimestre	−0,0578	0,8754

As interpretações para $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ e $\hat{\beta}_4$ são as seguintes:

- O intercepto de Y, $\hat{\beta}_0$, = 68,2182 (em bilhões de dólares), corresponde à previsão *não ajustada* para as receitas trimestrais no primeiro trimestre de 2004, o trimestre inicial na série temporal. *Não ajustada* significa que o componente sazonal não está incorporado na previsão.
- O valor $(\hat{\beta}_1 - 1) \times 100\% = 0{,}0235$, ou 2,35%, corresponde à *taxa de crescimento trimestral composta* estimada das receitas, depois do ajuste em relação ao componente sazonal.

- $\hat{\beta}_2 = 0{,}8658$ corresponde ao multiplicador sazonal para o primeiro trimestre em relação ao quarto trimestre; ele indica que existem 13,42% a menos, em termos de receitas no primeiro trimestre, em comparação com o quarto trimestre.
- $\hat{\beta}_3 = 0{,}9118$ corresponde ao multiplicador sazonal para o segundo trimestre em relação ao quarto trimestre; ele indica que existem 8,82% a menos, em termos de receitas no segundo trimestre, em comparação com o quarto trimestre.
- $\hat{\beta}_4 = 0{,}8754$ corresponde ao multiplicador sazonal para o terceiro trimestre em relação ao quarto trimestre; ele indica que existem 12,46% a menos, em termos de receitas no terceiro trimestre, em comparação com o quarto trimestre. Por conseguinte, o quarto trimestre, que inclui a temporada de compras de final de ano, apresenta o maior volume de vendas.

Utilizando os coeficientes de regressão, b_0, b_1, b_2, b_3 e b_4, e a Equação (16.17), você consegue fazer previsões para trimestres selecionados. Como um exemplo, para prever receitas para o quarto trimestre de 2009, ($X_i = 23$):

$$\log(\hat{Y}_i) = b_0 + b_1X_i$$
$$= 1{,}8339 + (0{,}0101)(23)$$
$$= 2{,}0662$$

Por conseguinte,

$$\log(\hat{Y}_i) = 10^{2{,}0662} = 116{,}4662$$

As receitas previstas para o quarto trimestre do ano fiscal de 2009 correspondem a \$116,4662 bilhões. Para realizar um prognóstico em relação a um período de tempo futuro, tal como o primeiro trimestre do ano fiscal de 2010 ($X_i = 24$, $Q_1 = 1$):

$$\log(\hat{Y}_i) = b_0 + b_1X_i + b_2Q_1$$
$$= 1{,}8339 + (0{,}0101)(24) + (-0{,}0626)(1)$$
$$= 2{,}0137$$

Consequentemente,

$$\hat{Y}_i = 10^{2{,}0137} = 103{,}2048$$

As receitas previstas para o primeiro trimestre do ano fiscal de 2010 totalizam \$103,2048 bilhões.

Problemas para a Seção 16.7

APRENDENDO O BÁSICO

16.40 Ao realizar uma previsão para uma série temporal mensal ao longo de um período de cinco anos, de janeiro de 2006 a dezembro de 2009, a equação de previsão de tendência exponencial para janeiro é

$$\log \hat{Y}_i = 2{,}0 + 0{,}01X_i + 0{,}10 \text{ (Janeiro)}$$

Aplique o antilogaritmo para o coeficiente apropriado, a partir dessa equação, e interprete

a. o intercepto de Y, $\hat{b}_0$.
b. a taxa de crescimento mensal composta.
c. o multiplicador para janeiro.

16.41 Ao se prever dados diários de séries temporais, quantas variáveis binárias (*dummy*) são necessárias para que se represente o componente sazonal dia da semana?

16.42 Ao se prever uma série temporal trimestral ao longo do período de cinco anos, desde o primeiro trimestre de 2005 até o quarto trimestre de 2009, a equação de previsão de tendência exponencial é fornecida por

$$\log \hat{Y}_i = 3{,}0 + 0{,}10X_i - 0{,}25Q_1 + 0{,}20Q_2 + 0{,}15Q_3$$

em que o trimestre zero corresponde ao primeiro trimestre de 2005. Aplique o antilogaritmo do coeficiente apropriado, a partir dessa equação, e interprete

a. o intercepto de Y, $\hat{b}_0$.
b. a taxa de crescimento trimestral composta.
c. o multiplicador para o segundo trimestre.

16.43 Reporte-se ao modelo exponencial apresentado no Problema 16.42.

a. Qual é o valor ajustado da série no quarto trimestre de 2007?
b. Qual é o valor ajustado da série no primeiro trimestre de 2007?
c. Qual é a previsão para o quarto trimestre de 2009?
d. Qual é a previsão para o primeiro trimestre de 2010?

APLICANDO OS CONCEITOS

AUTO-AVALIAÇÃO **16.44** Os dados no arquivo **Toys R Us** correspondem às receitas trimestrais (em milhões de dólares) correspondentes à Toys R Us, desde 1996 até 2008.

