A Course in Time Series Analysis

Edited by

DANIEL PEÑA
Universidad Carlos III de Madrid

GEORGE C. TIAO
University of Chicago

RUEY S. TSAY
University of Chicago

This page intentionally left blank

A Course in Time Series Analysis

A Course in Time Series Analysis

Edited by

DANIEL PEÑA
Universidad Carlos III de Madrid

GEORGE C. TIAO
University of Chicago

RUEY S. TSAY
University of Chicago

Contents

PART II ADVANCED TOPICS IN UNIVARIATE TIME SERIES

9. Heteroscedastic Models   **249**

R. S. Tsay

This page intentionally left blank

Preface

This book is based on the lectures of the ECAS' 97 Course in Time Series Analysis held at El Escorial, Madrid, Spain, from September 15 to September 19, 1997. The course was sponsored by the European Courses in Advanced Statistics (ECAS). In accordance with the objectives of ECAS, the lectures are directed to both researchers and teachers of statistics in academic institutions and statistical professionals in industry and government, with the goal of presenting an overview of the current status of the area. In particular, different approaches to time series analysis are discussed and compared. In editing the book, we have worked hard to uphold ECAS' objectives. In addition, special efforts have been made to unify the notation and to include as many topics as possible, so that readers of the book can have an overview of the current status of time series research and applications.

The book consists of three main components. The first component concern basic materials of univariate time series analysis presented in the first eight chapters. It includes recent developments in outlier detection, automatic model selection, and seasonal adjustment. The second component addresses advanced topics in univariate time series analysis such as conditional heteroscedastic models, nonlinear models, Bayesian analysis, nonparametric methods, and neural networks. This component represents current research activities in univariate time series analysis. The third and final component of the book concerns with multivariate time series, including vector ARMA models, cointegration, and linear systems.

The book can be used as a principal text or a complementary text for courses in time series. A basic time series course can be taught from the first part of the book that presents the basic material that can be found in the standard texts in time series. This part also includes topics not normally covered in these texts, such as the extended and inverse autocorrelation function, the decomposition of the forecast function of ARIMA models, a detailed analysis of outliers and influential observations and automatic methods for model building and model based seasonal adjustment. For a basic course this book should be complemented with some of the excellent texts available. The book would be very well suited for an advanced course in which some of the basic material can be quickly reviewed using the first part, that skips many details and concentrates in the main concepts of general applicability. Then the

course can concentrate in the topics in Parts 2 and 3. If the scope of the course is more in methodological extensions of univariate linear models the material in Part 2 can be useful, whereas if the objective is to introduce multivariate modeling Part 3 will be appropriate. To facilitate the use of the book as a text, all the time series data used in this book can be down loaded from the web address: *http://gsbwww. uchicago.edu/fac/ruey.tsay/teaching/ecas/*

We are grateful to all people who have made this book possible: (1) to the 11 authors of the chapters of the book who have been extremely helpful in the timely revisions of the drafts of the chapters and have made a big effort to unify the presentation and (2) to the organizers of the course and all the students from many different countries in four continents that made this one week of lectures a very enjoyable experience for all the participants. We are very grateful to our host in the Monastery of El Escorial, father Agustin Alonso, who did his best to make our staying in the monastery an unforgettable experience. The success of the course was in large part due to the enthusiastic work in all the organization details of Ana Justel, Regina Kaiser, Juan Romo, Esther Ruiz, and María Jesús Sánchez. In the preparation of the book we are also grateful to Monica Benito for her help in organizing the index and the references in the book.

The Editors

ABOUT ECAS

ECAS is a foundation of Statistical Societies within Europe that, according to its constitution, was founded in order to foster links and to promote cooperation between statisticians in Europe. In order to achieve these aims, courses on an advanced level covering varying aspects of statistics are organized every 2 years in different countries of Europe. In 1999 Statistical Societies members of ECAS belongs to the following countries: Austria, Belgium, Denmark, France, Finland, Germany, Italy, Portugal, Spain, Sweden, Switzerland, The Netherlands, and the United Kingdom.

The first ECAS course was held in Capri, Italy, on Multidimensional Data Analysis in 1987. Subsequent courses were held on robustness in statistics in 1989 in the castle Reisenburg, Germany; on experimental design in 1991 in Sète, France; on the analysis of categorical data in 1995 in Leiden, The Netherlands; on longitudinal data analysis and repeated measures in 1995 in Milton Keynes, United Kingdom; on time series analysis in 1997 in San Lorenzo del Escorial, Spain; and on environmental statistics in 1999 in Garpenberg, Sweden.

