

TURING

图灵原版数学·统计学系列 30

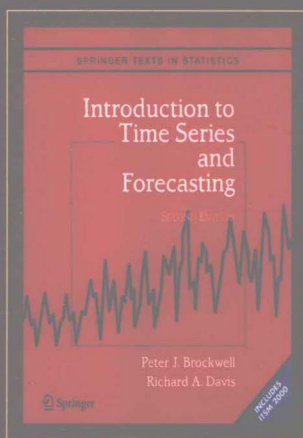
Springer

# Introduction to Time Series and Forecasting

# 时间序列与预测

(英文版·第2版)

[ 美 ] Peter J. Brockwell 著  
Richard A. Davis



人民邮电出版社  
POSTS & TELECOM PRESS



图灵原版数学·统计学系列

**Introduction to Time  
Series and Forecasting**

# 时间序列与预测

(英文版·第2版)

[美] Peter J. Brockwell 著  
Richard A. Davis

人民邮电出版社  
北 京

## 图书在版编目(CIP)数据

时间序列与预测=Introduction to Time Series and Forecasting: 第2版: 英文/(美)布罗克韦尔(Brockwell, P. J.), (美)戴维斯(Davis, R. A.)著. —北京: 人民邮电出版社, 2009.3

(图灵原版数学·统计学系列)

ISBN 978-7-115-19682-8

I. 时… II. ①布… ②戴… III. 时间序列分析—英文  
IV. O211.61

中国版本图书馆CIP数据核字(2009)第011822号

## 内 容 提 要

本书全面介绍了经济学、工程学、自然科学和社会科学中所用的时间序列和预测方法, 核心内容是平稳过程、ARMA过程、ARIMA过程、多变量时间序列、状态空间模型和谱分析. 另外, 本书还介绍了Burg算法、Hannan-Rissanen算法、EM算法、结构模型、指数平滑、转移函数模型、非线性模型、连续时间模型和长记忆模型等. 每章的末尾都有大量习题, 供读者巩固所学概念和方法. 本书强调方法和数据集的分析, 配有时间序列软件包ITSM2000的学生版.

本书适合作为各专业学生时间序列入门课程的教材, 也适合其他有兴趣的科研工作者阅读.

图灵原版数学·统计学系列

## 时间序列与预测(英文版·第2版)

- 
- ◆ 著 [美] Peter J. Brockwell Richard A. Davis  
责任编辑 明永玲
  - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街14号  
邮编 100061 电子函件 315@ptpress.com.cn  
网址 <http://www.ptpress.com.cn>  
北京铭成印刷有限公司印刷
  - ◆ 开本: 700×1000 1/16  
印张: 28.25  
字数: 450千字 2009年3月第1版  
印数: 1-2 000册 2009年3月北京第1次印刷

著作权合同登记号 图字: 01-2008-5816 号

ISBN 978-7-115-19682-8/O1

---

定价: 69.00元

读者服务热线: (010) 88593802 印装质量热线: (010) 67129223

反盗版热线: (010) 67171154

# 版 权 声 明

*Introduction to Time Series and Forecasting, Second Edition* by Peter J. Brockwell, Richard A. Davis, ISBN: 978-0-387-95351-9.

Copyright © 2002, 1996 by Springer Science+Business Media, LLC. All rights reserved.

This reprint has been authorized by Springer-Verlag (Berlin/Heidelberg/New York) for sale in the Mainland China only and not for export therefrom.

本书英文影印版由Springer-Verlag授权人民邮电出版社在中华人民共和国境内独家发行。本版仅限在中华人民共和国境内（不包括香港、澳门特别行政区和台湾地区）出版及标价销售。未经许可之出口，视为违反著作权法，将受法律之制裁。

版权所有，侵权必究。

# Preface

This book is aimed at the reader who wishes to gain a working knowledge of time series and forecasting methods as applied in economics, engineering and the natural and social sciences. Unlike our earlier book, *Time Series: Theory and Methods*, referred to in the text as TSTM, this one requires only a knowledge of basic calculus, matrix algebra and elementary statistics at the level (for example) of Mendenhall, Wackerly and Scheaffer (1990). It is intended for upper-level undergraduate students and beginning graduate students.

