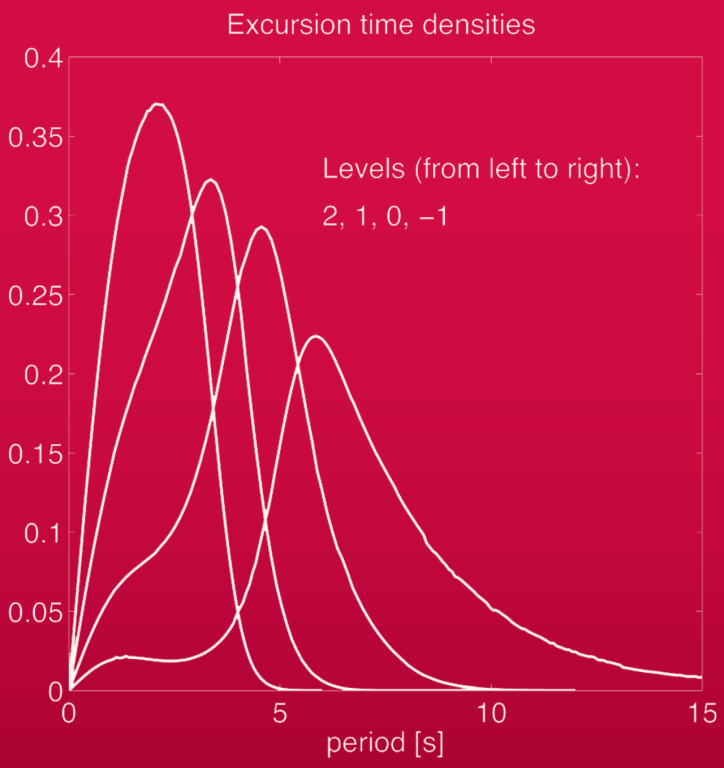
Texts in Statistical Science

Stationary Stochastic Processes

Theory and Applications



Georg Lindgren

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Stationary Stochastic Processes

Theory and Applications

CHAPMAN & HALL/CRC
Texts in Statistical Science Series

Analysis of Failure and Survival Data
P. J. Smith

The Analysis of Time Series — An Introduction, Sixth Edition
C. Chatfield

Applied Bayesian Forecasting and Time Series Analysis
A. Pole, M. West, and J. Harrison

Applied Categorical and Count Data Analysis
W. Tang, H. He, and X.M. Tu

Applied Nonparametric Statistical Methods, Fourth Edition
P. Sprent and N.C. Smeeton

Applied Statistics — Handbook of GENSTAT Analysis
E.J. Snell and H. Simpson

Applied Statistics — Principles and Examples
D.R. Cox and E.J. Snell

Applied Stochastic Modelling, Second Edition
B.J.T. Morgan

Bayesian Data Analysis, Second Edition
A. Gelman, J.B. Carlin, H.S. Stern, and D.B. Rubin

Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians
R. Christensen, W. Johnson, A. Branscum, and T.E. Hanson

Bayesian Methods for Data Analysis, Third Edition
B.P. Carlin and T.A. Louis

Beyond ANOVA — Basics of Applied Statistics
R.G. Miller, Jr.

The BUGS Book: A Practical Introduction to Bayesian Analysis
D. Lunn, C. Jackson, N. Best, A. Thomas, and D. Spiegelhalter

A Course in Categorical Data Analysis
T. Leonard

A Course in Large Sample Theory
T.S. Ferguson

Data Driven Statistical Methods
P. Sprent

Decision Analysis — A Bayesian Approach
J.Q. Smith

Design and Analysis of Experiments with SAS
J. Lawson

Elementary Applications of Probability Theory, Second Edition
H.C. Tuckwell

Elements of Simulation
B.J.T. Morgan

Epidemiology — Study Design and Data Analysis, Second Edition
M. Woodward

Essential Statistics, Fourth Edition
D.A.G. Rees

Exercises and Solutions in Biostatistical Theory
L.L. Kupper, B.H. Neelon, and S.M. O'Brien

Extending the Linear Model with R — Generalized Linear, Mixed Effects and Nonparametric Regression Models
J.J. Faraway

A First Course in Linear Model Theory
N. Ravishanker and D.K. Dey

Generalized Additive Models: An Introduction with R
S. Wood

Generalized Linear Mixed Models: Modern Concepts, Methods and Applications
W. W. Stroup

Graphics for Statistics and Data Analysis with R
K.J. Keen

Interpreting Data — A First Course in Statistics
A.J.B. Anderson

Introduction to General and Generalized Linear Models
H. Madsen and P. Thyregod

An Introduction to Generalized Linear Models, Third Edition
A.J. Dobson and A.G. Barnett

Introduction to Multivariate Analysis
C. Chatfield and A.J. Collins

Introduction to Optimization Methods and Their Applications in Statistics
B.S. Everitt

Introduction to Probability with R
K. Baclawski

Introduction to Randomized Controlled Clinical Trials, Second Edition
J.N.S. Matthews

Introduction to Statistical Inference and Its Applications with R
M.W. Trosset

Introduction to Statistical Limit Theory
A.M. Polansky

Introduction to Statistical Methods for Clinical Trials
T.D. Cook and D.L. DeMets

Introduction to the Theory of Statistical Inference
H. Liero and S. Zwanzig

Large Sample Methods in Statistics
P.K. Sen and J. da Motta Singer

Linear Models with R
J.J. Faraway

Logistic Regression Models
J.M. Hilbe

Markov Chain Monte Carlo — Stochastic Simulation for Bayesian Inference, Second Edition
D. Gamerman and H.F. Lopes

Mathematical Statistics
K. Knight

Modeling and Analysis of Stochastic Systems, Second Edition
V.G. Kulkarni

Modelling Binary Data, Second Edition
D. Collett

Modelling Survival Data in Medical Research, Second Edition
D. Collett

Multivariate Analysis of Variance and Repeated Measures — A Practical Approach for Behavioural Scientists
D.J. Hand and C.C. Taylor

Multivariate Statistics — A Practical Approach
B. Flury and H. Riedwyl

Multivariate Survival Analysis and Competing Risks
M. Crowder

Pólya Urn Models
H. Mahmoud

Practical Data Analysis for Designed Experiments
B.S. Yandell

Practical Longitudinal Data Analysis
D.J. Hand and M. Crowder

Practical Multivariate Analysis, Fifth Edition
A. Afifi, S. May, and V.A. Clark

Practical Statistics for Medical Research
D.G. Altman

A Primer on Linear Models
J.F. Monahan

Principles of Uncertainty
J.B. Kadane

Probability — Methods and Measurement
A. O'Hagan

Problem Solving — A Statistician's Guide, Second Edition
C. Chatfield

Randomization, Bootstrap and Monte Carlo Methods in Biology, Third Edition
B.F.J. Manly

Readings in Decision Analysis
S. French

Sampling Methodologies with Applications
P.S.R.S. Rao

Stationary Stochastic Processes: Theory and Applications
G. Lindgren

Statistical Analysis of Reliability Data
M.J. Crowder, A.C. Kimber, T.J. Sweeting, and R.L. Smith

Statistical Methods for Spatial Data Analysis
O. Schabenberger and C.A. Gotway

Statistical Methods for SPC and TQM
D. Bissell

Statistical Methods in Agriculture and Experimental Biology, Second Edition
R. Mead, R.N. Curnow, and A.M. Hasted

Statistical Process Control — Theory and Practice, Third Edition
G.B. Wetherill and D.W. Brown

Statistical Theory, Fourth Edition
B.W. Lindgren

Statistics for Accountants
S. Letchford

Statistics for Epidemiology
N.P. Jewell

Statistics for Technology — A Course in Applied Statistics, Third Edition
C. Chatfield

Statistics in Engineering — A Practical Approach
A.V. Metcalfe

Statistics in Research and Development, Second Edition
R. Caulcutt

Stochastic Processes: An Introduction, Second Edition
P.W. Jones and P. Smith

Survival Analysis Using S — Analysis of Time-to-Event Data
M. Tableman and J.S. Kim

The Theory of Linear Models
B. Jørgensen

Time Series Analysis
H. Madsen

Time Series: Modeling, Computation, and Inference
R. Prado and M. West

This page intentionally left blank

Texts in Statistical Science

Stationary Stochastic Processes

Theory and Applications

Georg Lindgren

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group an **informa** business

A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20120801

International Standard Book Number-13: 978-1-4665-5780-2 (eBook - PDF)

Statistics is a science in my opinion, and it is no more a branch of mathematics than are physics, chemistry and economics; for if its methods fail the test of experience – not the test of logic – they are discarded.

