图灵原版数学·统计学系列 30

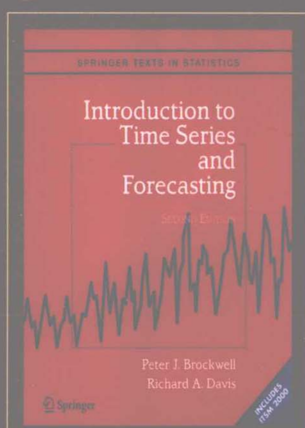
Springer

Introduction to Time Series and Forecasting

时间序列与预测

（英文版·第2版）

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



人民邮电出版社
POSTS & TELECOM PRESS

TURING 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
Richard A. Davis 著

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.

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[①] 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 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, çoauthor 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

1 Introduction

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

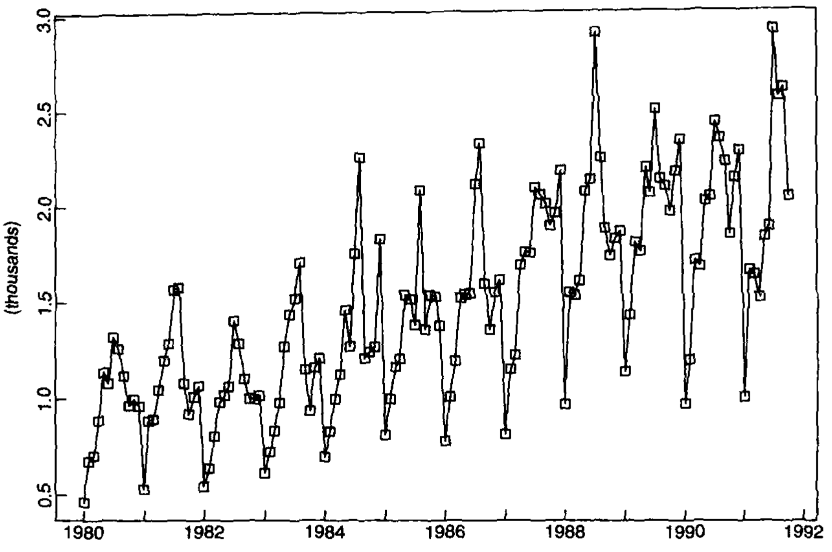


Figure 1-1 The Australian red wine sales, Jan. '80 – Oct. '91.

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 {(Jan. 1980), (Feb. 1980), ...,(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 $\{1, 2, \ldots, n\}$. In the present example this amounts to measuring time in months with (Jan. 1980) as month 1. Then T_0 is the set $\{1, 2, \ldots, 142\}$. 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. □

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

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

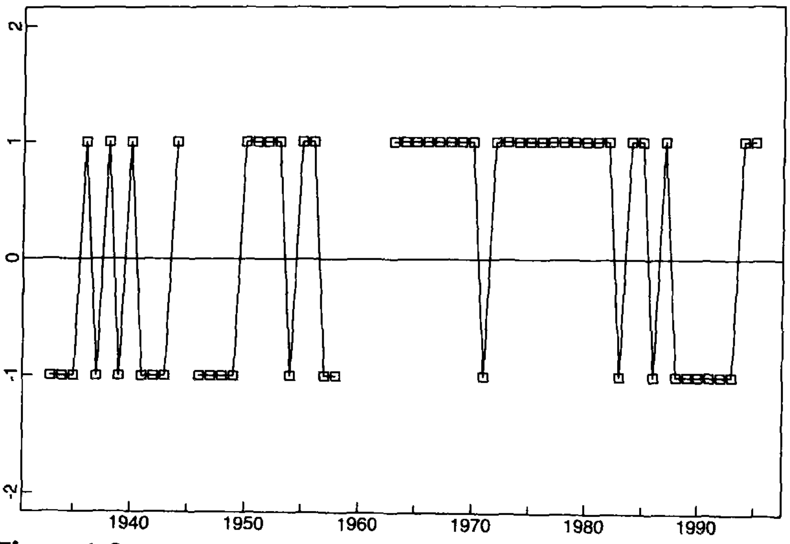


Figure 1-2 Results of the all-star baseball games, 1933–1995.

This is a series with only two possible values, ± 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. □

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. □

Example 1.1.4 A signal detection problem; SIGNAL.TSM

Figure 1.4 shows simulated values of the series

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

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

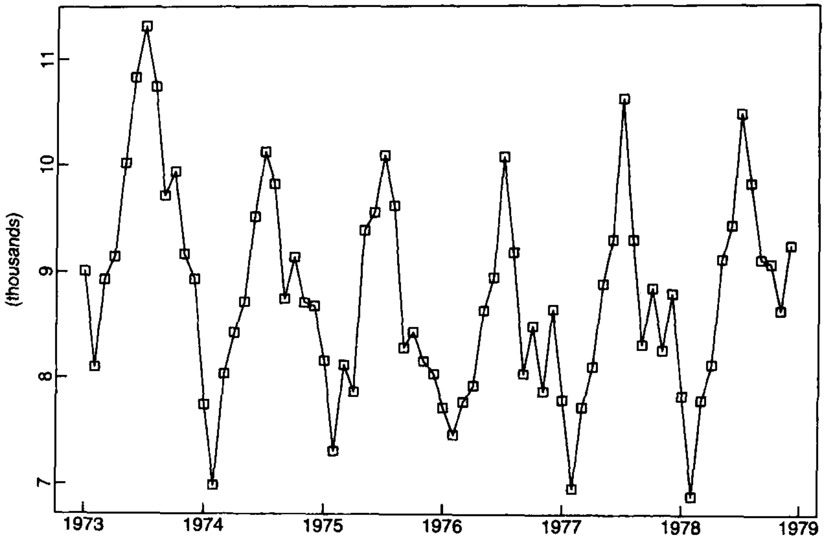


Figure 1-3 The monthly accidental deaths data, 1973–1978.