A Council has the overall responsibility for ECAS. Its members are nominated by the statistical societies of participating countries. The Presidents of ECAS have been Jean Jacques Debrosque (Belgium, 1987–1993) and Siegfried Heiler (Germany, 1994–1997). The current President is Daniel Peña (Spain, 1998–2001).

Contributors

Manfred Deistler
Institut für Ökonometrie, Operations
    Research und Systemtheorie
Technische Universität Wien
Wien, Austria

Víctor Gómez
Dirección General de Presupuestos
Ministerio de Hacienda
Madrid, Spain

Siegfried Heiler
Fakultaet fuer Mathematik und
    Informatik
Universität Konstanz
Konstanz, Germany

Kurt Hornik
Institut für Statistik
Technische Universität Wien
Wien, Austria

Søren Johansen
Economic Department
European University Institute
Florence, Italy

Friedrich Leisch
Institut für Statistik
Technische Universität Wien
Wien, Austria

Agustín Maravall
Servicio de Estudios
Banco de España
Madrid, Spain

Daniel Peña
Departamento de Estadística y
    Econometría
Universidad Carlos III de
    Madrid
Madrid, Spain

George C. Tiao
Graduate School of
    Business
University of Chicago
Chicago, IL, USA

Ruey S. Tsay
Graduate School of
    Business
University of Chicago
Chicago, IL, USA

G. Tunnicliffe Wilson
Department of Mathematics and
    Statistics
Lancaster University
Lancaster, UK

This page intentionally left blank

A Course in Time Series Analysis

This page intentionally left blank

CHAPTER 1

Introduction

Daniel Peña
Universidad Carlos III de Madrid

George C. Tiao
University of Chicago

1.1. EXAMPLES OF TIME SERIES PROBLEMS

Data in business, economics, engineering, environment, medicine, and other areas of scientific investigations are often collected in the form of time series, that is, a sequence of observations taken at regular intervals of time such as hourly temperature readings, daily stock prices, weekly traffic volume, monthly beer consumption, and annual growth rates. The main objectives of time series modeling and analysis are (1) understanding the dynamic or time-dependent structure of the observations of a single series—*univariate* time series analysis and (2) ascertaining the leading, lagging, and feedback relationships among several series—*multivariate* time series analysis.

Knowledge of the dynamic structure will help produce accurate forecasts of future observations and design optimal control schemes. This chapter presents first a number of univariate and multivariate time series data sets arisen from various scientific disciplines. These data examples are used to introduce and illustrate the following:

- Stationary versus nonstationary series
- Linear versus nonlinear dynamic relationship
- Homogeneity versus heterogeneity in variance
- Unidirectional versus feedback relation between series

- Outlier, level shift, structural change, and intervention
- Comovement and cointegration

These concepts motivate many of the topics discussed in this book.

1.1.1. Stationary series

Figure 1.1a shows a series of the yield of 70 consecutive batches of a chemical process given in Box and Jenkins (1976). The observations fluctuate about a fixed mean level with constant variance over the observational period. In other words, the overall behavior of the series remains the same over time. Such a series is called a *stationary series*. A formal definition of stationarity will be given later.