John W. Tukey (1953)

This page intentionally left blank

Contents

List of figures **xv**

Preface **xvii**

Acknowledgments **xxi**

List of notations **xxiii**

1 Some probability and process background **1**
1.1 Sample space, sample function, and observables 2
1.2 Random variables and stochastic processes 4
1.2.1 Probability space and random variables 4
1.2.2 Stochastic processes and their finite-dimensional distributions 5
1.2.3 The distribution of random sequences and functions 6
1.2.4 An important comment on probabilities on $\mathbb{R}^T$ 12
1.3 Stationary processes and fields 13
1.3.1 Stationary processes, covariance, and spectrum 14
1.3.2 Stationary streams of events 17
1.3.3 Random fields 17
1.4 Gaussian processes 18
1.4.1 Normal distributions and Gaussian processes 18
1.4.2 Conditional normal distributions 20
1.4.3 Linear prediction and reconstruction 21
1.5 Four historical landmarks 22
1.5.1 Brownian motion and the Wiener process 22
1.5.2 S.O. Rice and electronic noise 25
1.5.3 Gaussian random wave models 26
1.5.4 Detection theory and statistical inference 28
Exercises 29

2 Sample function properties **31**
2.1 Quadratic mean properties 32
2.1.1 Quadratic mean continuity 33
2.1.2 Quadratic mean differentiability 35
2.1.3 Higher order derivatives and their correlations 37
2.2 Sample function continuity 39
2.2.1 Countable and uncountable events 39
2.2.2 Conditions for sample function continuity 41
2.2.3 Moment conditions for continuity 45
2.2.4 Continuity of stationary Gaussian processes 47
2.2.5 Probability measures on $\mathbb{C}[0,1]$ 49
2.3 Derivatives, tangents, and other characteristics 50
2.3.1 Differentiability – general conditions 50
2.3.2 Differentiable Gaussian processes 51
2.3.3 Jump discontinuities, Hölder conditions, and tangents 53
2.3.4 Continuity conditions based on entropy 59
2.4 Stochastic integration 61
2.5 An ergodic result 65
Exercises 67

3 Spectral representations **69**
3.1 Complex-valued stochastic processes 69
3.1.1 Moments in complex-valued processes 69
3.1.2 Non-negative definite functions 71
3.1.3 Strict and weak stationarity for Gaussian processes 73
3.2 Bochner's theorem and the spectral distribution 74
3.2.1 The spectral distribution 74
3.2.2 The inversion theorem and the spectrum-covariance relation 78
3.2.3 The one-sided real form 81
3.2.4 Spectrum for stationary sequences 82
3.2.5 Sampling and the alias effect 83
3.3 Spectral representation of a stationary process 84
3.3.1 The spectral process 84
3.3.2 The spectral theorem 88
3.3.3 Discrete and continuous spectrum 92
3.3.4 One-sided spectral representation of a real process 94
3.3.5 Why negative frequencies? 95
3.3.6 Spectral representation of stationary sequences 96
3.3.7 Fourier transformation of data 97

3.4 Gaussian processes 101
3.4.1 The Gaussian spectral process 101
3.4.2 Gaussian white noise in continuous time 102
3.5 Stationary counting processes 105
3.5.1 Counting processes 105
3.5.2 The Poisson process 106
3.5.3 Correlation intensity function 108
3.5.4 Spectral distribution for counting processes 110
Exercises 113

4 Linear filters – general properties 117
4.1 Linear time invariant filters 117
4.1.1 Linear filter on a stationary process 117
4.1.2 Impulse response/frequency response relations 119
4.1.3 Correlation and spectral relations 121
4.1.4 Linear processes 123
4.1.5 Lévy processes 125
4.1.6 Fractional Brownian motion 127
4.2 Linear filters and differential equations 128
4.2.1 The RC-filter and exponential smoothing 128
4.2.2 The linear oscillator 130
4.2.3 Linear differential equations driven by a stochastic process 132
4.3 White noise in linear systems 135
4.3.1 White noise in linear differential equations 135
4.3.2 White noise and constant spectral density 137
4.3.3 Shot noise 139
4.4 Long range dependence, non-integrable spectra, and unstable systems 142
4.4.1 Long range dependence 143
4.4.2 Non-integrable spectral density - intrinsic stationarity 145
4.5 The ARMA-family 149
Exercises 152

5 Linear filters – special topics 155
5.1 The Hilbert transform and the envelope 155
5.1.1 The Hilbert transform 155
5.1.2 The envelope 158
5.1.3 The envelope of a narrow band process 159
5.2 The sampling theorem 161

5.3 Karhunen-Loève expansion 163
5.3.1 Principal components 163
5.3.2 Expansion along eigenfunctions 165
5.3.3 The Karhunen-Loève theorem 166
5.3.4 Testing the mean function 170
Exercises 172

6 Classical ergodic theory and mixing 173
6.1 The basic ergodic theorem in $\mathscr{L}^2$ 173
6.2 Stationarity and transformations 175
6.2.1 Pseudo randomness and transformation 175
6.2.2 Stationarity and measure preserving transformations 176
6.3 The ergodic theorem, transformation view 178
6.3.1 Invariant sets and invariant random variables 179
6.3.2 Ergodicity 180
6.3.3 The Birkhoff ergodic theorem 183
6.4 The ergodic theorem, process view 187
6.4.1 Ergodic stationary sequences 187
6.4.2 Ergodic stationary processes 191
6.5 Ergodic Gaussian sequences and processes 192
6.6 Mixing and asymptotic independence 194
6.6.1 Singularity and regularity 195
6.6.2 Asymptotic independence, regularity, and singularity 197
6.6.3 Uniform, strong, and weak mixing 201
6.6.4 Global mixing conditions for Gaussian processes 203
6.6.5 Adapted mixing conditions 204
Exercises 206

7 Vector processes and random fields 207
7.1 Spectral representation for vector processes 207
7.1.1 Spectral distribution 208
7.1.2 Spectral representation of $\mathbf{x}(t)$ 210
7.1.3 Cross-correlation in linear filters 214
7.2 Some random field theory 216
7.2.1 Homogeneous fields 217
7.2.2 Isotropic fields: covariance properties 221
7.2.3 Isotropic fields: spectral properties 224
7.2.4 Space-time fields 228
Exercises 233

8 Level crossings and excursions **235**
8.1 Level crossings and Rice's formula 236
8.1.1 Level crossings 236
8.1.2 Rice's formula for absolutely continuous processes 237
8.1.3 Alternative proof of Rice's formula 240
8.1.4 Rice's formula for differentiable Gaussian processes 242
8.2 Poisson character of high-level crossings 244
8.3 Marked crossings and biased sampling 247
8.3.1 Marked crossings 247
8.3.2 Biased sampling and empirical distributions 248
8.3.3 Prediction from a random crossing time 251
8.4 The Slepian model 253
8.4.1 Definition 253
8.4.2 A Slepian model in Gaussian processes 255
8.4.3 Local maxima in a Gaussian process 257
8.4.4 Excursion length and related distributions 261
8.5 Crossing problems for vector processes and fields 267
8.5.1 Crossings by vector processes 267
8.5.2 Level sets in fields 271
Exercises 273

A Some probability theory **275**
A.1 Events, probabilities, and random variables 275
A.1.1 Events and families of events 275
A.1.2 Probabilities 277
A.1.3 Sets with countable basis 279
A.1.4 Random variables 280
A.1.5 The σ-field generated by random variables 281
A.2 The axioms of probability 282
A.2.1 The axioms 282
A.2.2 Extension of a probability from field to σ-field 283
A.2.3 Kolmogorov's extension to $\mathbb{R}^\infty$ and $\mathbb{R}^T$ 285
A.3 Expectations 287
A.3.1 Expectation of a random variable 287
A.3.2 Expectations and the space $\mathscr{L}^p(\Omega, \mathscr{F}, P)$ 288
A.3.3 Conditional distribution and conditional expectation 288
A.4 Convergence 291
A.4.1 Stochastic convergence 291
A.4.2 The Borel-Cantelli lemmas 292
A.4.3 Criteria for convergence with probability one 292

A.4.4 Criteria for convergence in quadratic mean 294
A.4.5 Criteria for convergence in probability 295
A.5 Characteristic functions 296
A.5.1 The characterstic function 296
A.5.2 The inversion theorems 298
A.5.3 Plancherel's and Parseval's relations 300
A.6 Hilbert space and random variables 301
A.6.1 Hilbert space and scalar products 301
A.6.2 Stochastic processes and Hilbert spaces 304
A.6.3 Projection and the linear prediction problem 306
Exercises 307

B Spectral simulation of random processes 309
B.1 The Fast Fourier Transform, FFT 309
B.2 Random phase and amplitude 310
B.3 Simulation scheme 311
B.4 Difficulties and details 312
B.5 Summary 313

C Commonly used spectra 315

D Solutions and hints to selected exercises 317
1 Some probability and process background 317
2 Sample function properties 319
3 Spectral and other representations 320
4 Linear filters – general properties 323
5 Linear filters – special topics 323
6 Ergodic theory and mixing 323
7 Vector processes and random fields 325
8 Level crossings and excursions 325
A Some probability theory 326

Bibliography 327

Index 337

List of figures

1.1 Overview of the three types of worlds in which our processes live. 2

1.2 Top: Processes with very narrow band spectrum, oscillating covariance function, and "fading" sample function; Middle: Realistic ocean waves with moderate band width wave spectrum; Bottom: low frequency white noise spectrum. The spectra are illustrated by their one-sided spectral densities. 16

1.3 Three particle paths in Perrin's experiment. 24

2.1 Successive approximations with piecewise linear functions. 42

3.1 Drastic effect of undersampling a stationary process with sampling interval d larger than the Nyquist frequency; left: spectrum of sampled sequence, right: sampled sequence. 85

3.2 Two types of periodogram. Left: The tone "a," 220[Hz], together with overtones, played on a cello, sampled 8192 times per second; $N = 9288$. Right: Water waves outside the African west coast, sampled four times per second; $N = 9524$. 99

3.3 Realization of the Ornstein-Uhlenbeck process with different values of α. 104

4.1 Input $x(t)$ and output $y(t)$ in an exponential smoother (RC-filter). 128

4.2 A linear oscillator. 130

4.3 A resonance circuit. 131

5.1 Gaussian processes, their Hilbert transforms, and envelopes. Left: A process with Pierson-Moskowitz broad band spectrum. Right: A process with narrow triangular spectrum over $(0.8, 1.2)$. 158

7.1 Illustration of in-phase and out-of-phase correlations for frequency ω. 212
7.2 Left: Level curves for directional spectrum with frequency dependent spreading. Right: Level curves for simulated Gaussian space sea. 232

8.1 Upcrossing occurs between t and $t+\Delta t$ when $x(t)$ is between $u-x'(t)\Delta t$ and u. The probability is the integral of the joint density $f_{x(t),x'(t)}(u,z)$ over the dashed area. 242
8.2 Excursions above $u=1.5$, translated to the origin, contribute to the distribution $P^{1.5}(\cdot)$. 249
8.3 Density (8.37) of height of local maxima in a standard normal process; spectral parameter $\varepsilon=0$ (Rayleigh) to $\varepsilon=1$ (normal). 259
8.4 Probability densities for excursions above $u=-1,0,1,2$ for process with North Sea wave spectrum JONSWAP. 264
8.5 Probability density for crest period T_c and crest height A_c for excursions above the mean level for a process with North Sea JONSWAP spectrum together with 481 observed cycles. 267
8.6 Moderately narrow JONSWAP spectrum, and the average number of waves per envelope upcrossing as function of level. 271

A.1 The random variable x maps outcomes ω in the sample space Ω on to the real line $\mathbb{R}$; $A_a=\{\omega;x(\omega)\le a\}$. 281
A.2 Sub-σ-fields generated by random variables x and y on Ω. 282

Preface

This book has grown out of my own experiences as teacher and researcher at a department in mathematical statistics with responsibilities both to an engineering and a science community. The spirit of the text reflects those double responsibilities.

