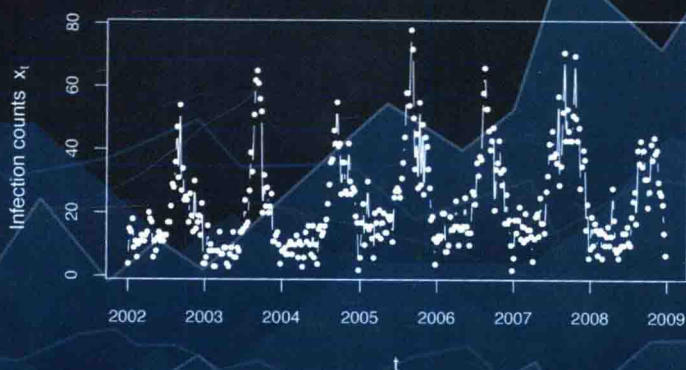


AN INTRODUCTION TO

DISCRETE-VALUED TIME SERIES

CHRISTIAN H. WEISS



WILEY

**A much-needed introduction to the field of discrete-valued time series,
with a focus on count-data time series**

Time series analysis is an essential tool in a wide array of fields, including business, economics, computer science, epidemiology, finance, manufacturing and meteorology, to name just a few. Despite growing interest in discrete-valued time series – especially those arising from counting specific objects or events at specified times – most books on time series give short shrift to that increasingly important subject area. *An Introduction to Discrete-Valued Time Series* seeks to rectify that state of affairs by providing a much needed introduction to discrete-valued time series, with particular focus on count-data time series.

The main focus of this book is on modeling. Throughout, numerous examples are provided illustrating models currently used in discrete-valued time series applications. Statistical process control, including various control charts (such as cumulative sum control charts) and performance evaluation are treated at length.

This book:

- Provides a balanced presentation of theory and practice, exploring both categorical and integer-valued series
- Covers common models for time series of counts as well as for categorical time series, and works out their most important stochastic properties
- Addresses statistical approaches for analyzing discrete-valued time series and illustrates their implementation with numerous data examples

An Introduction to Discrete-Valued Time Series is a valuable working resource for researchers and practitioners in a broad range of fields, including statistics, data science, machine learning and engineering. It will also be of interest to postgraduate students in statistics, mathematics and economics.

CHRISTIAN H. WEISS is a professor in the Department of Mathematics and Statistics, Helmut Schmidt University, Hamburg, Germany. His main area of research is discrete-valued time series. He has published numerous articles in this area and given lectures about time series analysis and discrete-valued time series. He has also written five lecture books in German.



Cover Design: Wiley

Cover Images: (Background) © RKaulitzki/Gettyimages;
(Foreground) Courtesy of Christian H. Weiss

www.wiley.com/go/weiss/discrete-valuedtimeseries

WILEY



Also available
as an e-book

ISBN 978-1-119-09696-2



9

781119

096962

WEISS

AN INTRODUCTION TO

DISCRETE-TIME SERIES



An Introduction to Discrete-Valued Time Series

Christian H. Weiss

WILEY

This edition first published 2018
© 2018 John Wiley & Sons Ltd

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, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

The right of Christian H. Weiss to be identified as the author of this work has been asserted in accordance with law.

Registered Offices

John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Office

9600 Garsington Road, Oxford, OX4 2DQ, UK

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Limit of Liability/Disclaimer of Warranty

While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

Library of Congress Cataloging-in-Publication Data:

Name: Weiss, Christian H., 1977– author.

Title: An introduction to discrete-valued time series / by Christian H. Weiss.

Description: Hoboken, NJ : John Wiley & Sons, 2017. | Includes bibliographical references and index. |

Identifiers: LCCN 2017040687 (print) | LCCN 2017046480 (ebook) | ISBN 9781119096986 (pdf) | ISBN 9781119096993 (epub) | ISBN 9781119096962 (cloth)

Subjects: LCSH: Time-series analysis. | Discrete-time systems—Mathematical models.

