Handbook of Statistics

VOLUME 5

General Editor

P. R. Krishnaiah

Time Series in the Time Domain

Edited by

E. J. Hannan

P. R. Krishnaiah

M. M. Rao

Time Series in the Time Domain

Edited by

E. J. Hannan
Department of Statistics
Australian National University, Canberra, Australia

P. R. Krishnaiah
Center for Multivariate Analysis
University of Pittsburgh, Pittsburgh, PA, U.S.A.

M. M. Rao
Department of Mathematics
University of California, Riverside, CA, U.S.A.

1985

NORTH-HOLLAND
AMSTERDAM · NEW YORK · OXFORD

ISBN: 0 444 87629 4

Published by:
ELSEVIER SCIENCE PUBLISHERS B.V.
P.O. Box 1991
1000 BZ Amsterdam
The Netherlands

Sole distributors for the U.S.A. and Canada:
ELSEVIER SCIENCE PUBLISHING COMPANY, INC.
52 Vanderbilt Avenue
New York, N.Y. 10017
U.S.A.

Library of Congress Cataloging-in-Publication Data

Main entry under titel:

Time series in the time domain.

(Handbook of statistics ; v. 5)
Includes bibliographies and index.
1. Time-series analysis. I. Hannan, E. J. (Edward James), 1921- . II. Krishnaiah, Paruchuri R. III. Rao, M. M. (Malempati Madhusudana), 1929- IV. Series.
QA280.T57 1985 519.5'5 85-10207
ISBN 0-444-87629-4

PRINTED IN THE NETHERLANDS

HANDBOOK OF STATISTICS
VOLUME 5

Handbook of Statistics

VOLUME 5

General Editor

P. R. Krishnaiah

NORTH-HOLLAND
AMSTERDAM · NEW YORK · OXFORD

Preface

The theory and practice of the analysis of time series has followed two lines almost since its inception. One of these proceeds from the Fourier transformation of the data and the other from a parametric representation of the temporal relationships. Of course, the two lines are interrelated. The frequency analysis of data was surveyed in Volume 3 of the present *Handbook of Statistics* series, subtitled, *Time Series in the Frequency Domain*, edited by D. R. Brillinger and P. R. Krishnaiah. Time domain methods are dealt with in this volume. The methods are old, going back at least to the ideas of Prony in the eighteenth century, and owe a great deal to the work of Yule early this century.

Several different techniques for classes of nonstationary processes have been developed by various analysts. By the very nature of the subject in these cases, the work tends to be either predominantly data analysis oriented with scant justifications, or mathematically oriented with inevitably advanced arguments. This volume contains descriptions of both these approaches by strengthening the former and minimizing the latter, and yet presenting the state-of-the-art in the subject. A brief indication of the work included is as follows.

One of the successful parametric models is the classical autoregressive scheme, going back to the pioneering work of G. U. Yule, early in this century. The model is a difference equation with constant coefficients, and much of the classical work is done if the roots of its characteristic equation are interior to the unit circle. If the roots are of unit modulus, the analysis presents many difficulties. The advances made in recent years in this area are described in W. Fuller's article. An important development in the time domain area is the work of R. Kalman. It led to the emphasis on a formalization of rational transfer function systems as defined by an underlying state vector generated in a Markovian manner and observed subject to noise. This representation is connected with a rich structure theory whose understanding is central in the subject. It is surveyed in the article by M. Deistler. The structure and analysis of several classes of nonstationary time series that are not of autoregressive type but for which the ideas of Fourier analysis extend is given in the article by M. M. Rao; and the filtering and smoothing problems are discussed by D. K. Chang. Related results on what may be termed "asymptotically stationary" and allied time series have been surveyed in C. S. K. Bahagavan's paper.