The background

The book *Stationary and Related Stochastic Processes* [35] appeared in 1967. Written by Harald Cramér and M.R. Leadbetter, it drastically changed the life of PhD students in mathematical statistics with an interest in stochastic processes and their applications, as well as that of students in many other fields of science and engineering. Through that book, they got access to tools and results for stationary stochastic processes that until then had been available only in rather advanced mathematical textbooks, or through specialized statistical journals. The impact of the book can be judged from the fact that still, after almost fifty years, it is a standard reference to stationary processes in PhD theses and research articles.

Even if many of the more specialized results in the Cramér-Leadbetter book have been superseded by more general results, and simpler proofs have been found for some of the statements, the general attitude in the book makes it enjoyable reading both for the student and for the teacher. It will remain a definite source of reference for many standard results on sample function and crossings properties of continuous time processes, in particular in the Gaussian case. Unfortunately, the book only appeared in a first edition, and it was out of print for many years. The Dover reprint from 2004 filled that gap.

Another book, that at its time created a link between the mathematical theory of stationary processes and their engineering use, is the book *An Introduction to the Theory of Stationary Random Functions*, by A.M. Yaglom from 1962, [124]. The idea of bridging the gap between a strict mathematical treatment and the many practically useful results in the theory of stationary processes is even more explicitly expressed in the later, two-volume work by the same author, *Correlation Theory of Stationary and Related Random*

Functions from 1987, [125]. That work, with its extensive literature survey, presents a rather complete overview of correlation and spectrum theory for stationary processes at its time, with lots of extensions, references, and historical notices.

The third influential book is *Probability*, by Leo Breiman [19], also long out of print until it was reprinted in 1992. In the preface of that book, the author mentions the two-armed character of probability theory. The right arm is rigorous mathematics, based on measure theory; the left arm sets probability in action with gambling, motions of physical particles, and, now, almost fifty years later, remote sensing interpretation and systems control.

This book is for the left handed student, who wants to learn what the right handed side of stationary process theory has to offer, without going very far into the mathematical details! It is my hope that also the right handed probabilist could appreciate some of the fascination that lies in the potential of probabilistic thinking in engineering and science. The Preface to Yaglom's [125], Volume II, expresses these thoughts in a judicious way.

This text has grown out from a series of PhD courses on *Stationary stochastic processes*, which have been given at the Department of Mathematical Statistics at Lund University. Previous versions of the text have also been used in Stockholm, Trondheim, and Umeå. The audiences have consisted of Masters and PhD students in Mathematics and Statistics, but also to a large extent of PhD students in Automatic control, Electrical engineering and Signal processing, Image analysis, and Economics. As far as possible, I have tried to incorporate the links between these concrete topics and the theoretical tools developed.

I am grateful to three decades of PhD students who have followed the courses in Lund and elsewhere. Their comments and suggestions have removed many unclear and obscure points and statements. Their demands have also had the result that the text has grown, with new topics along with the increasing interest in advanced statistical tools in applications.

The text

The text is intended for a "second course" in stationary processes, and the material has been chosen to give a fairly broad overview of the theory behind widely scattered applications in engineering and science. The reader I have in mind has typically some experience with stochastic processes and has felt an urge to know more about "what it really is" and "why."

The emphasis is on second order properties and Fourier methods. Neither time series analysis nor Markov assumptions play any prominent role.

Since the days of the three texts mentioned in the introduction, applications of second order stationary processes have expanded enormously. The basic theory remains, but the applications, in particular new measurement techniques, have inspired much new theory. This is certainly true for spatial and spatio-temporal models, counting processes in space and time, extremes and crossings, and, not least, on computational aspects.

The aim is to provide a reasonably condensed presentation of sample function properties, spectral representations for stationary processes, and fields, including a portion of stationary point processes. One of the most important applications of second order process theory is found in linear filtering theory. The material on linear filters is divided into two chapters, one with a general technique of how to use the spectral representations and one on three important applications.

One of the most difficult topics to explain in elementary courses on stationary processes is ergodicity. I have tried to give both a mathematical and a statistical view with the main aim to couple mathematics and probability to statistics and measurements.

A chapter on crossing problems and extremes with an introduction to some of its applications has grown out from my own research interests.

Some knowledge of the mathematical foundations of probability helps while reading the text; I have included some of it in a probability appendix, together with the existence and basic convergence properties, as well as some Hilbert space concepts. There is also an appendix on how to simulate stationary stochastic processes by spectral methods and the FFT algorithm.

Some suggestions for reading

The reader who wants to design a course on parts of the text may make use of the following selections:

- For a general and concrete level:
  - Section 1.1 (Section 1.2), Sections 1.3–1.5
  - Section 2.1
  - Sections 3.2–3.4
  - Chapter 4 except Section 4.4
  - Section 5.2
  - Sections 7.1.1, 7.2.1
  - Section 8.1.1
- For specialized and concrete levels:
  - Section 3.5

  - Section 4.4
  - Sections 5.1 and 5.3
  - Chapter 7
  - Chapter 8
  - Appendix B
- For a more abstract level one could add:
  - Section 1.2, Appendix A, plus a book on general probability theory
  - Sections 2.2–2.5
  - Section 3.1
  - Chapter 6

The website http://www.crcpress.com/product/isbn/97814665577796 contains corrections and updates, as well as worked applications.

MATLAB® and Simulink are registered trademarks of The Mathworks, Inc. For product information, please contact:

The Math Works, Inc.
3 Apple Hill Drive
Natick, MA 01760-2098 USA
Tel: 508-647-7000
Fax: 508-647-7001
E-mail: info@mathworks.com
Web: www.mathworks.com

Acknowledgments

To all those students who have been confronted with early versions of this text, as well as the colleagues who have read, lectured from, and commented on the text, I express my great gratitude. I have felt as we have been fellow students of the many facets of the topic. I mention in particular Anastassia Baxevani, Rikard Berthilsson, Klas Bogsjö, Jonas Brunskog, Halfdan Grage, Peter Gustafsson, Oskar Hagberg, Lars Holst, Ulla Holst, Henrik Hult, Pär Johannesson, Timo Koski, Torgny Lindström, Karl-Ola Lundberg, Ulla Machado, Anders Malmberg, Dan Mattsson, Sebastian Rasmus, Bengt Ringnér, Tord Rikte, Igor Rychlik, Jesper Rydén, Mikael Signahl, Eva Sjö, Martin Sköld, Martin Svensson, Jörg Wegener, Linda Werner, and Sofia Åberg.

During the preparation of the final text I have had fruitful discussions with David Bolin, Finn Lindgren, Johan Lindström, Jimmy Olsson, Krzysztof Podgórski, Magnus Wiktorsson; thank you!

But it all goes back to Harald Cramér and Ross Leadbetter!

Georg Lindgren
Lund

This page intentionally left blank

List of notations

Roman letters

A, A_k	random amplitudes
$A_{\mathbf{s}}(\mathbf{v})$	set of functions $y(s)$ with $y(s_j) \leq v_j$
$A_{xy}(\omega) = \lvert f_{xy}(\omega)\rvert$	cross-amplitude spectrum
$A(z)$	generating polynomial
$A \Delta B$	$(A-B) \cup (B-A)$, difference between sets
a_T, b_T	inverse scale, and location, parameter in asymptotic extreme value distribution
$BB_T(t)$	Brownian bridge
$\mathscr{B}, \mathscr{B}_n, \mathscr{B}_\infty, \mathscr{B}_T$	Borel fields in $\mathbb{R}, \mathbb{R}^n, \mathbb{R}^\infty, \mathbb{R}^T$
C	generic constant
$C(x_1, x_2)$	$E((x_1 - m_1)\overline{(x_2 - m_2)})$, covariance of complex variables
$C(z)$	generating polynomial
$\mathbb{C}$	space of continuous functions
$c_{xy}(\omega)$	co-spectrum density
$D(u_n), D'(u_n)$	mixing and non-clustering conditions in asymptotic extreme value theory
$d(\mathbf{s}, \mathbf{t})$	$\sqrt{E(\lvert x(\mathbf{s}) - x(\mathbf{t})\rvert^2)}$, distance measure
$\mathrm{d}Z(\omega)$	differential spectral component
$\mathrm{d}\widetilde{Z}(\omega) = \frac{\mathrm{d}Z(\omega)}{f(\omega)}$	normalized differential spectral component
$E(x\lvert y), E(x\lvert \mathscr{A})$	conditional expectation given y or given $\mathscr{A}$
$e(\mathbf{s}, \mathbf{t})$	Euclidian distance
$F_{t_1 \cdots_n}(b_1, \ldots, b_n)$	n-dimensional distribution function for $x(t_1), \ldots, x(t_n)$
$\mathbf{F} = \{F_{\mathbf{t}^n}\}$	family of finite-dimensional distribution functions
$F(\omega), F_x(\omega)$	spectral distribution function
F^{ac}, F^d, F^s	absolutely continuous, discrete, and singular part of distribution function
$\widetilde{F}(\omega)$	$(F(\omega) + F(\omega - 0))/2$, midpoint at spectral jump
$F_{xy}(\omega), f_{xy}(\omega)$	cross-spectral distribution and density functions
$\mathscr{F}_0, \mathscr{F}$	field and σ-field of events