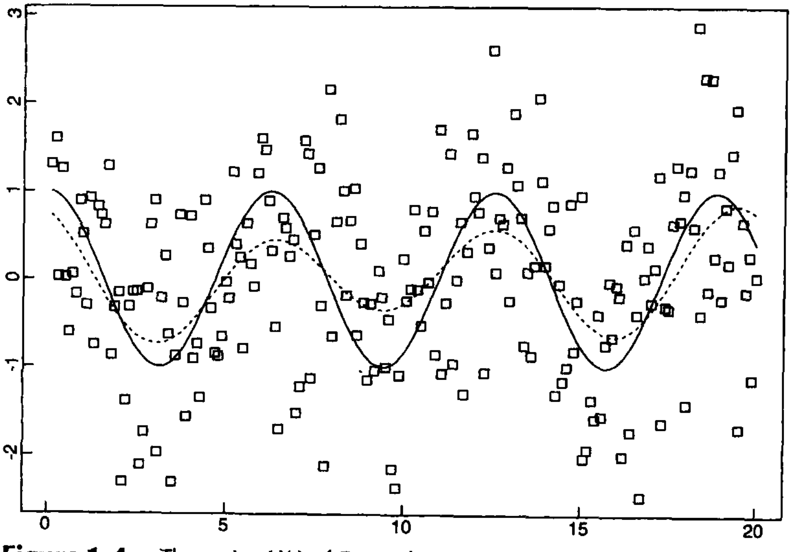


Figure 1-4 The series $\{X_t\}$ of Example 1.1.4.

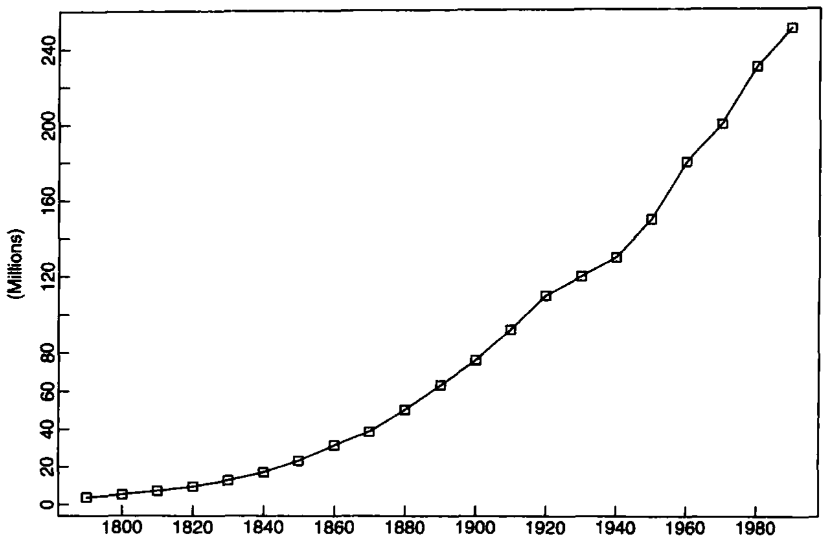


Figure 1-5 Population of the U.S.A. at ten-year intervals, 1790–1990.

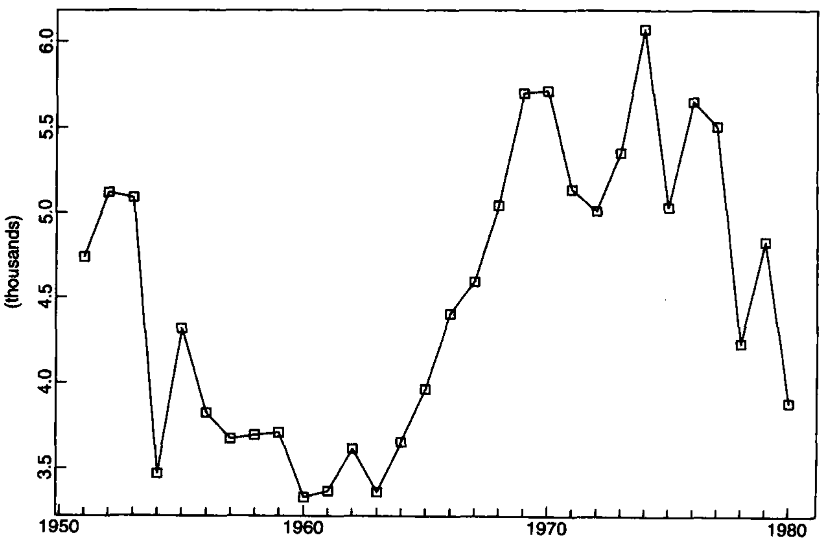


Figure 1-6 Strikes in the U.S.A., 1951–1980.