Fonte: *Dados extraídos de* Standard & Poor's Stock Reports, *novembro de 1995, novembro de 1998 e abril de 2002. New York: McGraw-Hill, Inc.; e Toys R Us Inc.,* ***www.toysrus.com.***

a. Você acredita que as receitas da Toys R Us estão sujeitas a variações sazonais? Explique.
b. Construa um gráfico com os dados. Esse gráfico respalda sua resposta para o item (a)?
c. Desenvolva uma equação para previsão de tendência exponencial com componentes trimestrais.
d. Interprete a taxa de crescimento trimestral composta.
e. Interprete os multiplicadores para os trimestres.
f. Quais são as previsões para todos os quatro trimestres de 2009?

16.45 Os preços da gasolina ficam mais altos durante a alta temporada das férias de verão? Os dados em **Gasolinanp** contêm os preços médios mensais (em dólares por galão) da gasolina sem chumbo, não poluente, nos Estados Unidos, de janeiro de 2003 a fevereiro de 2009.

Fonte: *Dados extraídos de Energy Information Administration, U.S. Department of Energy,* ***www.eia.doe.gov.***

a. Construa um gráfico de séries temporais.
b. Desenvolva uma equação de previsão de tendência exponencial para dados mensais.
c. Interprete a taxa de crescimento mensal composta.
d. Interprete os multiplicadores mensais.
e. Redija um resumo sucinto de suas descobertas.

16.46 O U.S. Bureau of Labor Statistics (Departamento de Estatísticas do Trabalho dos EUA) compila dados correspondentes a uma ampla gama de questões relacionadas à força de trabalho. Os dados em **Desemprego** apresentam as taxas mensais sazonais ajustadas do desemprego civil nos Estados Unidos, de janeiro de 2002 a abril de 2009.

Fonte: *Dados extraídos de Bureau of Labor Statistics, U.S. Department of Labor,* ***www.bls.gov.***

a. Elabore um gráfico com os dados da série temporal.
b. Desenvolva uma equação para previsão de tendência exponencial com componentes mensais.
c. Qual é o valor ajustado em abril de 2009?
d. Quais são as previsões para todos os últimos oito meses de 2009?
e. Interprete a taxa de crescimento mensal composta.
f. Interprete o multiplicador para julho.
g. Dirija-se à biblioteca ou à Internet e localize a taxa real de desemprego para os últimos oito meses de 2009. Discuta.

16.47 Os dados a seguir (dados no arquivo **Crédito**) representam os encargos financeiros mensais de cartões de crédito (em milhões de dólares) para um cartão de crédito popular emitido por um grande banco (cujo nome, a pedido, não é citado):

a. Construa o gráfico de séries temporais.
b. Descreva o padrão mensal que está evidente nos dados.
c. Em termos gerais, você afirmaria que o montante total, em dólares, relativo a encargos cobrados pelo cartão de crédito do banco está aumentando ou diminuindo? Explique.
d. Observe que as quantias gastas em dezembro de 2008 foram superiores a \$63 milhões, mas, em fevereiro de 2009, foram inferiores a \$40 milhões. O total de fevereiro foi próximo daquilo que seria esperado?

	Ano		
Mês	**2007**	**2008**	**2009**
Janeiro	31,9	39,4	45,0
Fevereiro	27,0	36,2	39,6
Março	31,3	40,5	
Abril	31,0	44,6	
Maio	39,4	46,8	
Junho	40,7	44,7	
Julho	42,3	52,2	
Agosto	49,5	54,0	
Setembro	45,0	48,8	
Outubro	50,0	55,8	
Novembro	50,9	58,7	
Dezembro	58,5	63,4	

e. Desenvolva uma equação de previsão de tendência exponencial com componentes mensais.
f. Interprete a taxa de crescimento mensal composta.
g. Interprete o multiplicador para janeiro.
h. Qual é o valor previsto para março de 2009?
i. Qual é o valor previsto para abril de 2009?
j. De que maneira esse tipo de previsão de séries temporais pode beneficiar o banco?

16.48 Os dados em **Índice S&P** representam o Índice S&P Composto para Preços de Ações, registrado ao final de cada trimestre, desde 1998 até o primeiro trimestre de 2009.

Fonte: *Dados extraídos de* ***finance.yahoo.com.***

a. Faça um gráfico com os dados.
b. Desenvolva uma equação para previsão de tendência exponencial com componentes trimestrais.
c. Interprete a taxa de crescimento trimestral composta.
d. Interprete o multiplicador para o primeiro trimestre.
e. Qual é o valor ajustado para o primeiro trimestre de 2009?
f. Quais são as previsões para os três trimestres remanescentes de 2009?
g. As previsões em (f) foram precisas? Explique

16.49 Os dados no arquivo **Ford** correspondem às receitas trimestrais (em milhões de dólares) da Ford Motor Company, de 1996 a 2008.