The emphasis is on methods and the analysis of data sets. The student version of the time series package ITSM2000, enabling the reader to reproduce most of the calculations in the text (and to analyze further data sets of the reader's own choosing), is included on the CD-ROM<sup>①</sup> which accompanies the book. The data sets used in the book are also included. The package requires an IBM-compatible PC operating under Windows 95, NT version 4.0, or a later version of either of these operating systems. The program ITSM can be run directly from the CD-ROM or installed on a hard disk as described at the beginning of Appendix D, where a detailed introduction to the package is provided.

Very little prior familiarity with computing is required in order to use the computer package. Detailed instructions for its use are found in the on-line help files which are accessed, when the program ITSM is running, by selecting the menu option Help>Contents and selecting the topic of interest. Under the heading Data you will find information concerning the data sets stored on the CD-ROM. The book can also be used in conjunction with other computer packages for handling time series. Chapter 14 of the book by Venables and Ripley (1994) describes how to perform many of the calculations using S-plus.

There are numerous problems at the end of each chapter, many of which involve use of the programs to study the data sets provided.

To make the underlying theory accessible to a wider audience, we have stated some of the key mathematical results without proof, but have attempted to ensure that the logical structure of the development is otherwise complete. (References to proofs are provided for the interested reader.)

---

① 该CD-ROM上的内容可以从图灵网站[www.turingbook.com](http://www.turingbook.com)下载。——编者注

Since the upgrade to ITSM2000 occurred after the first edition of this book appeared, we have taken the opportunity, in this edition, to coordinate the text with the new software, to make a number of corrections pointed out by readers of the first edition and to expand on several of the topics treated only briefly in the first edition.

Appendix D, the software tutorial, has been rewritten in order to be compatible with the new version of the software.

Some of the other extensive changes occur in (i) Section 6.6, which highlights the role of the innovations algorithm in generalized least squares and maximum likelihood estimation of regression models with time series errors, (ii) Section 6.4, where the treatment of forecast functions for ARIMA processes has been expanded and (iii) Section 10.3, which now includes GARCH modeling and simulation, topics of considerable importance in the analysis of financial time series. The new material has been incorporated into the accompanying software, to which we have also added the option *Autofit*. This streamlines the modeling of time series data by fitting maximum likelihood  $\text{ARMA}(p, q)$  models for a specified range of  $(p, q)$  values and automatically selecting the model with smallest AICC value.

There is sufficient material here for a full-year introduction to univariate and multivariate time series and forecasting. Chapters 1 through 6 have been used for several years in introductory one-semester courses in univariate time series at Colorado State University and Royal Melbourne Institute of Technology. The chapter on spectral analysis can be excluded without loss of continuity by readers who are so inclined.

We are greatly indebted to the readers of the first edition and especially to Matthew Calder, coauthor of the new computer package, and Anthony Brockwell for their many valuable comments and suggestions. We also wish to thank Colorado State University, the National Science Foundation, Springer-Verlag and our families for their continuing support during the preparation of this second edition.

Fort Collins, Colorado  
August 2001

Peter J. Brockwell  
Richard A. Davis

# Contents

<b>1. Introduction</b>	<b>1</b>
1.1. Examples of Time Series	1
1.2. Objectives of Time Series Analysis	6
1.3. Some Simple Time Series Models	7
1.3.1. Some Zero-Mean Models	8
1.3.2. Models with Trend and Seasonality	9
1.3.3. A General Approach to Time Series Modeling	14
1.4. Stationary Models and the Autocorrelation Function	15
1.4.1. The Sample Autocorrelation Function	18
1.4.2. A Model for the Lake Huron Data	21
1.5. Estimation and Elimination of Trend and Seasonal Components	23
1.5.1. Estimation and Elimination of Trend in the Absence of Seasonality	24
1.5.2. Estimation and Elimination of Both Trend and Seasonality	31
1.6. Testing the Estimated Noise Sequence Problems	35 40
<b>2. Stationary Processes</b>	<b>45</b>
2.1. Basic Properties	45
2.2. Linear Processes	51
2.3. Introduction to ARMA Processes	55
2.4. Properties of the Sample Mean and Autocorrelation Function	57
2.4.1. Estimation of $\mu$	58
2.4.2. Estimation of $\gamma(\cdot)$ and $\rho(\cdot)$	59
2.5. Forecasting Stationary Time Series	63
2.5.1. The Durbin–Levinson Algorithm	69
2.5.2. The Innovations Algorithm	71
2.5.3. Prediction of a Stationary Process in Terms of Infinitely Many Past Values	75