(a) Yield of Chemical Process - Box-Jenkins Series F

(b) Monthly Changes in 90-Day T-Bill Rate 1983-1993
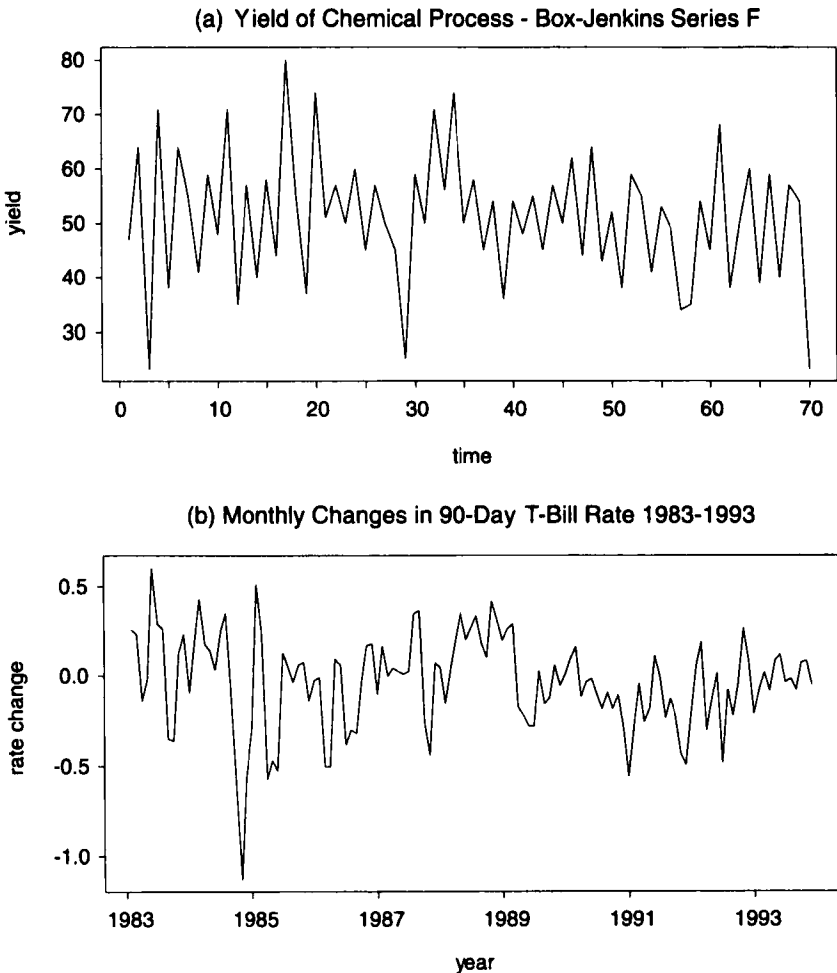
FIGURE 1.1   Two examples of stationary series.

As another example of stationary series, Figure 1.1b gives a series of the month to month changes in the interest rates of 90-day U.S. Treasury bills (T-bills) from 1983 to 1993. Except for the sharp dip near the end of 1984, this series appears to be quite stationary with a mean level close to zero over time.

In practice, temporal changes (week to week, month to month, or quarter to quarter) of many economic time series often exhibit this kind of stationary behavior. Good examples are stock returns and changes in exchange rates.

1.1.2. Nonstationary series

Instead of month to month changes, if we look at the series of monthly rates of the 90-day T-bills themselves, we see a vastly different behavior. This is shown in Figure 1.2a.
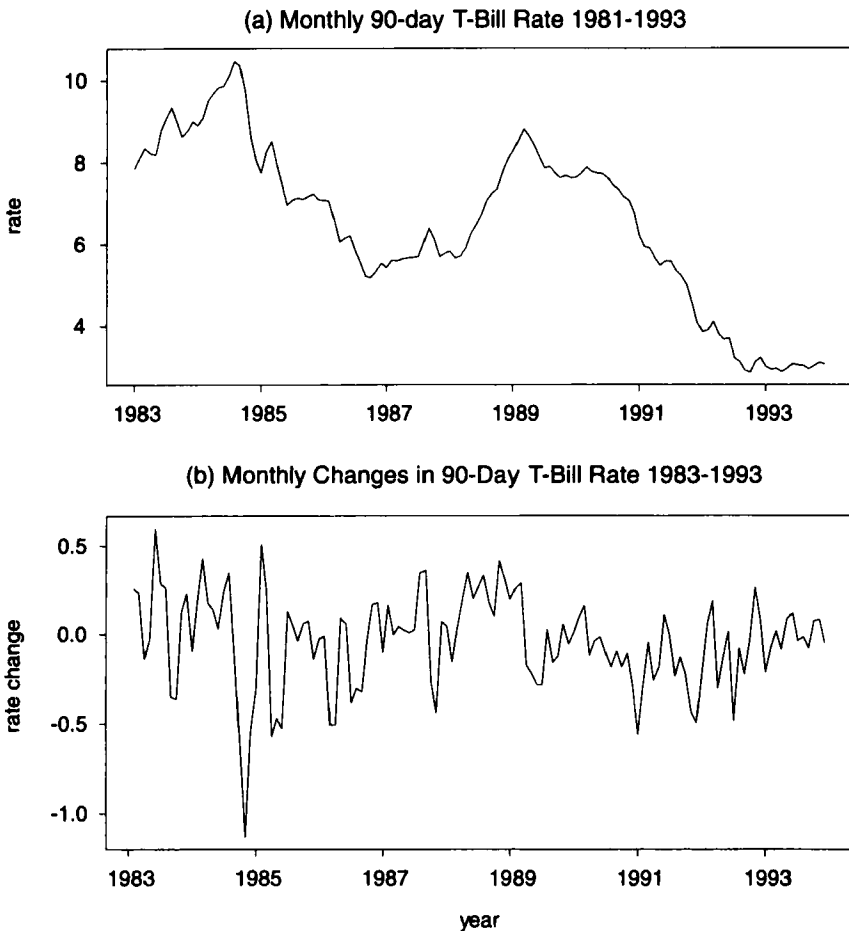
FIGURE 1.2   A nonstationary series and its first difference.