f	friction coefficient
$f_{x\|y=v}(u)$	conditional density function
$f(\omega), f_x(\omega)$	spectral density functions for stationary process
$f_x^{(d)}(\omega)$	spectral density after sampling with sampling distance d
$f_x^{\text{time}}(\omega)$	time wave spectrum
$f_x^{\text{space}}(\kappa)$	space wave spectrum
$f_{PM}(\omega), f_J(\omega)$	Pierson-Moskowitz and JONSWAP spectral densities
$f(\omega,\theta)$	$f_x^{\text{time}}(\omega)g(\omega,\theta)$, spectral density for random field with directional spreading
$f_c(\omega)$	$\frac{\lambda}{2\pi}+f(\omega)$, Bartlett spectrum for counting process
$f_n^{\text{per}}(\omega)$	periodogram computed from data
g	constant of gravity
$(g,h)_{\mathscr{H}(F)}$	$\int g(\omega)\overline{h(\omega)}\,\mathrm{d}F(\omega)$, inner product in Hilbert space $\mathscr{H}(F)$
$g(\omega)$	$\int e^{-i\omega u}h(u)\,\mathrm{d}u$ frequency response function
$g(\omega,\theta)$	directional spreading function
H	Hurst parameter, index of fractional Brownian motion
$H(\varepsilon)$	entropy
h	water depth
$\mathscr{H}, \mathscr{H}(\Omega)$	Hilbert space of random variable with finite variance
$\mathscr{H}(F)$	Hilbert space of functions with $\int \|g(\omega)\|^2\,\mathrm{d}F(\omega)<\infty$
$\mathscr{H}(x)$	$\mathscr{S}(x(s);s\in T)$, Hilbert space generated by $x(t), t\in T$
$\mathscr{H}(x,t)$	$\mathscr{S}(x(s);s\le t)$, Hilbert space generated by $x(s), s\le t$
$h(u)$	impulse response function
$\mathscr{I}$	field of unions of a finite number of intervals, or σ-field of invariant sets
$J_m(t)$	Bessel function of the first kind of order m
$K_\nu(t)$	Bessel function of the second kind of order ν
$\mathscr{M}_a^b$	σ-field generated by $x(t), a\le t\le b$
N	Avogadro's number
$N(t)$	Poisson process
$\widetilde{N}(I)$	normalized point process of crossings
$N_I(x,u), N_I^+(x,u)$	number of u-crossings/upcrossings by $x(t), t\in I$
$\mathbb{N}$	the set of natural numbers
P	general probability measure
$P^{hw}(A)$	horizontal window conditioned probability after level crossing

$P^u(A)$	distribution of Slepian model for process after u-upcrossing
$P_j^{\max}(A)$	distributions for Slepian models after local maximum, $j = 1,2$
$P_x, P_y, \ldots$	probability measure for processes $x(t), y(t), \ldots$
$P_x^{(n)}, P_y^{(n)}, \ldots$	n-dimensional probability measure for $x(t), y(t), \ldots$
$p^u(z)$	Rayleigh density for derivative at level upcrossing in Gaussian process
$q_{xy}(\omega)$	quadrature spectrum
R	Boltzmann's constant
$R(t)$	envelope process
$\mathbb{R}, \mathbb{R}^n, \mathbb{R}^\infty, \mathbb{R}^T$	real line, n-dimensional real space, space of infinite real sequences, space of real function defined on T
$r(s,t)$	$C(x(s), x(t))$ covariance function
$r(t)$	$C(x(s+t), x(s))$ covariance function for stationary process or field
$r_{xy}(t)$	cross-covariance between $x(s+t)$ and $y(s)$
$r_N^*(t)$	covariance estimate based on the periodogram
$r_{\Delta_1}(s,t), r_{\Delta_2}(s,t)$	covariance functions for Gaussian residual in Slepian models after local maximum
$r_\kappa(s,t)$	covariance function for Gaussian residual in Slepian model after level upcrossing
$\mathscr{S}(L)$	space spanned by random variables in L
$\mathscr{S}(x, \varepsilon)$	ε-surrounding around point or function x
$\operatorname{sinc}(x) = \frac{\sin x}{x}$	sinc function
T	parameter space, often $\mathbb{R}$ or $\mathbb{R}^+$, or absolute temperature
$T(\omega)$	transformation of a probability space
$T_2 = 2\pi\sqrt{\omega_0/\omega_2}$	mean period
$t_n^{k)} = k/2^n$	dyadic number
$(u,v)_{\mathscr{H}(x)}$	$E(u\overline{v})$, inner product in Hilbert space $\mathscr{H}(x)$
$V_f(I)$	total variation of $f(t), t \in I$
$w(t)$	Wiener process
$w'(t)$	Gaussian white noise
$w(\tau)$	covariance intensity function
$w_c(\tau) = \lambda\delta_0(\tau) + w(\tau)$	complete covariance intensity function
$\widehat{x}(t)$	Hilbert transform of $\{x(t), t \in \mathbb{R}\}$
$\mathbb{Z}$	the set of integer numbers
$Z(\omega), Z_j(\omega)$	spectral processes

Greek letters

$\alpha = \omega_2/\sqrt{\omega_0\omega_4}$	spectral width parameter
γ	peakedness parameter in JONSWAP spectrum
$\gamma(\mathbf{u},\mathbf{v})$	semi-variogram
$\Delta F_\omega, \Delta F_k$	spectral distribution jump at ω and ω_k
$\Delta Z_\omega, \Delta Z_k$	spectral process jump at ω and ω_k
δ_{jk}	Kronecker delta
$\delta(t), \delta_0(t), \delta_\tau(t)$	Dirac delta functions
$\varepsilon = \sqrt{1-\alpha^2}$	spectral width parameter
$\zeta = c/(2\sqrt{mk})$	relative damping in linear oscillator
θ	wave direction in random field
κ	space frequency, wave number
κ_k	spectral moment in space
$\kappa^2_{xy}(\omega)$	squared coherence spectrum
λ	intensity in point process
λ_k	eigenvalue to a covariance operator
$\mu^+ = \frac{1}{\pi}\sqrt{\omega_2/\omega_0}$	mean frequency
$\mu(u)$	mean number of u-crossings per time unit
$\mu^+(u)$	mean number of u-upcrossings per time unit
$\mu_{\min} = \mu_{\max}$	mean number of local extremes per time unit
$\xi_u(t)$	Slepian model for process after u-upcrossing
$\xi_1^{\max}(t), \xi_2^{\max}(t)$	Slepian models for process after local maximum
$\rho^2 = \omega_1^2/\omega_0\omega_2$	spectral width parameter
$\rho_{xy}(\omega)$	in-phase correlation in vector processes
$\widetilde{\rho}_{xy}(\omega)$	out-of-phase correlation in vector processes
$\mathbf{\Sigma}$	covariance matrix
$\sigma^2/2\pi$	constant spectral density of white noise
$\sigma(x)$	σ-field generated by random variable
$\Phi_{xy}(\omega)$	phase spectrum
ϕ, ϕ_k	random phases
$\phi_k(t)$	eigenfunction to a covariance operator
ϕ, Φ	standard normal density and distribution functions
$\varphi(s), \varphi(x \mid y)$	characteristic function and conditional expectation
χ_A	indicator function for event A
ψ	phase parameter in linear oscillator
$\psi_\varepsilon(t)$	characteristic function
Ω	sample space
$(\Omega, \mathscr{F})$	measurable space, sample space with σ-field
$(\Omega, \mathscr{F}, P)$	probability space

ω	frequency or element in sample space
$\omega_0 = \sqrt{k/m}$	resonance frequency in linear oscillator
ω_2	$-r''(0) = V(x(t))$, second spectral moment
ω_k	discrete frequency, or spectral moment $\int \lvert\omega\rvert^k \, \mathrm{d}F(\omega)$
$\omega_{\boldsymbol{\alpha}} = \Pi_k \omega_k^{\alpha_k}$	multiple spectral moment
$\omega_x(h)$	continuity modulus of function x
$\omega_{x,d}(h)$	continuity modulus of function x in d-distance

Special symbols

$\overset{a.s.}{\to}$	almost sure convergence, with probability one
$\overset{q.m.}{\to}$	quadratic mean convergence
$\overset{P}{\to}$	convergence in probability
$\overset{\mathscr{L}}{\to}$	convergence in distribution
$x \overset{\mathscr{L}}{=} y$	x and y have the same distribution

This page intentionally left blank

Chapter 1

Some probability and process background

This introductory chapter gives examples of some of the main concepts in stationary processes theory that will be dealt with in detail in later chapters. The intention here is not to dwell on the precise mathematical properties, but to point to the statistical content and interpretation of the mathematical constructions, and on the driving ideas. The mathematical details are instead presented in Appendix A, and refreshed when needed.

First of all, we deal with the finite-dimensional distribution functions (d.f.), which uniquely define probabilities for a sufficiently rich family of events, namely events that can be identified through process values at discrete sets of times. In particular, they allow us to find conditions for sample function continuity and differentiability and other properties that are essential in statistical model building. This is the topic of Chapter 2.

Another central topic in the book is the covariance properties of stationary processes and their spectral representations, dealt with in Chapter 3. These also form the tools in the chapters on filter manipulations of stationary processes, Chapters 4 and 5. Prediction and limiting behavior are the topics of Chapter 6. Extensions to multivariate models and stochastic fields are treated in Chapter 7, and the final Chapter 8 deals with techniques and important applications of extremes and crossings.

This chapter ends with four examples of path breaking innovative use of stochastic models in physics and engineering, innovations which helped to shape the presented tools and techniques. The examples are the Wiener process as a model for Brownian motion, stationary processes as models for noise in electronic communication systems, stochastic wave models used in ocean engineering, and finally the link between statistical inference and stochastic processes as a basis for signal processing.

