

0212
H4

8762434

Time Series in the Time Domain



E8762434

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.



NORTH-HOLLAND
AMSTERDAM · NEW YORK · OXFORD

2849870

© ELSEVIER SCIENCE PUBLISHERS B.V., 1985

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

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 title:

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
ISBN 0-444-87629-4

85-10207

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

Preface v

Contributors xiii

Ch. 1. Nonstationary Autoregressive Time Series 1

W. A. Fuller

- 1. Introduction 1
- 2. The first-order model 3
- 3. The p th-order model 12
- Acknowledgements 21
- References 21

Ch. 2. Non-Linear Time Series Models and Dynamical Systems 25

T. Ozaki

- 1. Introduction 25
- 2. Amplitude-dependent autoregressive models 26
- 3. Diffusion processes and their time discretizations 51
- 4. Estimation 73
- 5. Discussions 77
- Acknowledgement 81
- References 81

Ch. 3. Autoregressive Moving Average Models, Intervention Problems and Outlier Detection in Time Series 85

G. C. Tiao

- 1. Introduction 85
- 2. The univariate ARMA models 88
- 3. Transfer function models, intervention analysis and outlier detection 104
- 4. Illustrative examples 110
- 5. Some aspects of vector ARMA models 116
- References 117

Ch. 4. Robustness in Time Series and Estimating ARMA Models 119

R. D. Martin and V. J. Yohai

1. Robustness concepts 119
2. Estimates for perfectly observed ARMA models 126
3. Estimation of imperfectly observed ARMA models 132
4. General M-estimates 134
5. Residual autocovariance estimates 136
6. Approximate maximum-likelihood type estimates 140
7. Breakdown points and influence curves for time series 150
- References 153

Ch. 5. Time Series Analysis with Unequally Spaced Data 157

R. H. Jones

1. Introduction 157
2. State space and the Kalman filter 157
3. A state-space representation for an ARMA(1, 1) process 160
4. ARMA(p, q) processes 163
5. Stationarity and invertibility 165
6. ARIMA(p, d, q) processes 166
7. Continuous time models for unequally spaced data 168
8. Continuous time AR(p) process with observational error, CAR(p) 170
9. CARMA(p, q) and CARIMA(p, d, q) processes 171
10. Regression with stationary errors 171
11. Variance component models 173
12. Nonlinear optimization 175
13. Conclusion 175
- References 176

Ch. 6. Various Model Selection Techniques in Time Series Analysis 179

R. Shibata

1. Introduction 179
2. Statistics specific for each model 180
3. Other statistics not specific for type of models 183
4. Model selection 185
- References 186

Ch. 7. Estimation of Parameters in Dynamical Systems 189

L. Ljung

1. Introduction 189
2. Time-domain models of dynamical systems 190
3. Models and predictors 193
4. Guiding principles behind identification methods 195

5. Asymptotic properties of the estimates	199
6. Numerical schemes for determining the estimates	199
7. Recursive identification methods	202
8. Recursive prediction error methods	203
9. Pseudolinear regressions	205
10. Asymptotic properties	206
11. Implementation aspects	209
12. Some practical aspects	209
13. Conclusions	210
References	210

Ch. 8. Recursive Identification, Estimation and Control 213

P. Young

1. Introduction	213
2. The transfer function model	215
3. Recursive algorithms for estimating the TF model parameters	217
4. State-variable estimation	222
5. Other recursive algorithms and related topics	225
6. Recursive time-series analysis and the MICROCAPTAIN program package	231
7. Practical experiences with recursive estimation	233
8. State-of-the-art of recursive estimation	250
Acknowledgements	252
References	252

Ch. 9. General Structure and Parametrization of ARMA and State-Space Systems and its Relation to Statistical Problems 257

M. Deistler

1. Introduction	257
2. ARMA representations	259
3. State-space representations	263
4. Canonical forms	265
5. The manifold structure of $M(n)$	269
6. The case of additional a priori information on the parameters: Structural identifiability	271
7. The relation to estimation	272
References	275

Ch. 10. Harmonizable, Cramér, and Karhunen Classes of Processes 279

M. M. Rao

1. Introduction	279
2. Harmonizable processes	280
3. Karhunen class	286
4. Cramér class	291
5. Multivariate harmonizable processes	292
6. Class(KF) and harmonizability	295