This series does not seem to have a mean level and exhibits a drifting or wandering behavior. It is clearly not a stationary series. Financial time series such as stock prices, prices of derivatives, and exchange rates often behave in this manner. However, by taking successive differences of the observations, we obtain the series of monthly changes in Figure 1.1b, which is reproduced in Figure 1.2b for easy comparison. This example shows that a drifting nonstationary series can be transformed into a stationary one by the differencing operation. The series in Figure 1.2b is called the first difference of the series in Figure 1.2a. In practice, sometimes the first difference series may not be stationary and it may be necessary to difference the series again to make it stationary.

Figure 1.3a shows quarterly data of U.S. real GNP (gross national product) over the period 1946–1991. The series shows an exponential growth. By taking a logarithmic
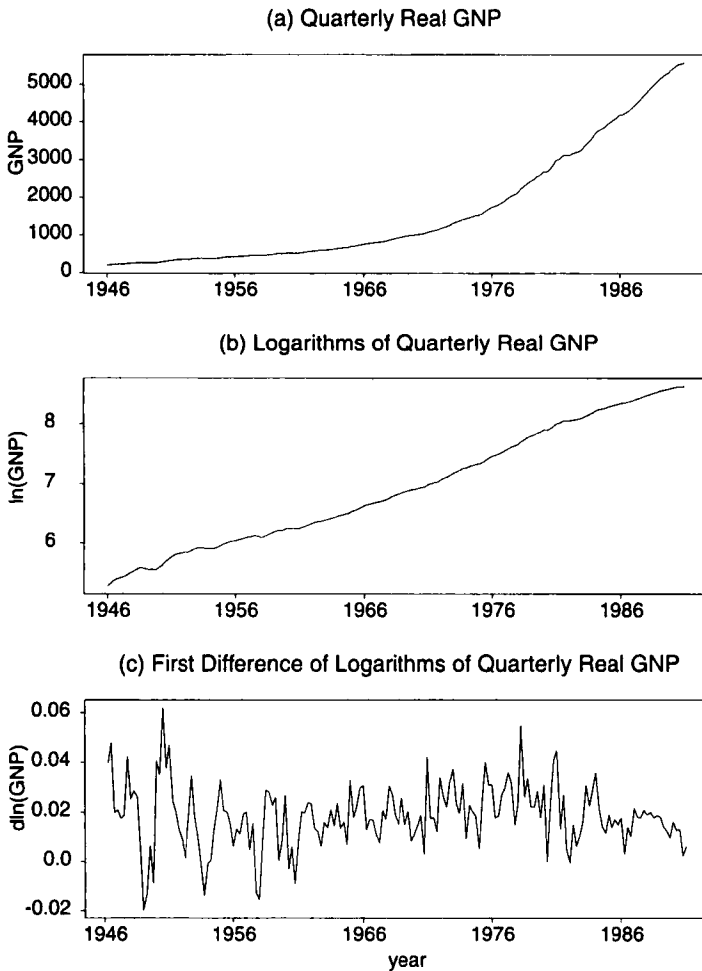
FIGURE 1.3   Quarterly U.S. real GNP 1946–1991.
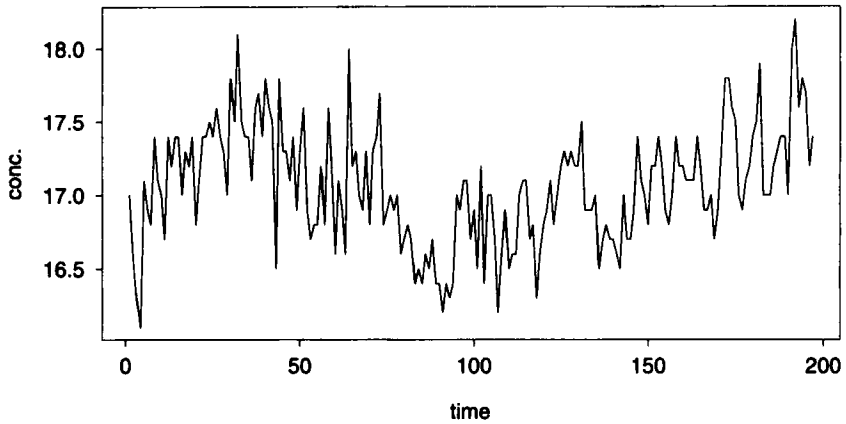
FIGURE 1.4 Concentration readings of a chemical process: Box–Jenkins series A.