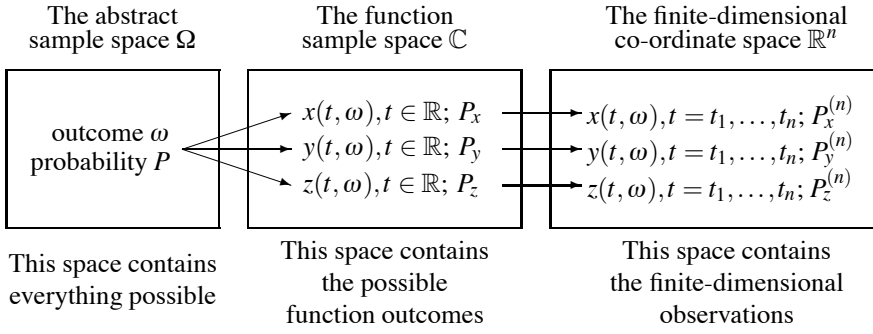


Figure 1.1: *Overview of the three types of worlds in which our processes live.*

1.1 Sample space, sample function, and observables

Stochastic processes are often called random functions; these two notions put emphasis on two different aspects of the theory, namely,

- stochastic processes as families of infinitely many random variables on the same sample space, usually equipped with a fixed probability measure,
- stochastic processes as a means to assign probabilities to sets of functions, for example some specified sets of continuous functions, or sets of piecewise constant functions with unit jumps.

These two aspects of stochastic processes can be illustrated as in Figure 1.1, corresponding to an experiment where the outcomes are continuous functions.

The figure illustrates the three levels of abstraction and observability for a random experiment. To be concrete, think of an experiment controlling the steering of a ship. The sample space Ω can be envisioned to describe the conditions that the ship will meet on the ocean, the wind and wave climate over the lifetime of the ship, the actual weather at a certain trip, etc. The function sample space $\mathbb{C}$ contains the different characteristics that physically characterize the movements of the ship and the steering process, while the co-ordinate space $\mathbb{R}^n$ describes the finite number of measurements that can be taken on these signals.

The general sample space Ω is an abstract set that contains all the possible outcomes of the experiment that can conceivably happen – and it may contain more. A probability measure P is defined on Ω that assigns probabilities to all interesting subsets – we need only one single probability measure to describe our whole world. We may choose to measure many different characteristics, with different statistical distributions, but they are all uniquely defined by the single underlying "world" probability measure P.

During the experiment one can record the time evolution of a number of things, such as rudder angle, which we call $\{x(t), t \in \mathbb{R}\}$, ship head angle, $\{y(t), t \in \mathbb{R}\}$, and roll angle $\{z(t), t \in \mathbb{R}\}$. Each observed function is an observation of a continuous random process. In the figure, the randomness is indicated by the dependence of the experiment outcome ω. The distributions of the different processes are P_x, P_y, P_z – we need one probability measure for each of the phenomena we have chosen to observe.[1]

In practice, the continuous functions are sampled in discrete time steps, $t = t_1, \ldots, t_n$, resulting in an observation vector, $(x_1, \ldots, x_n)$, with an n-dimensional distribution, $P_x^{(n)}$, etc. This familiar type is illustrated in the third box in Figure 1.1.

Since we do not always want to specify a finite value for n, the natural mathematical model for the practical situation is to replace the middle box, the sample space $\mathbb{C}$ of continuous functions, by the sample space $\mathbb{R}^\infty$ of infinite sequences of real numbers $(x_0, x_1, \ldots)$. This is close, as we shall see later, really very close, to the finite-dimensional space $\mathbb{R}^n$, and mathematically not much more complicated.

Taking the set $\mathbb{C}$ of continuous functions as a sample space and assigning probabilities P_x, etc., on it, is not as innocent as it may sound from the description above. Chapter 2 deals with conditions that guarantee that a stochastic process is continuous, i.e., has continuous sample functions. In fact, these conditions are all on the finite-dimensional distributions.

SUMMARY: *The abstract sample space Ω for an experiment contains everything that can happen and is therefore very complex and detailed. Each outcome $\omega \in \Omega$ is unique, and we need only one comprehensive probability measure P to describe every outcome of the experiment.*

An experiment is a way to "observe the world." The function (sequence) sample space $\mathbb{C}$ ($\mathbb{R}^\infty$) is simple. It can be used as sample space for a specified experiment for which the result is a function or a sequence of numbers. We have to define a specific probability measure for each experiment.

[1] The symbol ω is here used to represent the elementary experimental outcome, a practice that is standard in probability theory. In most parts of this book, ω will stand for (angular) frequency; no confusion should arise from this.

1.2 Random variables and stochastic processes

1.2.1 Probability space and random variables

A *probability space* $(\Omega, \mathscr{F}, P)$ is a triple consisting of: a sample space Ω of outcomes, a family $\mathscr{F}$ of events, i.e., subsets of Ω, and a probability measure P, that assigns probabilities to the events in $\mathscr{F}$. The family $\mathscr{F}$ may be chosen in many different ways; it may contain just a few subsets, or it may contain many. It is even possible to let $\mathscr{F}$ contain *all* subsets of Ω.[2] The obligatory properties of the family $\mathscr{F}$ are that it contains Ω and is closed under formation of complements, unions, and intersections of countably many of its elements. A family of subsets with these properties is called a σ*-field*.

A probability measure assigns probabilities to the events in $\mathscr{F}$ in such a way that the basic rules for adding probabilities, usually called the *Kolmogorov's axioms of probablities*, are fulfilled. The exact definitions of σ-field and the summation rules are formulated in Definition A.1 in Appendix A.1.

The link between the sample space Ω and the observable world is the random variable, a real valued function $\Omega \ni \omega \mapsto x(\omega) \in \mathbb{R}$. In order that x be called random variable it is essential that the *distribution* of x is well defined, i.e., that the *event* "$x \leq b$" has a probability for each value of $b \in \mathbb{R}$. This means that the set of outcomes $\{\omega \in \Omega; x(\omega) \leq b\}$ must be one of the sets in $\mathscr{F}$, those which have been assigned probability by P. Functions that fulfill this requirement are called *measurable*, or, more precisely, *Borel measurable*.

It is of course possible to define many random variables, $x_1, \ldots, x_n$, on the same probability space $(\Omega, \mathscr{F}, P)$. Since P assigns a probability to each of the individual events

$$\text{“}x_1 \leq b_1\text{”}, \ldots, \text{“}x_n \leq b_n\text{”}$$

it also assigns a probability to the joint occurrence of them all; this is by the property of the σ-field $\mathscr{F}$ that it contains intersections. The result is a multivariate random variable, $\mathbf{x} = \{x_k\}_{k=1}^n$ with an n-dimensional distribution.

In Appendix A.1.1 we explain what more events are given probabilities in this way. These events are the *Borel sets* in $\mathbb{R}^n$ and the family is called the *Borel* σ*-field*, denoted $\mathscr{B}_n$. Examples of such events are those defined by two-sided inequalities, like the half-open rectangle $a_1 < x_1 \leq b_1, a_2 < x_2 \leq b_2$, unions of rectangles, complements of unions, and many more. The Borel σ-

[2] If $\mathscr{F}$ contains all subsets of Ω every probability measure has to be discrete.

field is a very rich family of events, which all get probabilities defined by the finite-dimensional distribution functions.

1.2.2 Stochastic processes and their finite-dimensional distributions

We are now ready to define stochastic processes in general. There is no difficulty in considering an infinite sequence of random variables $\mathbf{x} = \{x_k\}_{k=1}^{\infty}$, on a single probability space $(\Omega, \mathscr{F}, P)$. One can even consider more than countably many random variables at the same time. Then it is common to index the variables by the letter t instead of k, and let the *parameter* t take values in a general *parameter space* T. Usually one thinks of t as "time" and T as the whole or part of the real line.

Thus, we can consider a family of functions, $\{x(t,\omega) \in \mathbb{R}\}_{t\in T}$, where each $x(t) = x(t,\cdot)$ is a random variable, i.e., a measurable function from Ω to $\mathbb{R}$. Hence, it has a distribution with a distribution function on $\mathbb{R}$, which we denote $F(\cdot;t)$, i.e.,

$$F(b;t) = \text{Prob}(x(t) \le b) = P(\{\omega; x(t,\omega) \le b\}).$$

Taking several variables, observed at times $t_1, \ldots, t_n$, we get an n-variate random variable $(x(t_1), \ldots, x(t_n))$ with an n-variate distribution in $\mathbb{R}^n$,

$$\begin{aligned} F_{t_1 \cdots t_n}(b_1, \ldots, b_n) &= \text{Prob}(x(t_1) \le b_1, \ldots, x(t_n) \le b_n) \\ &= P\left(\cap_1^n \{\omega; x(t_k,\omega) \le b_k\}\right). \end{aligned} \tag{1.1}$$

Write $F_{\mathbf{t}^n}$ for the n-dimensional distribution function of $(x(t_1), \ldots, x(t_n))$. We summarize the terminology in a formal, but simple, definition.

Definition 1.1. *Let T be a discrete or continuous real parameter set. A stochastic process $\{x(t), t \in T\}$ indexed by the parameter $t \in T$ is a family of random variables $x(t)$ defined on one and the same probability space $(\Omega, \mathscr{F}, P)$. In other words, a stochastic process is a function*

$$T \times \Omega \ni (t,\omega) \mapsto x(t,\omega) \in \mathbb{R},$$

such that, for fixed $t = t_0$, $x(t_0,\cdot)$ is a random variable, i.e., a Borel measurable function, $\Omega \ni \omega \mapsto x(t_0,\omega) \in \mathbb{R}$, and, for fixed $\omega = \omega_0$, $x(\cdot,\omega_0)$ is a function $T \ni t \mapsto x(t,\omega_0) \in \mathbb{R}$, called a sample path or sample function.

We now have the following concepts at our disposal in the three scenes in Figure 1.1:

sample space	events	probability
abstract space: Ω	σ-field $\mathscr{F}$	P
space of continuous functions: $\mathscr{C}$	?	?
real sequences: $\mathbb{R}^\infty$	?	?
real vectors: $\mathbb{R}^n$	Borel sets: $\mathscr{B}_n$	$P_{\mathbf{t^n}}$ from n-dimensional d.f.'s $F_{\mathbf{t^n}}$
real line: $\mathbb{R}$	Borel sets: $\mathscr{B}$	P from a d.f. F

In the table, the ? indicate what we yet have to define – or even show existence of – to reach beyond the elementary probability theory, and into the world of stochastic processes.