Classification: LCC QA280 (ebook) | LCC QA280 .W456 2017 (print) | DDC 519.5/5—dc23

LC record available at <https://lcn.loc.gov/2017040687>

Cover images: (Background) © RKaulitzki/Gettyimages; (Foreground) Courtesy of Christian H. Weiss
Cover design: Wiley

Set in 10/12pt WarnockPro by SPi Global, Chennai, India

Printed and bound in Malaysia by Vivar Printing Sdn Bhd

10 9 8 7 6 5 4 3 2 1

An Introduction to Discrete-Valued Time Series

*To Miia,
Maximilian, Tilman and Amalia*

Preface

People have long been interested in time series: data observed sequentially in time. See Klein (1997) for a historical overview. Nowadays, such time series are collected in diverse fields of science and practice, such as business, computer science, epidemiology, finance, manufacturing or meteorology. In line with the increasing potential for applications, more and more textbooks on time series analysis have become available; see for example the recent ones by Box et al. (2015), Brockwell & Davis (2016), Cryer & Chan (2008), Falk et al. (2012), Shumway & Stoffer (2011) and Wei (2006). These textbooks nearly exclusively concentrate on continuous-valued time series, where real numbers or vectors are the possible outcomes. During the last few decades, however, discrete-valued time series have also become increasingly important in research and applications. These are time series arising from counting certain objects or events at specified times, but they are usually neglected in the textbook literature. Among the few introductory or overview texts on discrete-valued time series are

- the books (or parts thereof) by Fahrmeir & Tutz (2001), Kedem & Fokianos (2002) and Cameron & Trivedi (2013) about regression models
- the book by Zucchini & MacDonald (2009) about hidden-Markov models
- the survey article by McKenzie (2003) in the *Handbook of Statistics*
- the textbook by Turkman et al. (2014), which includes a chapter about models for integer-valued time series
- the book by Davis et al. (2016), which provides a collection of essays about discrete-valued time series.

The present book intends to be an introductory text to the field of discrete-valued time series, and to present the subject with a good balance between theory and application. It covers common models for time series of counts as well as for categorical time series, and it works out their most important stochastic properties. It provides statistical approaches for analyzing discrete-valued time series, and it exemplifies their practical implementation in a number of data examples. It does not constitute a purely mathematical

treatment of the considered topics, but tries to be accessible to users from all those areas where discrete-valued time series arise and need to be analyzed. Inspired by the seminal time series book by Box & Jenkins (1970), there is a strong emphasis on models and methods “possessing maximum simplicity”, but it also provides background and references on more sophisticated approaches. Furthermore, following again the example of Box & Jenkins, the book also includes a part on methods from statistical process control, for the monitoring of a discrete-valued process.

The book is aimed at academics at graduate level having a basic knowledge of mathematics (calculus, linear algebra) and statistics. In addition, elementary facts about time series and stochastic processes are assumed, as they are typically taught in basic courses on time series analysis (also see the textbooks listed above on time series analysis). To allow the reader to refresh their knowledge and to make this book more self-contained, Appendix B contains background information on, for example, Markov chains and ARMA models. Besides putting the reader in a position to analyze and model the discrete-valued time series occurring in practice, the book can also be used as a textbook for a lecture on this topic. The author has already used parts of the book in courses about discrete-valued time series. To support both its application in practice and its use in teaching, ready-made software implementations for the data examples and numerical examples are available to accompany the book. Although such implementations are generally not restricted to a particular software package, the program codes are written in the R language (R Core Team, 2016), since R is freely available to everyone. But each of the examples in this book could have been done with another computational software like Matlab or Mathematica as well. All the R codes, and most of the datasets, are provided on a companion website, see Appendix C for details.