2.6. The Wold Decomposition	77
Problems	78
<b>3. ARMA Models</b>	<b>83</b>
3.1. ARMA( $p, q$ ) Processes	83
3.2. The ACF and PACF of an ARMA( $p, q$ ) Process	88
3.2.1. Calculation of the ACVF	88
3.2.2. The Autocorrelation Function	94
3.2.3. The Partial Autocorrelation Function	94
3.2.4. Examples	96
3.3. Forecasting ARMA Processes	100
Problems	108
<b>4. Spectral Analysis</b>	<b>111</b>
4.1. Spectral Densities	112
4.2. The Periodogram	121
4.3. Time-Invariant Linear Filters	127
4.4. The Spectral Density of an ARMA Process	132
Problems	134
<b>5. Modeling and Forecasting with ARMA Processes</b>	<b>137</b>
5.1. Preliminary Estimation	138
5.1.1. Yule–Walker Estimation	139
5.1.2. Burg’s Algorithm	147
5.1.3. The Innovations Algorithm	150
5.1.4. The Hannan–Rissanen Algorithm	156
5.2. Maximum Likelihood Estimation	158
5.3. Diagnostic Checking	164
5.3.1. The Graph of $\{\hat{R}_t, t = 1, \dots, n\}$	165
5.3.2. The Sample ACF of the Residuals	166
5.3.3. Tests for Randomness of the Residuals	166
5.4. Forecasting	167
5.5. Order Selection	169
5.5.1. The FPE Criterion	170
5.5.2. The AICC Criterion	171
Problems	174
<b>6. Nonstationary and Seasonal Time Series Models</b>	<b>179</b>
6.1. ARIMA Models for Nonstationary Time Series	180
6.2. Identification Techniques	187



6.3. Unit Roots in Time Series Models	193
6.3.1. Unit Roots in Autoregressions	194
6.3.2. Unit Roots in Moving Averages	196
6.4. Forecasting ARIMA Models	198
6.4.1. The Forecast Function	200
6.5. Seasonal ARIMA Models	203
6.5.1. Forecasting SARIMA Processes	208
6.6. Regression with ARMA Errors	210
6.6.1. OLS and GLS Estimation	210
6.6.2. ML Estimation	213
Problems	219
<b>7. Multivariate Time Series</b>	<b>223</b>
7.1. Examples	224
7.2. Second-Order Properties of Multivariate Time Series	229
7.3. Estimation of the Mean and Covariance Function	234
7.3.1. Estimation of $\mu$	234
7.3.2. Estimation of $\Gamma(h)$	235
7.3.3. Testing for Independence of Two Stationary Time Series	237
7.3.4. Bartlett's Formula	238
7.4. Multivariate ARMA Processes	241
7.4.1. The Covariance Matrix Function of a Causal ARMA Process	244
7.5. Best Linear Predictors of Second-Order Random Vectors	244
7.6. Modeling and Forecasting with Multivariate AR Processes	246
7.6.1. Estimation for Autoregressive Processes Using Whittle's Algorithm	247
7.6.2. Forecasting Multivariate Autoregressive Processes	250
7.7. Cointegration	254
Problems	256
<b>8. State-Space Models</b>	<b>259</b>
8.1. State-Space Representations	260
8.2. The Basic Structural Model	263
8.3. State-Space Representation of ARIMA Models	267
8.4. The Kalman Recursions	271
8.5. Estimation For State-Space Models	277
8.6. State-Space Models with Missing Observations	283
8.7. The EM Algorithm	289
8.8. Generalized State-Space Models	292
8.8.1. Parameter-Driven Models	292