1.2.3 The distribution of random sequences and functions

In the previous section we discussed finite-dimensional distributions for one, two, and (finitely) many random variables, how these distributions are defined by the finite-dimensional distribution functions, and what finite-dimensional events get their probabilites defined.

Our aim now is to find the events in $\mathbb{R}^\infty$ (= all real sequences) and $\mathbb{R}^T$ (= all real functions defined on T), and see how one can define a probability measure for these. When this is done, it is legitimate to talk about *the distribution of an infinite random sequence or function*, and of the probability that an experiment results in a sample function that satisfies some rather general requirement, e.g., that its maximum is less than a given constant.

It was this step, from probabilities for one-dimensional or n-dimensional real sets and events, to probabilities in $\mathbb{R}^\infty$ and $\mathbb{R}^T$, that was made axiomatic in A.N. Kolmogorov's celebrated *Grundbegriffe der Wahrscheinlichkeitsrechnung* from 1933 [71].

Generalized rectangles, intervals, and the Borel field

The basic requirement on the events in $\mathbb{R}^\infty$ or $\mathbb{R}^T$ is that they should not be simpler than the events in the finite-dimensional spaces $\mathbb{R}^n$, which means that if an event $B_n \in \mathscr{B}_n$ in $\mathbb{R}^\infty$ is expressed by means of a finite set of random variables $x_1,\ldots,x_n$, then it should be an event also in the space $\mathbb{R}^\infty$. Now, such an event can be written

$$\begin{aligned}\{\mathbf{x} = (x_1,x_2,\ldots) \in \mathbb{R}^\infty; (x_1,x_2,\ldots,x_n) \in B_n\} &= B_n \times \mathbb{R} \times \mathbb{R} \times \ldots \\ &= B_n \times \mathbb{R}^\infty.\end{aligned}$$

A set of this form is called a *generalized rectangle in* $\mathbb{R}^\infty$.

Similarly, we define a generalized rectangle in $\mathbb{R}^T$ as a set defined by a condition on any finite set of $x(t)$-values,

$$\{\mathbf{x} \in \mathbb{R}^T; (x(t_1), x(t_2), \ldots, x(t_n)) \in B_n\}.$$

Hence, we have to require that the σ-field of events contains at least all generalized rectangles. The natural event field is exactly the smallest σ-field which contains all such sets; cf. Example A.2 in Appendix A.1. This σ-field is denoted $\mathscr{B}_\infty$ and $\mathscr{B}_T$, respectively, and is, naturally, called the *Borel field* in $\mathbb{R}^\infty$ or $\mathbb{R}^T$. Symbolically, we can write, for sequences,

$$\mathscr{B}_\infty = \sigma\left(\cup_{n=1}^\infty (\mathscr{B}_n \times \mathbb{R}^\infty)\right).$$

A particularly simple form of rectangles is the *intervals*. These are, in $\mathbb{R}^\infty$, the sets of form

$$I = (a_1, b_1] \times (a_2, b_2] \times \ldots \times (a_n, b_n] \times \mathbb{R}^\infty,$$

where each $(a_j, b_j]$ is a real half-open interval. Thus, the half-open intervals in $\mathbb{R}^\infty$ and $\mathbb{R}^T$ are defined by inequalities like

$$a_1 < x_1 \leq b_1, a_2 < x_2 \leq b_2, \ldots, a_n < x_n \leq b_n. \tag{1.2}$$

$$a_1 < x(t_1) \leq b_1, a_2 < x(t_2) \leq b_2, \ldots, a_n < x(t_n) \leq b_n. \tag{1.3}$$

Sets which are unions of a finite number of intervals are important; they form a field, which we denote $\mathscr{I}$. The σ-field generated by $\mathscr{I}$ is exactly $\mathscr{B}_\infty$, and $\mathscr{B}_T$, respectively, in $\mathbb{R}^\infty$ and $\mathbb{R}^T$. The reader should draw a picture of some sets in the field $\mathscr{I}$ in $\mathbb{R}^2$.

The sample spaces we are mostly going to work with are $\mathbb{R}^\infty$ and $\mathbb{R}^T$, infinite real sequences, and real functions defined on T, and the Borel σ-fields of events to be given probabilities are $\mathscr{B}_\infty$ and $\mathscr{B}_T$, respectively, generated by the finite-dimensional intervals.

These spaces are important since they provide us with a natural world of objects for stochastic modeling.

Definition 1.2 (The co-ordinate process)**.** *The simplest process on* $(\mathbb{R}^\infty, \mathscr{B}_\infty)$ *is the co-ordinate process, also known as "the canonical process,"* $\{x_n, n \in \mathbb{N}\}$ *defined by* $x_k(\omega) = \omega_k$*, on the outcome* $\omega = (\omega_1, \omega_2, \ldots)$.

Similarly, the simplest process on $(\mathbb{R}^T, \mathscr{B}_T)$ *is the co-ordinate process* $x(t,\omega) = \omega(t)$ *for* $\omega \in \mathbb{R}^T$.

Finding a probability measure

The next big step is to assign probabilities to the Borel sets in $\mathscr{B}_\infty$ and $\mathscr{B}_T$, and this can be done in either of two ways, from the abstract side or from the observable, finite-dimensional side. Going back to Figure 1.1 we shall fill in the two remaining question marks in the following scheme:

$$(\Omega, \mathscr{F}, P) \overset{x}{\mapsto} \left\{ \begin{array}{l} (\mathbb{R}^\infty, \mathscr{B}_\infty, ?) \\ (\mathbb{R}^T, \mathscr{B}_T, ?) \end{array} \right\} \leftarrow\!\shortmid (\mathbb{R}^n, \mathscr{B}_n, P_x^{(n)}), n = 1, 2, \ldots .$$

The mapping $\overset{x}{\mapsto}$ symbolizes that we have already defined a probability P on some measurable space $(\Omega, \mathscr{F})$ and want to define probabilities on a sample function space for a stochastic process x with discrete or continuous parameter. The mapping $\leftarrow\!\shortmid$ symbolizes that we shall define the probabilities from a family of finite-dimensional distributions on $(\mathbb{R}^n, \mathscr{B}_n), n = 1, 2, \ldots$ so that the co-ordinate process gets exactly these finite-dimensional distributions.

Definition from a random process

Suppose the probability space $(\Omega, \mathscr{F}, P)$ is given a priori, and a process x is defined on $(\Omega, \mathscr{F})$, either as $\{x_n, n \in \mathbb{N}\}$ or as $\{x(t), t \in T\}$. We shall use P on $(\Omega, \mathscr{F})$ to define the distribution of x on the sample function spaces $(\mathbb{R}^\infty, \mathscr{B}_\infty)$ or $(\mathbb{R}^T, \mathscr{B}_T)$. The following lemma may look complicated, but its content is simple. It says that there is a direct connection between the events in $\mathscr{B}_\infty$ or $\mathscr{B}_T$ and certain events in $\mathscr{F}$ on Ω.

Lemma 1.1. *For every set* $A \in \mathscr{B}_\infty$ *or* $A \in \mathscr{B}_T$*, the inverse set* $x^{-1}(A) \in \mathscr{F}$*:*

$$x^{-1}(A) = \{\omega \in \Omega; \mathbf{x}(\cdot, \omega) \in A\} \in \mathscr{F}.$$

Proof. Let $\mathscr{C}$ be the class of sets C in $\mathscr{B}_\infty$ or $\mathscr{B}_T$ such that $x^{-1} \in \mathscr{F}$. By the definition of a stochastic process, $\mathscr{C}$ contains all rectangles, and it is easy to show (do that as Exercise 1:3) that it is a σ-field. Since the Borel field is the smallest σ-field that contains the rectangles, the claim is shown. □

It is now easy to define probabilities on the sample function spaces. If x is a stochastic process, i.e., a function from Ω to $\mathbb{R}^\infty$ or $\mathbb{R}^T$, then a probability measure P_x is defined on $(\mathbb{R}^\infty, \mathscr{B}_\infty)$ and $(\mathbb{R}^T, \mathscr{B}_T)$, respectively, by

$$P_x(B) = P(x^{-1}(B)), \quad \text{for } B \in \mathscr{B}_\infty \text{ or } \mathscr{B}_T. \tag{1.4}$$

Thus, each stochastic process produces a probability measure P_x on $(\mathbb{R}^\infty, \mathscr{B}_\infty)$ or $(\mathbb{R}^T, \mathscr{B}_T)$.

Example 1.1 ("Random amplitude and phase")**.** *Here is a simple example of a stochastic process that is often used as a building block in more complicated models. Take two random variables A, ϕ, with $A \geq 0$ and ϕ uniformly distributed between 0 and 2π, and define a stochastic process $\{x(t), t \in \mathbb{R}\}$ as $x(t) = A\cos(t+\phi)$. This procedure defines the process x as a function from the sample space of (A,ϕ)-outcomes, $\Omega = [0,\infty) \times [0,2\pi] = \{(A,\phi)\}$, to the function space $\mathbb{R}^T$, with $T = \mathbb{R}$. Of course, only cosine functions with different amplitudes and phases can appear as sample functions, and the probability measure on $\mathbb{R}^T$ gives probability one to the subset of such functions. The finite-dimensional distributions of $x(t)$ can be calculated from the joint distribution of A and ϕ.*

Example 1.2 ("Poisson process")**.** *The sample functions of a common Poisson process are piecewise constant and increase in unit steps. It can be explicitly constructed from a sequence of independent exponential variables defining the time that elapses between two jumps, and the finite-dimensional distributions of the Poisson process can be calculated from the distribution of these intervals in $(R^\infty, \mathscr{B}_\infty)$.*

It is essential, in the two examples, that we already have a probability space that allows us to construct the amplitude and phase variables, and the inter-event times, respectively. It is also worth to notice, that the more complex a process we wanted to generate, the more complex did we choose the basic sample space. This causes no problem, since one can always expand the sample space to include the desired variability. There is no need to look for the *simplest* probability space for a particular application.