- 7. The Cramér-Hida approach and multiplicity 298
- 8. Prediction and related questions 302
- 9. Some inference problems with normal processes 307
- Acknowledgement 309
- References 309

Ch. 11. On Non-Stationary Time Series 311

C. S. K. Bhagavan

- 1. Introduction 311
- 2. Stationarity 311
- 3. Spectrum 312
- 4. Spectra of non-stationary processes 313
- 5. Spectrum and ergodic theorems 317
- 6. Concluding remarks 319
- References 319

Ch. 12. Harmonizable Filtering and Sampling of Time Series 321

D. K. Chang

- 1. Introduction 321
- 2. The linear filtering problem 324
- 3. Optimal signal estimation 328
- 4. Sampling a harmonizable process 330
- 5. A numerical illustration 332
- References 335

Ch. 13. Sampling Designs for Time Series 337

S. Cambanis

- 1. Introduction 337
- 2. The time series setup 339
- 3. The sampling designs 342
- 4. The estimators 344
- 5. Optimal fixed sample size designs and asymptotically optimal designs 349
- 6. Discussion and extension 355
- 7. Related topics 358
- References 361

Ch. 14. Measuring Attenuation 363

M. A. Cameron and P. J. Thomson

- 1. Introduction 363
- 2. The model 364
- 3. Estimation 367

- 4. Estimation in the presence of delays 375
- 5. Applying the methods 381
- References 386

Ch. 15. Speech Recognition Using LPC Distance Measures 389

P. J. Thomson and P. de Souza

- 1. Introduction 389
- 2. The LPC model—a review 393
- 3. Comparative tests for LPC models 397
- 4. Power functions of the tests 406
- 5. Computational costs of the tests 408
- 6. An isolated word recognition experiment 409
- Acknowledgements 410
- References 410

Ch. 16. Varying Coefficient Regression 413

D. F. Nicholls and A. R. Pagan

- 1. Introduction 413
- 2. Random coefficient variation 415
- 3. Evolving coefficient variation 426
- 4. Some special models 443
- 5. Conclusion 445
- Appendix 445
- References 446

Ch. 17. Small Samples and Large Equation Systems 451

H. Theil and D. G. Fiebig

- 1. Introduction 451
- 2. How asymptotic tests can be misleading 451
- 3. Simultaneous equation estimation from undersized samples 453
- 4. The ME distribution of a univariate sample 455
- 5. The ME distribution of a multivariate sample 458
- 6. Experiments in simultaneous equation estimation 461
- 7. Canonical correlations and symmetry-constrained estimation 466
- 8. Conclusions 478
- Appendix 478
- References 479

Subject Index 481

Contents of Previous Volumes 485

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 p th-order real valued autoregressive time series is assumed to satisfy

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

where the e_t , $t = 1, 2, \dots$, 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 $(Y_0, Y_{-1}, \dots, Y_{-p+1})$, by the function $g(t)$, by the coefficients $(\alpha_1, \alpha_2, \dots, \alpha_p)$, 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 $\{Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}\}$ of the Y_t be denoted by

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

The time series is strictly stationary if

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

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

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

and

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

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, \dots, e_t) and $(Y_0, Y_{-1}, \dots, Y_{-p+1})$. The difference equation

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

with initial conditions

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

has solution of the form

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

where m_j are the roots of the characteristic equation

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

the coefficients c_{ji} are of the form

$$c_{ji} = b_j i^{k_j}, \quad (1.5)$$

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 $k_j = 0, 1, \dots, r-1$.

Using the ω_p the time series Y_t can be written as

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

The mean of Y_t is

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

Therefore, if $(Y_0, Y_{-1}, \dots, Y_{-p+1})$ 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 $(Y_0, Y_{-1}, \dots, Y_{-p+1})$ is a vector of random variables with common mean

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

common variance

$$\sigma^2 \sum_{i=0}^{\infty} \omega_i^2 \quad (1.9)$$

and covariances

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

If $g(t) = \alpha_0$, if $(Y_0, Y_{-1}, \dots, Y_{-p+1})$ is independent of (e_0, e_1, \dots) , 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

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

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

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