Fonte: *Dados extraídos de* Standard & Poor's Stock Reports, *novembro de 2000 e abril de 2002. Nova York: McGraw-Hill, Inc., e Ford Motor Company,* ***www.ford.com.***

a. Você acredita que as receitas da Ford Motor Company estão sujeitas a variações sazonais? Explique.
b. Faça um gráfico com os dados. Esse gráfico respalda a sua resposta para (a)?
c. Desenvolva uma equação para previsão de tendência exponencial com componentes trimestrais.
d. Interprete a taxa de crescimento trimestral composta.
e. Interprete os multiplicadores para os trimestres.
f. Quais são as previsões para todos os quatro trimestres de 2009?

16.8 Tópico *Online*: Números-Índice

Números-índice medem o valor de um item (ou grupo de itens) em um ponto específico no tempo, sob a forma de uma percentagem do valor de um item (ou grupo de itens) em algum outro ponto no tempo. Para estudar esse tópico, leia o arquivo com o tópico *online* intitulado **Seção 16.8** que está disponível no site da LTC Editora para este livro. (Veja a Seção D.8 do Apêndice D para aprender a acessar os arquivos com os tópicos *online*.)

PENSE SOBRE ISSO Alertas para o Usuário de Modelos

Ao utilizar um modelo, você deve sempre examinar os pressupostos contidos no modelo e deve sempre deixar claro o modo como circunstâncias novas ou variáveis podem tornar o modelo menos útil. Nenhum modelo consegue eliminar completamente o risco envolvido na tomada de uma decisão.

Está implícito nos modelos de séries temporais desenvolvidos neste capítulo que dados do passado podem ser utilizados para ajudar a prever o futuro. Embora o uso de dados do passado dessa maneira seja uma aplicação legítima de modelos de séries temporais, com frequência bastante alta, uma crise nos mercados financeiros ilustra que o uso de modelos que se baseiam no passado para prever o futuro não se dá com isenção de riscos.

Por exemplo, durante agosto de 2007, muitos fundos alavancados sofreram prejuízos sem precedentes. Aparentemente, muitos administradores de fundos alavancados utilizaram modelos que basearam sua estratégia de investimentos em padrões de negociações ao longo de extensos períodos de tempos. Esses modelos não refletiram — e não poderiam refletir — padrões de negociações contrários a padrões históricos (G. Morgenson, "A Week When Risk Came Home to Roost", *The New York Times,* 12 de agosto de 2007, pp. B1, B7). Quando administradores de fundos, no início de agosto de 2007, precisaram vender ações em razão de prejuízos em suas carteiras de renda fixa, ações que anteriormente eram mais fortes se tornaram mais fracas e ações mais fracas se tornaram mais fortes — o inverso do esperado com base nos modelos. O que piorou a questão foi o fato de que muitos administradores de fundos estavam utilizando modelos similares e de modo inflexível tomaram decisões sobre investimentos exclusivamente com base no que prescreviam os modelos. Essas ações semelhantes multiplicaram o efeito da pressão de venda, um efeito que os modelos não haviam considerado e que, portanto, não poderia ser verificado nos resultados dos modelos.

Esse exemplo ilustra que o uso de modelo não exime seu usuário da responsabilidade de ser um tomador de decisão criterioso. Não hesite em utilizar modelos — quando utilizados de modo apropriado, eles reforçarão sua tomada de decisão —, mas não faça uso deles indiscriminadamente, porque, nas palavras de um famoso anúncio de serviços públicos, "uma mente é uma coisa terrível de se desperdiçar".

UTILIZANDO A ESTATÍSTICA @ The Principled Revisitada

No cenário Utilizando a Estatística, você era o analista financeiro da The Principled, uma empresa de grande porte que presta serviços financeiros. Você precisava fazer previsões para receitas da Cabot Corporation, da Coca-Cola Company e da Wal-Mart Stores, Inc. para poder avaliar melhor as oportunidades de investimentos de seus clientes. No que diz respeito à Cabot Corporation, você utilizou os métodos de médias móveis e de ajuste exponencial para desenvolver prognósticos. Você previu que a receita da Cabot Corporation, durante 2009, seria de $2.419,1 milhões.

Quanto à Coca-Cola Company, você utilizou o método dos mínimos quadrados, os modelos linear, quadrático e exponencial e autorregressivo para desenvolver prognósticos. Você avaliou esses modelos alternativos e determinou que o modelo autorregressivo de primeira ordem proporcionou a melhor previsão de acordo com vários critérios. Você previu que a receita da Coca-Cola Company seria de $36,5471 bilhões em 2009 e de $42,7217 bilhões em 2010.

No tocante à Wal-Mart Stores, Inc., você utilizou um modelo de regressão dos mínimos quadrados com componentes sazonais, para desenvolver prognósticos. Você previu que a Wal-Mart Stores teria receitas da ordem de $103,2048 bilhões no primeiro trimestre do ano fiscal de 2010.