8.8.2. Observation-Driven Models	299
Problems	311
<b>9. Forecasting Techniques</b>	<b>317</b>
9.1. The ARAR Algorithm	318
9.1.1. Memory Shortening	318
9.1.2. Fitting a Subset Autoregression	319
9.1.3. Forecasting	320
9.1.4. Application of the ARAR Algorithm	321
9.2. The Holt–Winters Algorithm	322
9.2.1. The Algorithm	322
9.2.2. Holt–Winters and ARIMA Forecasting	324
9.3. The Holt–Winters Seasonal Algorithm	326
9.3.1. The Algorithm	326
9.3.2. Holt–Winters Seasonal and ARIMA Forecasting	328
9.4. Choosing a Forecasting Algorithm	328
Problems	330
<b>10. Further Topics</b>	<b>331</b>
10.1. Transfer Function Models	331
10.1.1. Prediction Based on a Transfer Function Model	337
10.2. Intervention Analysis	340
10.3. Nonlinear Models	343
10.3.1. Deviations from Linearity	344
10.3.2. Chaotic Deterministic Sequences	345
10.3.3. Distinguishing Between White Noise and iid Sequences	347
10.3.4. Three Useful Classes of Nonlinear Models	348
10.3.5. Modeling Volatility	349
10.4. Continuous-Time Models	357
10.5. Long-Memory Models	361
Problems	365
<b>A. Random Variables and Probability Distributions</b>	<b>369</b>
A.1. Distribution Functions and Expectation	369
A.2. Random Vectors	374
A.3. The Multivariate Normal Distribution	377
Problems	381

<b>B. Statistical Complements</b>	<b>383</b>
B.1. Least Squares Estimation	383
B.1.1. The Gauss-Markov Theorem	385
B.1.2. Generalized Least Squares	386
B.2. Maximum Likelihood Estimation	386
B.2.1. Properties of Maximum Likelihood Estimators	387
B.3. Confidence Intervals	388
B.3.1. Large-Sample Confidence Regions	388
B.4. Hypothesis Testing	389
B.4.1. Error Probabilities	389
B.4.2. Large-Sample Tests Based on Confidence Regions	390
 <b>C. Mean Square Convergence</b>	 <b>393</b>
C.1. The Cauchy Criterion	393
 <b>D. An ITSM Tutorial</b>	 <b>395</b>
D.1. Getting Started	396
D.1.1. Running ITSM	396
D.2. Preparing Your Data for Modeling	396
D.2.1. Entering Data	397
D.2.2. Information	397
D.2.3. Filing Data	397
D.2.4. Plotting Data	398
D.2.5. Transforming Data	398
D.3. Finding a Model for Your Data	403
D.3.1. Autofit	403
D.3.2. The Sample ACF and PACF	403
D.3.3. Entering a Model	404
D.3.4. Preliminary Estimation	406
D.3.5. The AICC Statistic	408
D.3.6. Changing Your Model	408
D.3.7. Maximum Likelihood Estimation	409
D.3.8. Optimization Results	410
D.4. Testing Your Model	411
D.4.1. Plotting the Residuals	412
D.4.2. ACF/PACF of the Residuals	412
D.4.3. Testing for Randomness of the Residuals	414
D.5. Prediction	415
D.5.1. Forecast Criteria	415
D.5.2. Forecast Results	415

D.6. Model Properties	416
D.6.1. ARMA Models	417
D.6.2. Model ACF, PACF	418
D.6.3. Model Representations	419
D.6.4. Generating Realizations of a Random Series	420
D.6.5. Spectral Properties	421
D.7. Multivariate Time Series	421
<b>References</b>	<b>423</b>
<b>Index</b>	<b>429</b>