The papers by L. Ljung, P. Young and G. C. Tiao relate to the estimation problems in the dynamical modelling systems. Here Young's paper deals with the on-line (real time) calculations. One of the uses of these models has been to analyze the consequences of an intervention (such as the introduction of exhaust emission laws) and another to consider the outlier detection problems. These are discussed by Tiao and T. Ozaki. Though rational transfer function models are parametric, it is seldom the case that the model set contains the truth and the problem may better be viewed as one of selecting a structure from an infinite set in some asymptotically optimal manner. This point of view is explored by R. Shibata. Though least squares techniques, applied to the prediction errors, have dominated, there is a need to modify these to obtain estimators less influenced by discrepant observations. This is treated by Tiao and, in an extensive discussion, by R. D. Martin and V. J. Yohai. The model selection and unequally spaced data are natural problems in this area confronting the experimenter, and these are discussed by R. H. Jones. Since the time points may sometimes be under control of the experimenter, their optimal choice must be considered. This problem is treated by S. Cambanis. The modelling in the papers referred to above has been essentially linear. Ozaki presents an approach to the difficult problem of nonlinear modelling.

The autoregressive models may have time varying parameters, and this is considered by D. F. Nicholls and A. R. Pagan. Their paper has special reference to econometric data as does also the paper by H. Theil and D. G. Fiebig who treat the problem where the regressor vectors in a multivariate system may be of a dimension higher than the number of time points for observation. The final two papers on applications by M. A. Cameron, P. J. Thomson and P. de Souza complement the areas covered by the preceding ones. These are designed to show two special applications, namely in signal attenuation estimation and speech recognition.

Thus several aspects of the time domain analysis and the current trends are described in the different chapters of this volume. So they will be of interest not only to the research workers in the area of time series, but also to data analysts who use these techniques in their work.

We wish to express our sincere appreciation to the authors for their excellent cooperation. We also thank the North-Holland Publishing Company for their cooperation.

E. J. Hannan
P. R. Krishnaiah
M. M. Rao

Contributors

C. S. K. Bhagavan, *Dept. of Statistics, Andhra University, Waltair, India 530003* (Ch. 11)

S. Cambanis, *Dept. of Statistics, University of North Carolina, Chapel Hill, NC 27514, USA* (Ch. 13)

M. A. Cameron, CSIRO, *Division of Mathematics & Statistics, P.O. Box 218, Lindfield, N.S.W., Australia 2070* (Ch. 14)

D. K. Chang, *Dept. of Mathematics, California State University, Los Angeles, CA 90023, USA* (Ch. 12)

M. Deistler, *Institute of Econometrics, Technical University of Vienna, Argentinierstr. 8, A 1040 Vienna, Austria* (Ch. 9)

P. de Souza, *Dept. of Mathematics, Victoria University, Wellington, New Zealand* (Ch. 15)

D. G. Fiebig, *University of Sydney, Sydney, N.S.W., Australia 2006* (Ch. 17)

W. A. Fuller, *Dept. of Statistics, Iowa State University, Ames, IA 50011, USA* (Ch. 1)

R. H. Jones, *Scientific Computing Center, University of Colorado Medical Center, Box B-119, Denver, CO 80262, USA* (Ch. 5)

L. Ljung, *Dept. of Electrical Engineering, Linköping University, S-581 83 Linköping, Sweden* (Ch. 7)

R. D. Martin, *Dept. of Statistics, GN22, B313 Padelford Hall, University of Washington, Seattle, WA 98195, USA* (Ch. 4)

D. F. Nicholls, *Statistics Dept., Australian National University, G.P.O. Box 4, Canberra, A.C.T., Australia 2601* (Ch. 16)

T. Ozaki, *The Institute of Statistical Mathematics, 4-6-7-Minami-Azabu, Minato-Ku, Tokyo, Japan* (Ch. 2)

A.R. Pagan, *Statistics Dept., Australian National University, G.P.O. Box 4, Canberra, A.C.T., Australia 2601* (Ch. 16)

M.M. Rao, *Dept. of Mathematics, University of California, Riverside, CA 92521, USA* (Ch. 10)