I am very grateful to Prof. Dr. Konstantinos Fokianos (University of Cyprus), Prof. Dr. Robert Jung (University of Hohenheim), Prof. Dr. Dimitris Karlis (Athens University of Economics and Business) and to M. Sc. Tobias Möller (Helmut Schmidt University Hamburg) for reading the entire manuscript and for many valuable comments. I also wish to thank Prof. Dr. Sven Knoth (Helmut Schmidt University Hamburg) for useful feedback on Part III of this book, as well as M. Sc. Boris Aleksandrov and M. Sc. Sebastian Ottenstreuer (ibid.) for making me aware of some typographical errors. I want to thank Prof. Dr. Kurt Brännäs (Umeå University) for allowing me to share the transactions counts data in Example 4.1.5, Alexander Jonen (Helmut Schmidt University Hamburg) for making me aware of the rig counts data in Example 2.6.2, and again Prof. Dr. Dimitris Karlis for contributing the accidents counts data in Example 3.4.2. Thanks go to the Helmut Schmidt University in Hamburg, to the editorial staff of Wiley, especially to Blesy Regulas and Shyamala Venkateswaran for the production of this book, and to Andrew Montford

(Anglosphere Editing Limited) for the copyediting of the book. Finally, I wish to thank my wife Miia and my children Maximilian, Tilman and Amalia for their encouragement and welcome distraction during this work.

Christian H. Weiss
Hamburg
February 2017

About the Companion Website

Don't forget to visit the companion website for this book:

www.wiley.com/go/weiss/discrete-valuedtimeseries



There you will find valuable material designed to enhance your learning, including:

- codes and datasets

Scan this QR code to visit the companion website



Contents

Preface *xi*

About the Companion Website *xv*

1 Introduction *1*

Part I Count Time Series *9*

2 A First Approach for Modeling Time Series of Counts: The Thinning-based INAR(1) Model *11*

2.0 Preliminaries: Notation and Characteristics of Count Distributions *11*

2.1 The INAR(1) Model for Time-dependent Counts *16*

2.1.1 Definition and Basic Properties *17*

2.1.2 The Poisson INAR(1) Model *20*

2.1.3 INAR(1) Models with More General Innovations *22*

2.2 Approaches for Parameter Estimation *26*

2.2.1 Method of Moments *26*

2.2.2 Maximum Likelihood Estimation *28*

2.3 Model Identification *29*

2.4 Checking for Model Adequacy *32*

2.5 A Real-data Example *34*

2.6 Forecasting of INAR(1) Processes *37*

3 Further Thinning-based Models for Count Time Series *43*

3.1 Higher-order INARMA Models *43*

3.2 Alternative Thinning Concepts *54*

3.3 The Binomial AR Model *59*

3.4 Multivariate INARMA Models *64*

4 INGARCH Models for Count Time Series 73

- 4.1 Poisson Autoregression 73
- 4.2 Further Types of INGARCH Models 85
- 4.3 Multivariate INGARCH Models 93

5 Further Models for Count Time Series 95

- 5.1 Regression Models 95
- 5.2 Hidden-Markov Models 107
- 5.3 Discrete ARMA Models 116

Part II Categorical Time Series 119**6 Analyzing Categorical Time Series 121**

- 6.1 Introduction to Categorical Time Series Analysis 122
- 6.2 Marginal Properties of Categorical Time Series 126
- 6.3 Serial Dependence of Categorical Time Series 128

7 Models for Categorical Time Series 133

- 7.1 Parsimoniously Parametrized Markov Models 133
- 7.2 Discrete ARMA Models 139
- 7.3 Hidden-Markov Models 146
- 7.4 Regression Models 151

Part III Monitoring Discrete-Valued Processes 161**8 Control Charts for Count Processes 163**

- 8.1 Introduction to Statistical Process Control 163
- 8.2 Shewhart Charts for Count Processes 165
 - 8.2.1 Shewhart Charts for i.i.d. Counts 166
 - 8.2.2 Shewhart Charts for Markov-Dependent Counts 171
- 8.3 Advanced Control Charts for Count Processes 177
 - 8.3.1 CUSUM Charts for i.i.d. Counts 178
 - 8.3.2 CUSUM Charts for Markov-dependent Counts 182
 - 8.3.3 EWMA Charts for Count Processes 186

9 Control Charts for Categorical Processes 193

- 9.1 Sample-based Monitoring of Categorical Processes 194
 - 9.1.1 Sample-based Monitoring: Binary Case 194
 - 9.1.2 Sample-based Monitoring: Categorical Case 198