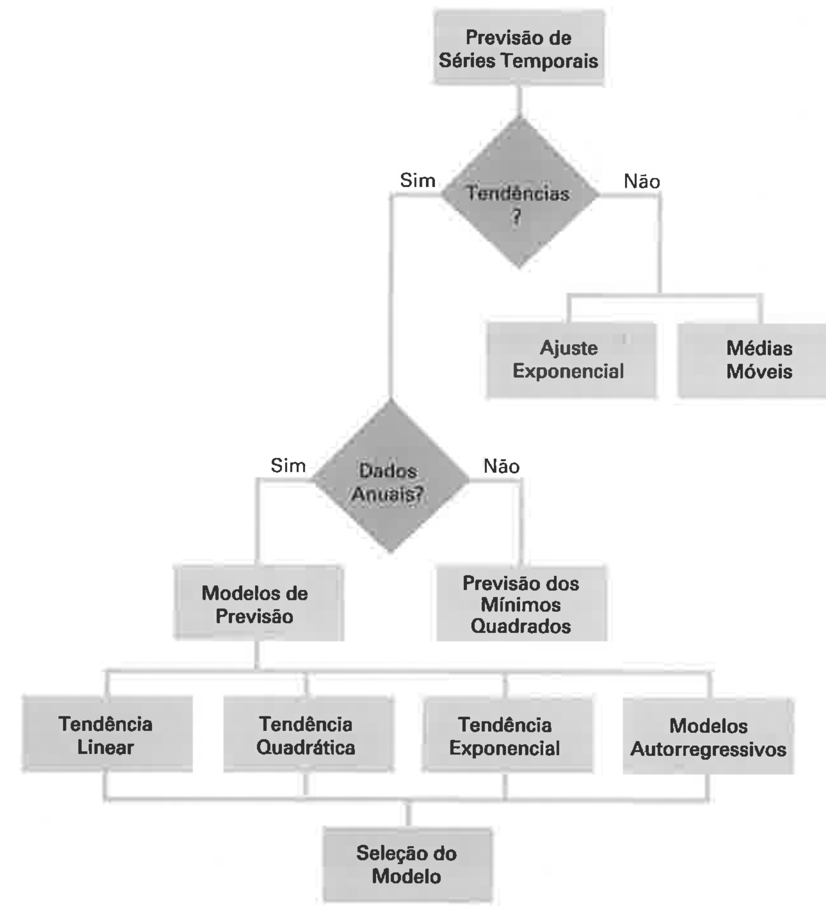
Dadas essas previsões em relação a essas três empresas, você precisa agora determinar se seus clientes deveriam investir e, em caso afirmativo, o montante que eles deveriam investir nessas empresas.

RESUMO

Neste capítulo, você utilizou métodos de séries temporais para desenvolver prognósticos para a Cabot Corporation, a Coca-Cola Company e a Wal-Mart Stores, Inc. Você estudou técnicas de ajuste, o ajuste de tendências dos mínimos quadrados, modelos autorregressivos e a previsão de dados sazonais. A Figura 16.22 fornece um gráfico de resumo para os métodos de séries temporais discutidos neste capítulo.

Ao utilizar previsões para séries temporais, você precisa elaborar um gráfico com a série temporal e responder à seguinte pergunta: Existe uma tendência nos dados? Caso exista uma tendência, você pode então utilizar o modelo autorregressivo ou os modelos de tendência linear, quadrática ou exponencial. Caso não exista nenhuma tendência evidente no gráfico de séries temporais, você deve então utilizar médias móveis ou o ajuste exponencial para amenizar as consequências dos efeitos aleatórios e possíveis efeitos cíclicos. Depois de ajustar os dados, caso ainda não se faça presente uma tendência, você pode utilizar então o ajuste exponencial para prever valores futuros. Caso o ajuste nos dados revele uma tendência, você pode então utilizar o modelo autorregressivo, ou os modelos de tendência linear, quadrática ou exponencial.

FIGURA 16.22
Gráfico de resumo dos métodos de previsão de séries temporais.



EQUAÇÕES-CHAVE

Calculando um Valor Exponencialmente Ajustado no Período de Tempo *i*

$$E_1 = Y_1 \quad (16.1)$$

$$E_i = WY_i + (1 - W)E_{i-1} \quad i = 2, 3, 4, \ldots$$

Fazendo a Previsão para o Período de Tempo *i* + 1

$$\hat{Y}_{i+1} = E_i \quad (16.2)$$

Equação para Previsão de Tendência Linear

$$\hat{Y}_i = b_0 + b_1X_i \quad (16.3)$$

Equação para Previsão de Tendência Quadrática

$$\hat{Y}_i = b_0 + b_1X_i + b_2X_i^2 \quad (16.4)$$

Modelo de Tendência Exponencial

$$Y_i = \beta_0\beta_1^{X_i}\varepsilon_i \quad (16.5)$$

Modelo de Tendência Exponencial Transformado

$$\begin{aligned}\log(Y_i) &= \log(\beta_0\beta_1^{X_i}\varepsilon_i) \\ &= \log(\beta_0) + \log(\beta_1^{X_i}) + \log(\varepsilon_i) \\ &= \log(\beta_0) + X_i\log(\beta_1) + \log(\varepsilon_i)\end{aligned} \quad (16.6)$$