R. Shibata, *Dept. of Mathematics, Keio University, 3-14-1 Hiyoshi, Kohoku, Yokohama 223, Japan* (Ch. 6)

H. Theil, *College of Business Administration, Dept. of Economics, University of Florida, Gainesville, FL 32611, USA* (Ch. 17)

P. J. Thomson, *Institute of Statistics and Operations Research, Victoria University, Wellington, New Zealand* (Ch. 14, 15)

G. C. Tiao, *Graduate School of Business, University of Chicago, Chicago, IL 60637, USA* (Ch. 3)

V. J. Yohai, *Department of Mathematics, Piso 7, University of Buenos Aires, Argentina* (Ch. 4)

P. Young, *Dept. of Environmental Sciences, University of Lancaster, Lancaster LA1 4YQ, England* (Ch. 8)

Table of Contents

Ch. 11. On Non-Stationary Time Series 311
C. S. K. Bhagavan

Ch. 12. Harmonizable Filtering and Sampling of Time Series 321
D. K. Chang

Ch. 13. Sampling Designs for Time Series 337
S. Cambanis

Ch. 14. Measuring Attenuation 363
M. A. Cameron and P. J. Thomson

E. J. Hannan, P. R. Krishnaiah, M. M. Rao, eds., *Handbook of Statistics*, Vol. 5
© Elsevier Science Publishers B.V. (1985) 1–23

Nonstationary Autoregressive Time Series

Wayne A. Fuller

1. Introduction

A model often used to describe the behavior of a variable over time is the autoregressive model. In this model it is assumed that the current value can be expressed as a function of preceding values and a random error. If we let Y_t denote the value of the variable at time t, the pth-order real valued autoregressive time series is assumed to satisfy

$$Y_t = g(t) + \sum_{i=1}^{p} \alpha_i Y_{t-i} + e_t, \quad t = 1, 2, \ldots, \tag{1.1}$$

where the e_t, $t = 1, 2, \ldots$, are random variables and $g(t)$ is a real valued fixed function of time. We have chosen to define the autoregressive time series on the positive integers, but the time series might be defined on other domains. The statistical behavior of the time series is determined by the initial values $(Y_0, Y_{-1}, \ldots, Y_{-p+1})$, by the function $g(t)$, by the coefficients $(\alpha_1, \alpha_2, \ldots, \alpha_p)$, and by the stochastic properties of the e_t. We shall, henceforth, assume that the e_t have zero mean and variance σ^2. At a minimum we assume the e_t to be uncorrelated. Often we assume the e_t to be independently and identically distributed.

Let the joint distribution function of a finite set $\{Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}\}$ of the Y_t be denoted by

$$F_{Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}}(y_{t_1}, y_{t_2}, \ldots, y_{t_n}) .$$

The time series is strictly stationary if

$$F_{Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}}(y_{t_1}, y_{t_2}, \ldots, y_{t_n}) = F_{Y_{t_1+h}, Y_{t_2+h}, \ldots, Y_{t_n+h}}(y_{t_1}, y_{t_2}, \ldots, y_{t_n})$$

for all possible sets of indices $t_1, t_2, \ldots, t_n$ and $t_1 + h, t_2 + h, \ldots, t_n + h$ in the set $\{1, 2, \ldots\}$. The time series is said to be covariance stationary if

$$E\{Y_t\} = \mu, \quad t = 1, 2, \ldots,$$

and

$$E\{(Y_t - \mu)(Y_{t+h} - \mu)\} = \gamma(h), \quad t = 1, 2, \ldots; \; h = 0, 1, \ldots,$$

where μ is a real number and $\gamma(h)$ is a real valued function of h.