transformation of the observations, we see a persistent linear growth in Figure 1.3b. The first difference series of the logged data is shown in Figure 1.3c, which is fairly stationary, although there appears to have some changes in the variability of the series over the data period.

Figure 1.4 presents series A and Figure 1.5 shows series C of Box and Jenkins (1976). The first appears to lie in the gray area of a stationary or a nonstationary series. For the second, differencing may be called for. Both have been used in the literature by other authors to illustrate novel methods for modeling time series.

1.1.3. Seasonal series

Time series data in business, economics, environment, and other disciplines often exhibit a strong cyclical or seasonal behavior. Modeling and analyzing such series
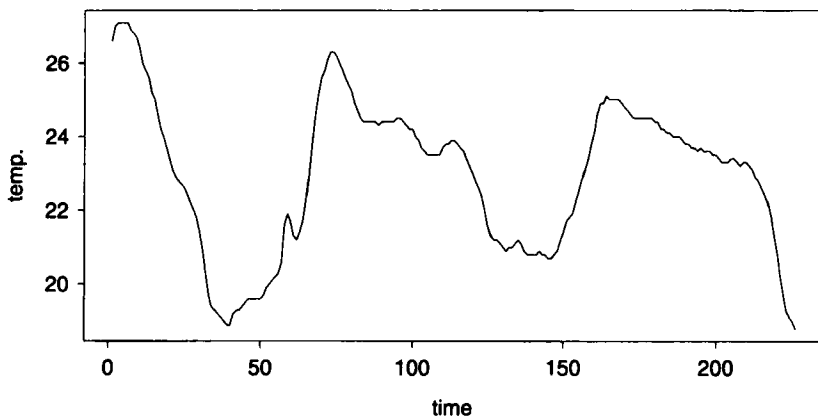
FIGURE 1.5 Temperature readings: Box–Jenkins series C.