Equação para Previsão de Tendência Exponencial

$$\log(\hat{Y}_i) = b_0 + b_1X_i \quad (16.7a)$$

$$\hat{Y}_i = \hat{\beta}_0\,\hat{\beta}_1^{X_i} \quad (16.7b)$$

Modelo Autorregressivo de Primeira Ordem

$$Y_i = A_0 + A_1Y_{i-1} + \delta_i \quad (16.8)$$

Modelo Autorregressivo de Segunda Ordem

$$Y_i = A_0 + A_1Y_{i-1} + A_2Y_{i-2} + \delta_i \quad (16.9)$$

Modelo Autorregressivo de *p*-ésima Ordem

$$Y_i = A_0 + A_1Y_{i-1} + A_2Y_{i-2} + \cdots + A_pY_{i-p} + \delta_i \quad (16.10)$$

Teste *t* para a Significância do Parâmetro Autorregressivo de Ordem Mais Elevada, A_p

$$t_{ESTAT} = \frac{a_p - A_p}{S_{a_p}} \quad (16.11)$$

Equação Autorregressiva de *p*-ésima Ordem Ajustada

$$\hat{Y}_i = a_0 + a_1Y_{i-1} + a_2Y_{i-2} + \cdots + a_pY_{i-p} \quad (16.12)$$

Equação para Previsão Autorregressiva de *p*-ésima Ordem

$$\hat{Y}_{n+j} = a_0 + a_1\hat{Y}_{n+j-1} + a_2\hat{Y}_{n+j-2} + \cdots + a_p\hat{Y}_{n+j-p} \quad (16.13)$$

Desvio Médio Absoluto

$$DMA = \frac{\sum_{i=1}^{n} |Y_i - \hat{Y}_i|}{n} \quad (16.14)$$

Modelo Exponencial com Dados Trimestrais

$$Y_i = \beta_0\beta_1^{X_i}\beta_2^{Q_1}\beta_3^{Q_2}\beta_4^{Q_3}\varepsilon_i \quad (16.15)$$

Modelo Exponencial Transformado com Dados Trimestrais

$$\begin{aligned}\log(Y_i) &= \log(\beta_0\beta_1^{X_i}\beta_2^{Q_1}\beta_3^{Q_2}\beta_4^{Q_3}\varepsilon_i) \\ &= \log(\beta_0) + \log(\beta_1^{X_i}) + \log(\beta_2^{Q_1}) + \log(\beta_3^{Q_2}) \\ &\quad + \log(\beta_4^{Q_3}) + \log(\varepsilon_i) \\ &= \log(\beta_0) + X_i\log(\beta_1) + Q_1\log(\beta_2) \\ &\quad + Q_2\log(\beta_3) + Q_3\log(\beta_4) + \log(\varepsilon_i)\end{aligned} \quad (16.16)$$

Equação para Previsão de Crescimento Exponencial com Dados Trimestrais

$$\log(\hat{Y}_i) = b_0 + b_1X_i + b_2Q_1 + b_3Q_2 + b_4Q_3 \quad (16.17)$$

Modelo Exponencial com Dados Mensais

$$Y_i = \beta_0\beta_1^{X_i}\beta_2^{M_1}\beta_3^{M_2}\beta_4^{M_3}\beta_5^{M_4}\beta_6^{M_5}\beta_7^{M_6}\beta_8^{M_7}\beta_9^{M_8}\beta_{10}^{M_9}\beta_{11}^{M_{10}}\beta_{12}^{M_{11}}\varepsilon_i \quad (16.18)$$

Modelo Exponencial Transformado com Dados Mensais

$$\begin{aligned}\log(Y_i) &= \log(\beta_0\beta_1^{X_i}\beta_2^{M_1}\beta_3^{M_2}\beta_4^{M_3}\beta_5^{M_4}\beta_6^{M_5}\beta_7^{M_6}\beta_8^{M_7}\beta_9^{M_8}\beta_{10}^{M_9}\beta_{11}^{M_{10}}\beta_{12}^{M_{11}}\varepsilon_i) \\ &= \log(\beta_0) + X_i\log(\beta_1) + M_1\log(\beta_2) + M_2\log(\beta_3) \\ &\quad + M_3\log(\beta_4) + M_4\log(\beta_5) + M_5\log(\beta_6) + M_6\log(\beta_7) \\ &\quad + M_7\log(\beta_8) + M_8\log(\beta_9) + M_9\log(\beta_{10}) \\ &\quad + M_{10}\log(\beta_{11}) + M_{11}\log(\beta_{12}) + \log(\varepsilon_i)\end{aligned} \quad (16.19)$$