If we start with a huge sample space $(\Omega, \mathscr{F})$ with a single but complex probability measure P we can construct a rich variety of processes on the single probability space $(\Omega, \mathscr{F}, P)$, and, by (1.4), obtain the probabilities of

sets B in $\mathscr{B}_\infty$ or $\mathscr{B}_T$. We can also derive the finite-dimensional distributions. The negative side of this approach is that we don't know how to construct the measure P from experiments or from simple process models like Poisson or Gaussian models. We therefore usually rely on the other alternative to construct process distributions.

Definition from a family of finite-dimensional distributions:

The family $\{F_{\mathbf{t}^n}\}_{n=1}^{\infty}$ of finite-dimensional distributions is the family of functions

$$F_{t_1\cdots t_n}(b_1,\ldots,b_n) = \text{Prob}(x(t_1)\le b_1,\ldots,x(t_n)\le b_n),$$

for $t_j \in T, t_1 < t_2 < \ldots < t_n, n = 1,2,\ldots$. A family of distribution functions for a stochastic process has to be *consistent* to satisfy certain, trivial requirements.

Definition 1.3. *A family $\{F_{\mathbf{t}^n}\}_{n=1}^{\infty}$ of finite-dimensional distributions is called consistent if, for every n, and $t_1, b_1, \ldots, t_n, b_n$,*

$$\lim_{b_k\uparrow\infty} F_{t_1,\ldots,t_n}(b_1,\ldots,b_n) = F_{t_1,\ldots,\widetilde{t_k},\ldots,t_n}(b_1,\ldots,\widetilde{b_k},\ldots,b_n),$$

$$F_{t_1,t_2}(b_1,b_2) = F_{t_2,t_1}(b_2,b_1),\ \textit{etc.}$$

where the notation $\widetilde{\ }$ indicates that the variable is missing.

Suppose a family of finite-dimensional distribution functions,

$$\mathbf{F} = \{F_{\mathbf{t}^n}\}_{n=1}^{\infty}, t_k \in T,$$

$$F_{t_1\cdots t_n}(b_1,\ldots,b_n) = \text{Prob}(x(t_1)\le b_1,\ldots,x(t_n)\le b_n),$$

is given a priori, with one distribution for each n and for each set $(t_1,\ldots,t_n)$ of times. If the family is consistent (Definition 1.3), one can use it to define probabilities for all half-open n-dimensional intervals in $\mathbb{R}^\infty$ or $\mathbb{R}^T$, for example by, for $n = 1,2,\ldots$, taking (cf. (1.1))

$$P_{\mathbf{x}}((a_1,b_1]\times(a_2,b_2]\times\ldots\times(a_n,b_n]\times\mathbb{R}^\infty) = P_n((a_1,b_1]\times(a_2,b_2]\times\ldots\times(a_n,b_n]).$$

This will give us a unique probability measure for each n and each n-dimensional selection of time points $(t_1,\ldots,t_n)$.

Theorem 1.1 (Existence of stochastic process; Kolmogorov, 1933)**.** *To every consistent family of finite-dimensional distribution functions,* $\mathbf{F} = \{F_{\mathbf{t^n}}\}_{n=1}^{\infty}$, $t_k \in T$, *there exists one and only one probability measure* P_x *on* $(\mathbb{R}^T, \mathscr{B}_T)$ *with*

$$P_x(x(t_1) \le b_1, \ldots, x(t_n) \le b_n) = F_{t_1 \cdots t_n}(b_1, \ldots, b_n), \text{ for } n = 1, 2, \ldots.$$

In Appendix A.2 we present an outline of the proof of the existence theorem. The main steps are to show that the finite dimensional distributions give a countably additive probability measure on the field $\mathscr{I}$ of finite unions of intervals, defined for different times $(t_1, \ldots, t_n)$. By the extension property of probability measures on fields, Theorem A.1, one can then conclude that this defines a unique probability measure P_x on the σ-field generated by the intervals, i.e., on the Borel field $\mathscr{B}_\infty$ or $\mathscr{B}_T$, respectively. The finite-dimensional distributions are the distributions in $\mathbf{F}$. The proof of the countable additivity is a significant part of Kolmogorov's existence theorem for stochastic processes; see Theorem A.3.

By this, we have defined events in $\mathbb{R}^\infty$ and $\mathbb{R}^T$ and know how to define probability measures in $(\mathbb{R}^\infty, \mathscr{B}_\infty)$ and $(\mathbb{R}^T, \mathscr{B}_T)$.

Remark 1.1. *It is worth noting that all the well-known processes in elementary probability theory can be introduced by finite dimensional distributions. Going back to Example 1.2, we saw there an explicit way to construct a Poisson process. On the other hand, following the route via the finite-dimensional distribution functions the Poisson process on the real line is characterized by the fact that the increments over a finite number of disjoint intervals are independent Poisson variables with expectation proportional to the interval length. This property can easily be formulated as a property of the finite-dimensional distributions at the interval end points, even if one never does so explicitly.*

Both approaches lead to a probability measure P_x, *defined for all* $B \in \mathscr{B}_T$, *either by* (1.4) *or by the Kolmogorov existence theorem. There is an important difference between the two approaches. With the explicit construction* all *sample functions are piecewise constant with unit jumps, by construction, and the probability measure* P_x *has the potential to take that into account and it can easily be extended to more events (subsets) than just those in* $\mathscr{B}_T$. *The construction by the Kolmogorov theorem does not allow any strong statement about sample function properties. In the next section and in Chapter 2 we discuss some consequences of this fact in more detail.*

Finally, also the explicit construction relies on the Kolmogorov existence theorem, since it requires the existence of an infinite sequence of independent exponential random variables.

1.2.4 An important comment on probabilities on $\mathbb{R}^T$

Processes with continuous time need some extra comments. Even if the existence of such processes can be deduced from Kolmogorov's theorem, the practical applications need some care. The sample space Ω is an abstract space and a mathematical construction, and the link to reality is provided by the random variables. In an experiment, one can observe the values of one or more random variables, x_1, x_2, etc. and also find their distribution, by some statistical procedure. There is no serious difficulty to allow the outcome to be any real number, and to define probability distributions on $\mathbb{R}$ or $\mathbb{R}^n$.

When the result of an experiment is a whole function with continuous parameter, the situation is more complicated. In principle, *all functions* of $t \in T$ are potential outcomes, and the sample space of all functions on T is simply so big that one needs to be restrictive on which sets of outcomes that should be given probability. There are too many realizations that ask for probability.

Here, practice comes to our assistance. In an experiment one can only observe the values of $x(t)$ at a finite number of times, $t_1, t_2, \ldots, t_n$; with $n = \infty$ we allow an unlimited series of observations. The construction of processes with continuous time was built on exactly this fact: the observable events are those that can be defined by countably many $x(t_j)$, $j \in \mathbb{N}$, and the probability measure we constructed assigns probabilities to only such events.

A set of functions that is defined by restrictions on countably many $x(t_j)$ is said to have a *countable basis*; see more on this in Appendix A.1. There are many important function sets that do not have a countable basis and are not Borel sets in $\mathbb{R}^T$. These sets are not given a unique probability by the finite-dimensional distribution functions.

Example 1.3. *Here are some examples of function sets with and without countable basis, when $T = [0,1]$:*

- $\{x \in \mathbb{R}^T; \lim_{n\to\infty} x(1/n) \text{ exists}\} \in \mathscr{B}_T$,
- $\{x \in \mathbb{R}^T; \lim_{t\to 0} x(t) \text{ exists}\} \notin \mathscr{B}_T$,
- $\{x \in \mathbb{R}^T; x \text{ is a continuous function}\} \notin \mathscr{B}_T$,
- $\{x \in \mathbb{R}^T; x(t) \le 2 \text{ for all rational } t\} \in \mathscr{B}_T$.

Approximation by finite-dimensional events

The events in the σ-field $\mathscr{B}_\infty$ in $\mathbb{R}^\infty$ can be approximated in probability by finite-dimensional sets. If $(\mathbb{R}^\infty, \mathscr{B}_\infty, P)$ is a probability space, and $B \in \mathscr{B}_\infty$, then, for every $\varepsilon > 0$, there is a finite n and an event $B_n \in \mathscr{B}_n$ such that

$$P(B \Delta \widetilde{B}_n) \leq \varepsilon,$$

where $\widetilde{B}_n = \{x \in \mathbb{R}^\infty; (x_1, \ldots, x_n) \in B_n\}$ and $A \Delta B = (A - B) \cup (B - A)$.

Similarly, events in $\mathscr{B}_T$ in $\mathbb{R}^T$ can be approximated arbitrarily close by events defined by the values of $x(t)$ for a finite number of t-values: $P(B \Delta \widetilde{B}_n) \leq \varepsilon$, with

$$\widetilde{B}_n = \{x \in \mathbb{R}^T; (x(t_1), \ldots, x(t_n)) \in B_n\}.$$

Remember that every probability measure on $(\mathbb{R}^\infty, \mathscr{B}_\infty)$ is uniquely determined by its finite-dimensional distributions, which implies that also every probability measure P on $(\mathbb{R}^T, \mathscr{B}_T)$ is determined by the finite-dimensional distributions, $\{F_{\mathbf{t}^n}\}_{n=1}^\infty$. In particular, the probability

$$P(\lim_{n \to \infty} x(t_0 + 1/n) \text{ exists and is equal to } x(t_0))$$

is determined by the finite-dimensional distributions. But $x(t_0 + 1/n) \to x(t_0)$ as $n \to \infty$ is almost, but not quite, the same as $x(t) \to x(t_0)$ as $t \to t_0$. To deal with sample function continuity we need a more refined construction of the probability measure from the finite-dimensional distributions. We deal with these matters in Chapter 2.