# 1

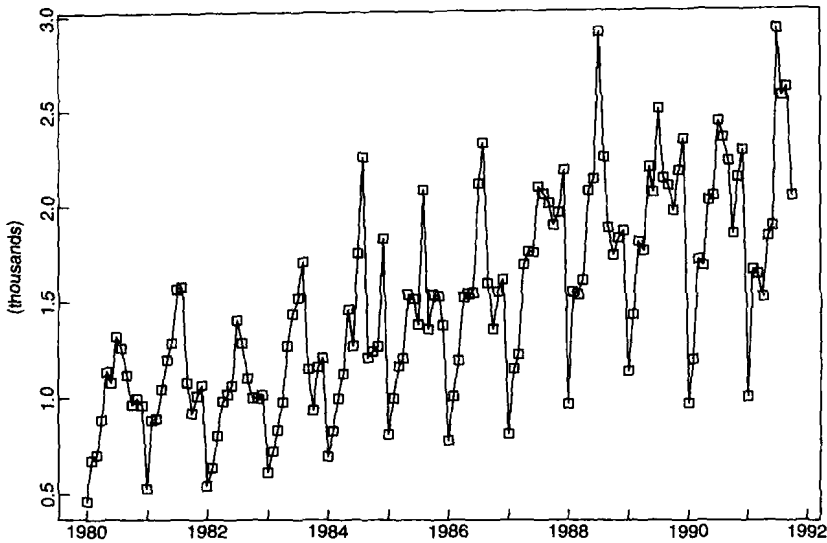
## Introduction

- 1.1 Examples of Time Series
- 1.2 Objectives of Time Series Analysis
- 1.3 Some Simple Time Series Models
- 1.4 Stationary Models and the Autocorrelation Function
- 1.5 Estimation and Elimination of Trend and Seasonal Components
- 1.6 Testing the Estimated Noise Sequence

In this chapter we introduce some basic ideas of time series analysis and stochastic processes. Of particular importance are the concepts of stationarity and the autocovariance and sample autocovariance functions. Some standard techniques are described for the estimation and removal of trend and seasonality (of known period) from an observed time series. These are illustrated with reference to the data sets in Section 1.1. The calculations in all the examples can be carried out using the time series package ITSM, the student version of which is supplied on the enclosed CD. The data sets are contained in files with names ending in .TSM. For example, the Australian red wine sales are filed as WINE.TSM. Most of the topics covered in this chapter will be developed more fully in later sections of the book. The reader who is not already familiar with random variables and random vectors should first read Appendix A, where a concise account of the required background is given.

### 1.1 Examples of Time Series

A **time series** is a set of observations  $x_t$ , each one being recorded at a specific time  $t$ . A *discrete-time time series* (the type to which this book is primarily devoted) is one in which the set  $T_0$  of times at which observations are made is a discrete set, as is the



**Figure 1-1** The Australian red wine sales, Jan. '80 – Oct. '91.

case, for example, when observations are made at fixed time intervals. *Continuous-time time series* are obtained when observations are recorded continuously over some time interval, e.g., when  $T_0 = [0, 1]$ .

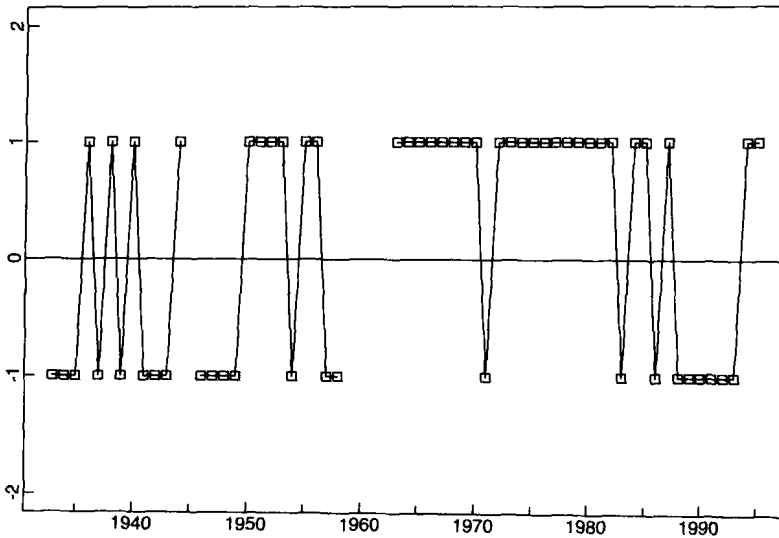
### Example 1.1.1 Australian red wine sales; WINE.TSM

Figure 1.1 shows the monthly sales (in kiloliters) of red wine by Australian winemakers from January 1980 through October 1991. In this case the set  $T_0$  consists of the 142 times  $\{(\text{Jan. 1980}), (\text{Feb. 1980}), \dots, (\text{Oct. 1991})\}$ . Given a set of  $n$  observations made at uniformly spaced time intervals, it is often convenient to rescale the time axis in such a way that  $T_0$  becomes the set of integers  $\{1, 2, \dots, n\}$ . In the present example this amounts to measuring time in months with (Jan. 1980) as month 1. Then  $T_0$  is the set  $\{1, 2, \dots, 142\}$ . It appears from the graph that the sales have an upward trend and a seasonal pattern with a peak in July and a trough in January. To plot the data using ITSM, run the program by double-clicking on the ITSM icon and then select the option `File>Project>Open>Univariate`, click OK, and select the file WINE.TSM. The graph of the data will then appear on your screen.  $\square$