(a) International Airline Passenger Totals - Box-Jenkins Series G

(b) Monthly Readings of Ozone at Downtown L.A. 1955-1972
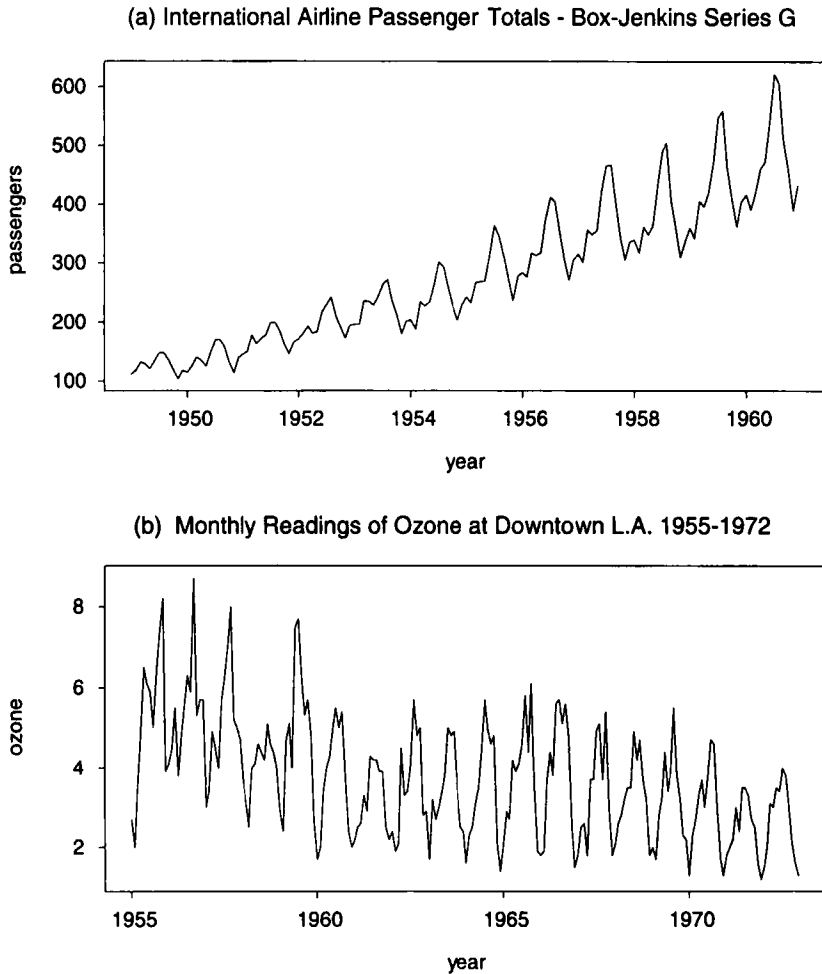
FIGURE 1.6   Two examples of seasonal series.

is an important topic in time series study. Figure 1.6a shows monthly international airline passenger totals in 1949–1959, which were used by Box and Jenkins (1976) to illustrate their innovative seasonal models. In practice, the user of the data may wish to remove the seasonality from the series in order to discern the "underlying trend," and this has led to the vast literature on seasonal decomposition and seasonal adjustment, which will be discussed later.

Figure 1.6b shows monthly averages of ozone in downtown Los Angeles during the period 1955–1972. Ambient ozone is an indicator of air pollution and is strongly seasonal: high in the summer months and low in the winter. In addition to seasonal cycles, there appears to be a level shift in the beginning of the sixth year and a down trend in the last 7 years of the data. The level shift may be associated with changes in

FIGURE 1.7 The Crest market share weekly data 1958–1963.

the traffic pattern and/or changes in the composition of gasoline sold in Los Angeles, and the down trend may be the result of progressively more stringent air q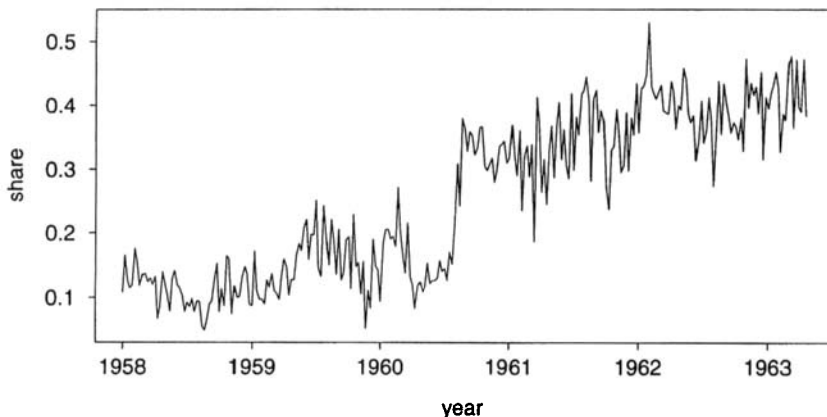uality standards at that time. This series was used by Box and Tiao (1975) to motivate intervention analysis in time series.

1.1.4. Level shifts and outliers in time series

Another example of a level shift is shown in Figure 1.7, the weekly market share data of Crest toothpaste from January 1958 to April 1963. In August 1960, the American Dental Association publicly endorsed Crest, and this led to a substantial jump in its market share as it is clearly seen in the figure. If the timing of this event is known, as in this case, the intervention analysis techniques can be applied to estimate its effect. In practice, such interventions are often unknown to the investigator, and detection of level shifts, outliers, and other types of structural changes becomes an important problem in time series analysis.