Equação para Previsão de Crescimento Exponencial com Dados Mensais

$$\begin{aligned}\log(\hat{Y}_i) &= b_0 + b_1X_i + b_2M_1 + b_3M_2 + b_4M_3 + b_5M_4 + b_6M_5 \\ &\quad + b_7M_6 + b_8M_7 + b_9M_8 + b_{10}M_9 + b_{11}M_{10} + b_{12}M_{11}\end{aligned} \quad (16.20)$$

TERMOS-CHAVE

ajuste exponencial
autocorrelação de *p*-ésima ordem
autocorrelação de primeira ordem
autocorrelação de segunda ordem
desvio médio absoluto (*DMA*)
efeito aleatório
efeito cíclico
efeito irregular
efeito sazonal
médias móveis
método causal de previsão
métodos de previsão de séries temporais
métodos de previsão qualitativos
métodos de previsão quantitativos
modelo autorregressivo
modelo autorregressivo de *p*-ésima ordem
modelo autorregressivo de primeira ordem
modelo autorregressivo de segunda ordem
modelo de tendência exponencial
modelo de tendência linear
modelo de tendência quadrática
parcimônia
previsão
séries temporais
tendência

PROBLEMAS DE REVISÃO DO CAPÍTULO 16

AVALIANDO O SEU ENTENDIMENTO

16.50 O que significa uma série temporal?

16.51 Quais são os diferentes componentes de um modelo de séries temporais?

16.52 Qual é a diferença entre médias móveis e ajuste exponencial?

16.53 Sob quais circunstâncias o modelo de tendência exponencial é o mais apropriado?

16.54 De que modo o modelo de previsão de tendência linear dos mínimos quadrados desenvolvido neste capítulo difere do modelo de regressão linear dos mínimos quadrados considerado no Capítulo 13?

16.55 De que modo a modelagem autorregressiva difere dos outros métodos de previsão?

16.56 Quais são os diferentes métodos para se escolher um modelo de previsão apropriado?

16.57 Qual é a principal diferença entre a utilização de S_{YX} e do *DMA* para avaliar a qualidade do ajuste de um determinado modelo aos dados?

16.58 De que modo a previsão para dados mensais ou trimestrais difere da previsão para dados anuais?

APLICANDO OS CONCEITOS

16.59 A tabela a seguir (contida no arquivo **Pólio**) representa as taxas de incidência anuais (para cada 100.000 pessoas) para casos informados de poliomielite aguda, registrados ao longo de períodos de cinco anos, de 1915 a 1955:

Ano	1915	1920	1925	1930	1935	1940	1945	1950	1955
Taxa	3,1	2,2	5,3	7,5	8,5	7,4	10,3	22,1	17,6

Fonte: *Dados extraídos de B. Wattenberg, ed.*, The Statistical History of the United States: From Colonial Times to the Present, *ser. B303 (New York: Basic Books, 1976).*

a. Faça um gráfico com os dados.
b. Desenvolva a equação para previsão de tendência linear e faça um gráfico para a linha de tendência.
c. Quais são as suas previsões para 1960, 1965 e 1970?
d. Utilizando uma biblioteca ou a Internet, encontre as taxas de incidência de poliomielite aguda efetivamente registradas para 1960, 1965 e 1970. Registre os seus resultados.
e. Por que razão as previsões feitas por você em (c) não são úteis? Discuta.

16.60 O U.S. Department of Labor coleta e publica estatísticas relacionadas ao mercado de trabalho. O arquivo **Força de Trabalho** contém a população civil não institucionalizada dos EUA de pessoas com 16 anos de idade ou mais (em milhares) e a força de trabalho civil não institucionalizada dos EUA de pessoas com 16 anos ou mais (em milhares) para 1984-2008. A variável força de trabalho relata o número de pessoas na população que têm algum emprego ou que estão efetivamente buscando um emprego.

Fonte: *Dados extraídos de Bureau of Labor Statistics, U.S. Department of Labor, www.bls.gov.*

a. Elabore um gráfico de séries temporais para a população civil não institucionalizada dos EUA de pessoas com 16 anos de idade ou mais.
b. Desenvolva a equação para previsão de tendência linear.
c. Faça a previsão para a população civil não institucionalizada dos EUA de pessoas com 16 anos de idade ou mais para 2009 e 2010.
d. Repita (a) a (c) para a força de trabalho civil não institucionalizada dos EUA de pessoas com 16 anos de idade ou mais.

16.61 O preço do gás natural (dólares por 40 unidades de calor) cobrado trimestralmente nos Estados Unidos, de 1994 a 2008, é apresentado no arquivo **Gás Natural**.