To study the behavior of the time series Y_t, we solve the difference equation (1.1) and express Y_t as a function of $(e_1, e_2, \ldots, e_t)$ and $(Y_0, Y_{-1}, \ldots, Y_{-p+1})$. The difference equation

$$\omega_i = \sum_{j=1}^{p} \alpha_j \omega_{i-j} \tag{1.2}$$

with initial conditions

$$\omega_0 = 1, \qquad \omega_i = 0, \quad i = -1, -2, \ldots,$$

has solution of the form

$$\omega_i = \sum_{j=1}^{p} c_{ji} m_j^i , \tag{1.3}$$

where m_j are the roots of the characteristic equation

$$m^p - \sum_{j=1}^{p} \alpha_j m^{p-j} = 0 , \tag{1.4}$$

the coefficients c_{ji} are of the form

$$c_{ji} = b_j i^{k_j} , \tag{1.5}$$

and the b_j are such that the initial conditions are satisfied. The exponent k_j is zero if the root m_j is a distinct root. A root with multiplicity r has r coefficients with $k_j = 0, 1, \ldots, r-1$.

Using the ω_i, the time series Y_t can be written as

$$Y_t = \sum_{i=0}^{t-1} \omega_i e_{t-i} + \sum_{i=0}^{p-1} \omega_{t+i} Y_{-i} + \sum_{i=0}^{t-1} \omega_i g(t-i) . \tag{1.6}$$

The mean of Y_t is

$$E\{Y_t\} = \sum_{i=0}^{t-1} \omega_i g(t-i) + \sum_{i=0}^{p-1} \omega_{t+i} E\{Y_{-i}\} . \tag{1.7}$$

Therefore, if $(Y_0, Y_{-1}, \ldots, Y_{-p+1})$ is a fixed vector, the variance of Y_t is a function of t and Y_t is not stationary.

If the roots of (1.4) are less than one in absolute value, then ω_i goes to zero as i goes to infinity. One common model is that in which $g(t) \equiv \alpha_0$. Assume that $(Y_0, Y_{-1}, \ldots, Y_{-p+1})$ is a vector of random variables with common mean

$$\alpha_0\left(1 - \sum_{i=1}^{p} \alpha_i\right)^{-1} \tag{1.8}$$

common variance

$$\sigma^2 \sum_{i=0}^{\infty} \omega_i^2 \tag{1.9}$$

and covariances

$$E\{Y_t Y_{t+h}\} = \sigma^2 \sum_{i=0}^{\infty} \omega_i \omega_{i+h}, \quad t, t+h = 0, -1, \ldots, -p+1 \,. \tag{1.10}$$

If $g(t) = \alpha_0$, if $(Y_0, Y_{-1}, \ldots, Y_{-p+1})$ is independent of $(e_0, e_1, \ldots)$, and if the initial conditions satisfy (1.8), (1.9) and (1.10), then Y_t is covariance stationary.

If the initial conditions do not satisfy (1.8), (1.9) and (1.10), the time series will display a different behavior for small t than for large t. However, if $g(t) = \alpha_0$ and the roots of the characteristic equation are less than one in absolute value, the nonstationarity is transitory. In such a situation, the large-t behavior is that of a stationary time series.

2. The first-order model

We begin our discussion with the first-order model

$$\begin{aligned} Y_t &= \alpha_0 + \alpha_1 Y_{t-1} + e_t, & t &= 1, 2, \ldots, \\ &= Y_0, & t &= 0 \,. \end{aligned} \tag{2.1}$$

Given n observations on the process, several inference problems can be considered. One is the estimation of α_1. Closely related to the estimation problem is the problem of testing hypotheses about α_1, particularly the hypothesis that $\alpha_1 = 1$. Finally, one may be interested in predicting future observations.

A natural estimator for (α_0, α_1) is the least squares estimator obtained by regressing Y_t on Y_{t-1}, including an intercept in the regression. The estimators are

$$\begin{aligned} \hat{\alpha}_1 &= \left[\sum_{t=1}^{n} (Y_{t-1} - \bar{Y}_{(-1)})^2\right]^{-1} \sum_{t=1}^{n} (Y_{t-1} - \bar{Y}_{(-1)})(Y_t - \bar{Y}_{(0)}) \,, \\ \hat{\alpha}_0 &= \bar{Y}_{(0)} - \bar{Y}_{(-1)} \hat{\alpha}_1 \,, \end{aligned} \tag{2.2}$$