1.1.5. Variance changes

Figure 1.8 shows monthly returns of value-weighted S&P (Standard and Poor) 500 stocks from 1926 to 1991. While the mean level stayed close to zero over the entire period, it is clear that changes in the variance, called *volatility* in the finance literature, occurred. There has been an intense interest in modeling data of this kind in recent (at the time of writing) years, and some of the methods will be discussed later in the book.

1.1.6. Asymmetric time series

Two time series are shown in Figure 1.9. The first, in panel (a) is a series of annual sunspot numbers from 1700 to 1979; the other in panel (b), shows seasonally adjusted quarterly U.S. unemployment rates from 1948 to 1993. Both series share a common
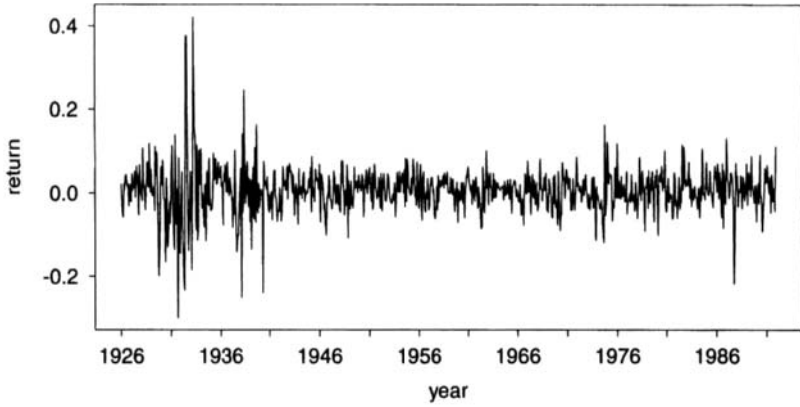
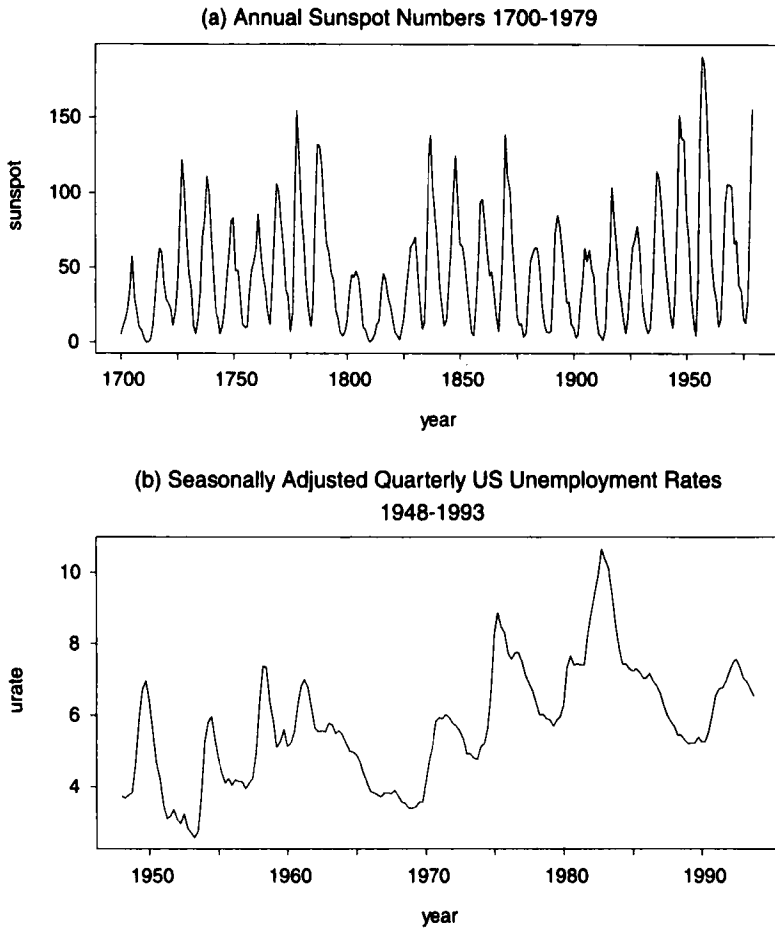**FIGURE 1.8**   Value-weighted S&P 500 returns 1926–1991.

(a) Annual Sunspot Numbers 1700-1979

**(b) Seasonally Adjusted Quarterly US Unemployment Rates
1948-1993**

FIGURE 1.9   Two nonlinear time series.