### Example 1.1.2 All-star baseball games, 1933–1995

Figure 1.2 shows the results of the all-star games by plotting  $x_t$ , where

$$x_t = \begin{cases} 1 & \text{if the National League won in year } t, \\ -1 & \text{if the American League won in year } t. \end{cases}$$



**Figure 1-2** Results of the all-star baseball games, 1933–1995.

This is a series with only two possible values,  $\pm 1$ . It also has some missing values, since no game was played in 1945, and two games were scheduled for each of the years 1959–1962.  $\square$

**Example 1.1.3** Accidental deaths, U.S.A., 1973–1978; DEATHS.TSM

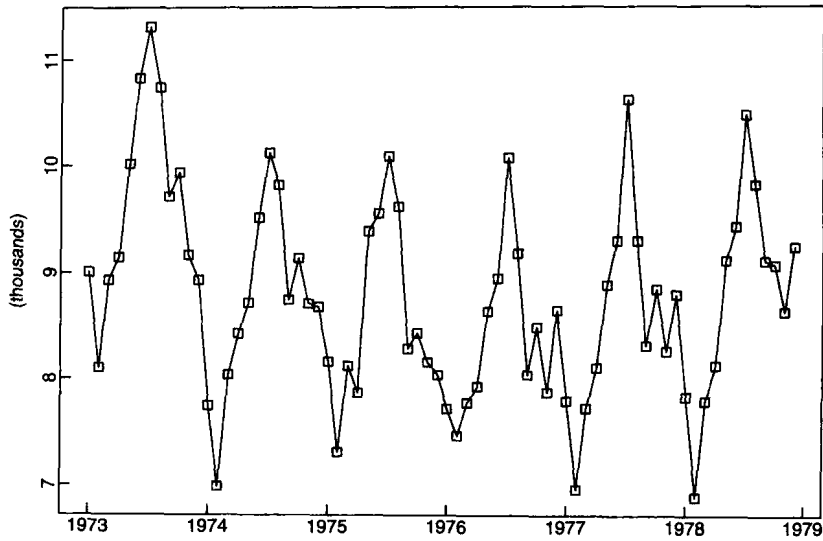
Like the red wine sales, the monthly accidental death figures show a strong seasonal pattern, with the maximum for each year occurring in July and the minimum for each year occurring in February. The presence of a trend in Figure 1.3 is much less apparent than in the wine sales. In Section 1.5 we shall consider the problem of representing the data as the sum of a trend, a seasonal component, and a residual term.  $\square$

**Example 1.1.4** A signal detection problem; SIGNAL.TSM

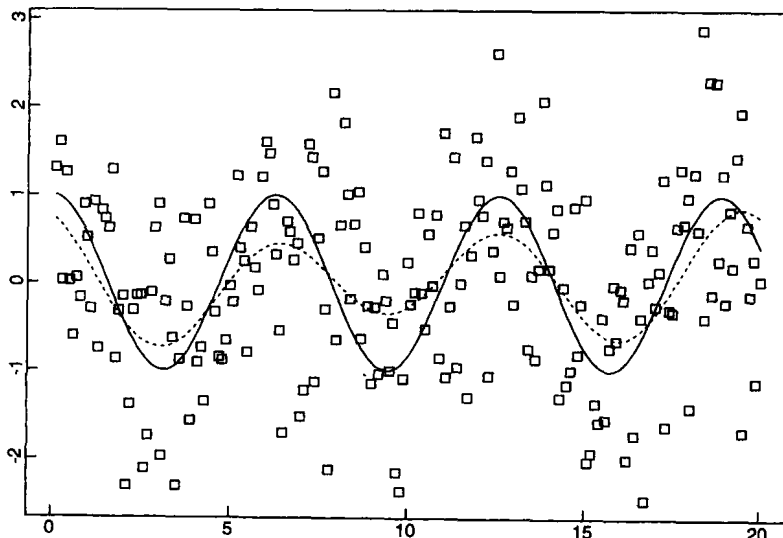
Figure 1.4 shows simulated values of the series