Fonte: *Dados extraídos de Energy Information Administration, U.S. Department of Energy, www eia.gov.*

a. Você acredita que o preço do gás natural apresenta um componente sazonal?
b. Elabore um gráfico para a série temporal. Esse gráfico respalda a sua resposta em (a)?
c. Desenvolva a equação para previsão de tendência exponencial para dados trimestrais.
d. Interprete a taxa de crescimento trimestral composta.
e. Interprete os multiplicadores trimestrais. Os multiplicadores respaldam suas respostas para (a) e (b)?

16.62 Os dados na tabela a seguir (armazenados em **McDonalds**) representam as receitas brutas (em bilhões de dólares correntes) da McDonald's Corporation ao longo do período de 31 anos, de 1975 a 2005:

Ano	Receitas	Ano	Receitas	Ano	Receitas
1975	1,0	1987	4,9	1999	13,3
1976	1,2	1988	5,6	2000	14,2
1977	1,4	1989	6,1	2001	14,8
1978	1,7	1990	6,8	2002	15,2
1979	1,9	1991	6,7	2003	16,8
1980	2,2	1992	7,1	2004	18,6
1981	2,5	1993	7,4	2005	19,8
1982	2,8	1994	8,3	2006	20,9
1983	3,1	1995	9,8	2007	22,8
1984	3,4	1996	10,7	2008	23,5
1985	3,8	1997	11,4		
1986	4,2	1998	12,4		

Fonte: *Dados extraídos de* Moody's Handbook of Common Stocks, *1980, 1989 e 1999*; Mergent's Handbook of Common Stocks, *Spring 2002; e* www.mcdonalds.com.

a. Elabore um gráfico com os dados.
b. Desenvolva a equação para previsão de tendência linear.
c. Desenvolva a equação para previsão de tendência quadrática.
d. Desenvolva a equação para previsão de tendência exponencial.
e. Encontre o modelo autorregressivo com o melhor ajuste, utilizando $\alpha = 0{,}05$.
f. Realize uma análise de resíduos para cada um dos modelos em (b) até (e).
g. Calcule o erro-padrão da estimativa (S_{YX}) e o *DMA* para cada um dos modelos correspondentes em (f).

h. Com base em seus resultados em (f) e (g), juntamente com uma avaliação sobre o princípio da parcimônia, qual modelo você selecionaria para fins de previsão? Discuta.
i. Utilizando o modelo selecionado em (h), faça uma previsão das receitas brutas para 2009.

16.63 O U.S. Census Bureau (Departamento de Censo dos EUA) acompanha o número de novas moradias para as quais as construções têm início a cada trimestre. Essa unidade de medida métrica para a atividade econômica é geralmente conhecida como "início de construção de moradias". O arquivo **Moradia** contém dados trimestrais de 2000 a 2008.

a. Você acredita que a construção de novas moradias é sazonal? Por que sim ou por que não?
b. Elabore um gráfico com os dados.
c. O seu gráfico em (b) respalda a sua resposta em (a)?
d. Desenvolva uma equação para previsão de tendência exponencial com dados trimestrais.
e. Com base em (d), qual trimestre apresenta a maior quantidade de inícios de construção de moradias? E o menor número de inícios de construção de moradias?
f. Existe uma tendência de longo prazo de 2000 a 2008, em termos do número de novas moradias sendo construídas?

16.64 O Sistema de Aposentadoria de Professores da Cidade de Nova York oferece a seus membros vários tipos de investimentos. Entre as opções estão investimentos com taxas de retorno fixas e com taxas de retorno variáveis. Existem atualmente duas categorias de investimentos com retornos variáveis. A variável *A* consiste em investimentos que são primordialmente compostos por ações, enquanto a variável *B* consiste em investimentos em títulos de empresas e outros tipos de instrumentos de mais baixo risco. Os dados (armazenados no arquivo **TRSNYC**) representam o valor de uma unidade de cada tipo de investimento com retorno variável ao início de cada ano, de 1984 a 2009:

Fonte: *Dados extraídos de Teachers' Retirement System of the City of New York, www.trs.nyc.ny.us.*

Para cada uma das duas séries temporais,

a. Elabore um gráfico com os dados.
b. Desenvolva a equação para previsão de tendência linear.
c. Desenvolva a equação para previsão de tendência quadrática.
d. Desenvolva a equação para previsão de tendência exponencial.
e. Encontre o modelo autorregressivo com o melhor ajuste, utilizando $\alpha = 0{,}05$.
f. Realize uma análise de resíduos para cada um dos modelos em (b) até (e).
g. Calcule o erro-padrão da estimativa (S_{YX}) e o *DMA* para cada modelo correspondente em (f).
h. Com base em seus resultados em (f) e (g), juntamente com uma avaliação sobre o princípio da parcimônia, qual modelo você selecionaria para fins de previsão? Discuta.
i. Utilizando o modelo selecionado em (h), faça uma previsão dos valores unitários para 2010.
j. Com base nos resultados de (a) a (i), qual estratégia de investimento você recomendaria a um membro do Sistema de Aposentadoria dos Professores da Cidade de Nova York? Explique.