$$X_t = \cos\left(\frac{t}{10}\right) + N_t, \quad t = 1, 2, \dots, 200,$$

where  $\{N_t\}$  is a sequence of independent normal random variables, with mean 0 and variance 0.25. Such a series is often referred to as *signal plus noise*, the signal being the smooth function,  $S_t = \cos(\frac{t}{10})$  in this case. Given only the data  $X_t$ , how can we determine the unknown signal component? There are many approaches to this general problem under varying assumptions about the signal and the noise. One simple approach is to *smooth* the data by expressing  $X_t$  as a sum of sine waves of various frequencies (see Section 4.2) and eliminating the high-frequency components. If we do this to the values of  $\{X_t\}$  shown in Figure 1.4 and retain only the lowest

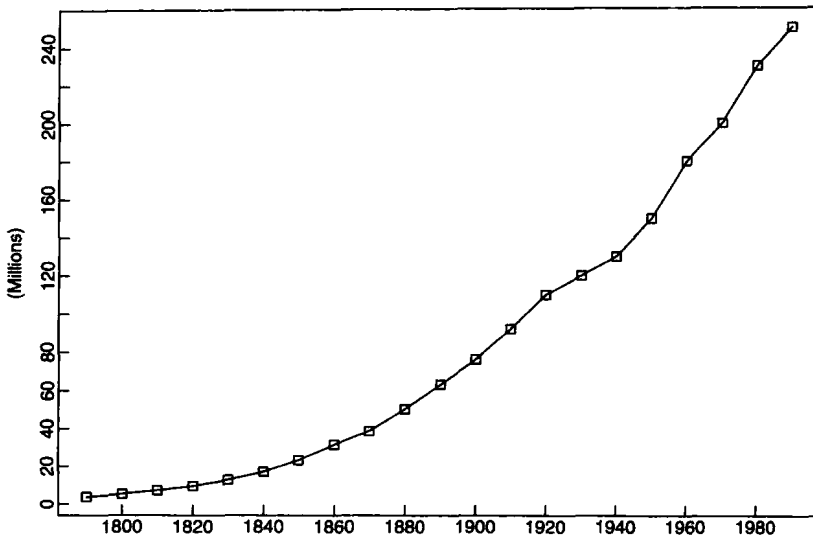


**Figure 1-3** The monthly accidental deaths data, 1973–1978.

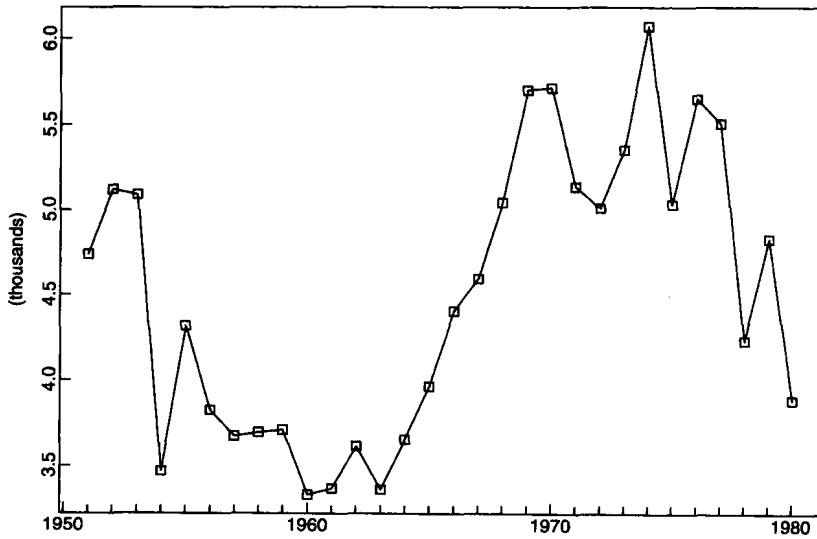


**Figure 1-4** The series  $\{X_t\}$  of Example 1.1.4.





**Figure 1-5** Population of the U.S.A. at ten-year intervals, 1790–1990.



**Figure 1-6** Strikes in the U.S.A., 1951–1980.