EXERCÍCIOS DE REDAÇÃO DE RELATÓRIOS

16.65 Como consultor de uma empresa de investimentos que realiza transações em várias moedas correntes, você recebeu a tarefa de estudar as tendências de longo prazo nas taxas de câmbio do dólar canadense, do iene japonês e da libra inglesa. Os dados de 1980 a 2008 estão contidos no arquivo **Moeda**, onde o dólar canadense, o iene japonês e a libra inglesa estão expressos em unidades de dólares norte-americanos.

Desenvolva um modelo de previsão para a taxa de câmbio de cada uma dessas três moedas correntes e forneça previsões para 2009 e 2010 para cada uma delas. Escreva um resumo executivo para uma apresentação a ser feita junto à empresa de investimentos. Anexe a esse resumo executivo uma discussão sobre possíveis limitações que possam existir nesses modelos.

ADMINISTRANDO O SPRINGVILLE HERALD

Como parte da iniciativa estratégica continuada de fazer crescer o número de assinaturas com entrega domiciliar, o departamento de circulação está monitorando de perto o número de assinaturas nessa modalidade. O departamento de circulação deseja prever futuras assinaturas com entrega domiciliar. Para realizar essa tarefa, o departamento de circulação compilou o número de assinaturas com entrega domiciliar para o período mais recente de 24 meses e armazenou os dados no arquivo **SH16**.

EXERCÍCIO

SH16.1 a. Analise esses dados e desenvolva um modelo para prever assinaturas com entrega domiciliar. Apresente suas descobertas em um relatório que inclua os pressupostos do modelo e suas respectivas limitações. Faça a previsão de assinaturas com entrega domiciliar para os próximos 4 meses.

b. Você estaria disposto a utilizar o modelo desenvolvido para prever assinaturas com entrega domiciliar um ano à frente no futuro? Explique.

c. Compare a tendência de assinaturas com entrega domiciliar com o número de novas assinaturas por mês, fornecido no arquivo **SH13**. Que explicação você pode dar quanto a eventuais diferenças?

CASO DE INTERNET

Aplique seus conhecimentos sobre previsões de séries temporais neste Caso de Internet, que representa uma extensão do caso em curso "Administrando o Springville Herald", *que aparece em capítulos selecionados deste livro.*

O *Springville Herald* está disputando novos leitores na área de TriCities com o mais novo concorrente, o *Oxford Glen Journal (OGJ)*. Recentemente, a equipe de circulação do *OGJ* declarou que a circulação e a base de assinaturas de seu jornal estão crescendo mais rapidamente do que no *Herald* e que os anunciantes locais fariam melhor ao transferir seus anúncios e propagandas do *Herald* para o *OGJ*. O departamento de circulação do *Herald* entrou com uma reclamação junto à Câmara de Comércio de Springville sobre as declarações feitas pelo *OGJ* e solicitou que a Câmara fizesse investigações, solicitação que foi bem aceita pela equipe de circulação do *OGJ*.

Utilizando um navegador na Web, acesse o site da LTC Editora para este livro e o Caso de Internet para o Capítulo 16 (ou abra diretamente o arquivo **SCC_CirculationDispute.htm** caso já tenha baixado para seu computador os arquivo com os Casos de Internet) para examinar as informações sobre a disputa entre os departamentos de circulação, coletadas pela Câmara de Comércio de Springville. Depois disso, responda ao seguinte:

1. Qual dos jornais você diria que tem o direito de declarar ter a circulação e a base de assinaturas com crescimento mais rápido? Respalde sua resposta realizando e sintetizando uma análise estatística apropriada.
2. Qual é o fato singular mais positivo em relação à circulação e à base de assinaturas do *Herald*? Qual é o fato singular mais positivo em relação à circulação e à base de assinaturas do *OGJ*? Explique suas respostas.
3. Que tipo de dados adicionais seriam úteis ao investigar as declarações sobre circulação feitas pelas equipes de cada um dos jornais?

REFERÊNCIAS

1. Bowerman, B. L., R. T. O'Connell, and A. Koehler, *Forecasting, Time Series, and Regression*, 4th ed. (Belmont, CA: Duxbury Press, 2005).
2. Box, G. E. P., G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd ed. (Upper Saddle River, NJ: Prentice Hall, 1994).
3. Frees, E. W., *Data Analysis Using Regression Models: The Business Perspective* (Upper Saddle River, NJ: Prentice Hall, 1996).
4. Hanke, J. E., D. W. Wichern, and A. G. Reitsch, *Business Forecasting*, 7th ed. (Upper Saddle River, NJ: Prentice Hall, 2001).
5. *Microsoft Excel 2007* (Redmond, WA: Microsoft Corp., 2007).