1.3 Stationary processes and fields

This section summarizes some elementary notation for stationary processes. More details, properties, and proofs of the most important facts will be given in Chapter 3. To avoid too cumbersome notation we sometimes allow ourselves to write "the process $x(t)$," when we should have used the notation $\{x(t), t \in \mathbb{R}\}$ for the process (or any other relevant parameter set T instead of $\mathbb{R}$). The simple notation x will also be frequently used, when the parameter space is clear from the context. If we mean "the random variable $x(t)$," we will say so explicitly. At this stage, we deal with real valued processes, only. Later, we will need to consider also complex valued processes.

We write $m(t) = E(x(t))$ for the *mean value function* and $r(s,t) = C(x(s),x(t))$ for the *covariance function* of the process $\{x(t), t \in \mathbb{R}\}$. If $\{x(t), t \in \mathbb{R}\}$ and $\{y(t), t \in \mathbb{R}\}$ are two processes, $r_{x,y}(s,t) = C(x(s), y(t))$ is the *cross-covariance function*.

1.3.1 Stationary processes, covariance, and spectrum

Definition 1.4 (Stationarity). *A stochastic process* $\{x(t), t \in \mathbb{R}\}$ *is strictly stationary if all n-dimensional distributions of* $x(t_1 + \tau), \ldots, x(t_n + \tau)$ *are independent of* τ. *It is called weakly stationary (or second order stationary) if its mean is constant,* $E(x(t)) = m$, *and if its covariance function,*

$$r(t) = C(x(s+t), x(s)),$$

is a function only of the time lag t.

Here are some simple properties of a covariance function of a stationary process.

Theorem 1.2. *If* $r(t)$ *is the covariance function of a (real) stationary process, then,*

a) $V(x(t+h) \pm x(t)) = 2(r(0) \pm r(h))$, *so* $|r(h)| \leq r(0)$,

b) if $|r(\tau)| = r(0)$ *for some* $\tau \neq 0$, *then* $r(t)$ *is periodic,*

c) if $r(t)$ *is continuous at* $t = 0$, *then it is continuous everywhere.*

Proof. Property (a) is clear, and it implies $|r(t)| \leq r(0)$. From (a) it also follows that if $r(\tau) = \pm r(0)$, then $x(t+\tau) = \pm x(t)$ for all t, so $x(t)$ is periodic with period τ and 2τ, respectively, which is (b). For (c), use the inequality, $C(x,y)^2 \leq V(x)V(y)$, on

$$\begin{aligned}(r(t+h) - r(t))^2 &= (C(x(0), x(t+h)) - C(x(0), x(t)))^2 \\ &= (C(x(0), x(t+h) - x(0)))^2 \\ &\leq V(x(0))\, V(x(t+h) - x(y)) = 2r(0) \cdot (r(0) - r(h)),\end{aligned}$$

which goes to 0 with h if $r(t)$ is continuous for $t = 0$. □

The covariance function is said to define the *time domain* properties of the process. Equally important for the theory is the *frequency domain* properties, expressed by the *spectral distribution*, studied in detail in Chapter 3.

Every continuous covariance function is a Fourier integral,

$$r(t) = \int_{-\infty}^{\infty} e^{i\omega t}\, \mathrm{d}F(\omega), \tag{1.5}$$

where the function $F(\omega)$ is the spectral distribution function.

The spectral distribution is characterized by the properties: [3]

- symmetry: $\mathrm{d}F(-\omega) = \mathrm{d}F(\omega)$,
- monotonicity: $\omega \leq \omega'$ implies $F(\omega) \leq F(\omega')$,
- boundedness: $F(+\infty) - F(-\infty) = r(0) = V(x(t)) < \infty$.

As indicated by the way we write the three properties, $F(\omega)$ is defined only up to an additive constant, and we usually take $F(-\infty) = 0$. The spectral distribution function is then equal to a cumulative distribution function multiplied by a positive constant, equal to the variance of the process. As we shall see in later chapters, ω is in a natural way interpreted as an *angular frequency*, not to be confused with the elementary event ω in basic probability theory.

In this work we will mainly use the symmetric form (1.5) of the spectral representation when we deal with theoretical aspects of the covariance function and of the spectrum. However, in many examples we will use a one-sided real form with $\omega \geq 0$; for details on this, see Section 3.2.3. In Section 3.3.5 we will meet examples where negative frequencies play an important role in their own right.

If $F(\omega)$ is absolutely continuous with $F(\omega) = \int_{s=-\infty}^{\omega} f(s)\,\mathrm{d}s$, then the spectrum is said to be (absolutely) continuous, and

$$f(\omega) = \frac{d}{d\omega} F(\omega)$$

is the *spectral density function*.

The *spectral moments* are defined as

$$\omega_k = \int_{-\infty}^{\infty} |\omega|^k \,\mathrm{d}F(\omega).$$

Note that the odd spectral moments are defined as absolute moments. Since F is symmetric around 0 the signed odd moments are always 0, when they exist.

[3] The differential notation in $\mathrm{d}F(-\omega) = \mathrm{d}F(\omega)$ will be used throughout the book as a convenient way to simplify increment notation.

Spectral moments may be finite or infinite. As we shall see in the next chapter, the finiteness of the spectral moments are coupled to the smoothness properties of the process $x(t)$. For example, the process is differentiable (in quadratic mean), see Section 2.1, if $\omega_2 = -r''(0) < \infty$, and the spectral density of the derivative is

$$f_{x'}(\omega) = \omega^2 f_x(\omega).$$

To ensure that sample functions are differentiable functions one needs just a little bit more than $\omega_2 = \int_{-\infty}^{\infty} \omega^2 f_x(\omega)\,\mathrm{d}\omega < \infty$, namely $-r''(t) = \omega_2 - C|t|^\alpha + o(|t|^\alpha)$ as $t \to 0$, for some $\alpha \in (1,2]$; see Theorem 2.10, Section 2.3.1.

The *mean frequency* of a stationary process,

$$\mu^+ = \frac{1}{2\pi}\sqrt{\omega_2/\omega_0}, \tag{1.6}$$

and its inverse, $1/\mu^+$, the *mean period*, characterize the main feature of the spectrum. The width of the spectrum around the mean frequency is measured by the *spectral width* parameter $1/\alpha = (\omega_0\omega_4)^{1/2}/\omega_2$.

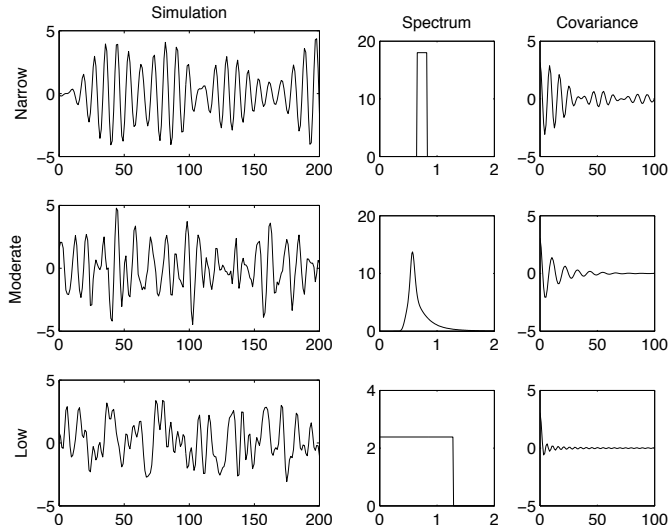


Figure 1.2 *Top: Processes with very narrow band spectrum, oscillating covariance function, and "fading" sample function; Middle: Realistic ocean waves with moderate band width wave spectrum; Bottom: low frequency white noise spectrum. The spectra are illustrated by their one-sided spectral densities.*

Example 1.4. *Here is a first example on the visual characteristics of spectrum, covariance function, and simulated sample function. Figure 1.2 illustrates one very narrow spectrum with oscillating covariance function and*

regular "fading" sample path, one realistic water wave spectrum with less regular sample paths, and one irregular process with "low frequency white noise" spectrum.

1.3.2 Stationary streams of events

The Poisson process is a *counting process* that counts the number of events that happen at discrete time points. Its characteristic property is that the number of events in any interval of length h has a Poisson distribution with mean λh, with parameter λ called the *intensity*. Further, the number of events in disjoint intervals are statistically independent. The events counted by the Poisson process is an example of a *stationary point process*. This is a natural generalization of the notion "stationary process," since the distribution of the number of events in a set of intervals is the same, regardless of their absolute location.

If one relaxes the independence condition, one obtains a more general class of point processes, where there may be causal or statistical dependence between events. For example, the intensity in a Poisson process may be allowed to vary like a stationary process. It is then possible to describe this dependence by means of a type of covariance function that describes the conditional intensity of events a distance t apart. We shall deal with this covariance density in Section 3.5, where we also introduce the corresponding spectral distribution. For general accounts of points processes, see [37, 39].

1.3.3 Random fields

A random field is a stochastic process $\{x(\mathbf{t})\}$ with multi-dimensional parameter $\mathbf{t} = (t_1, \ldots, t_p) \in \mathbf{T}$, which can be discrete or continuous. For example, if $\mathbf{t} = (t_1, t_2)$ is two-dimensional we can think of $(t_1, t_2, x(\mathbf{t}))$, $(t_1, t_2) \in \mathbb{R}^2$, as a random surface. The mean value and covariance functions are defined in the natural way, $m(\mathbf{t}) = E(x(\mathbf{t}))$ and $r(\mathbf{t}, \mathbf{u}) = C(x(\mathbf{t}), x(\mathbf{u}))$.

A random field is called *homogeneous* if it has constant mean value $m(\mathbf{t}) = m$ and the covariance function $r(\mathbf{t}, \mathbf{u})$ depends only on the vector $\mathbf{t} - \mathbf{u}$ between the two observation points, i.e., assuming $m = 0$,

$$r(\mathbf{t}) = r(\mathbf{u} + \mathbf{t}, \mathbf{u}) = E(x(\mathbf{u} + \mathbf{t}) \cdot x(\mathbf{u})).$$

The covariance of the process values at two parameter points depends on *distance* as well as on *direction* of the vector between the two points. This is the equivalent of a stationary process on the real line.

If the covariance between $x(\mathbf{u})$ and $x(\mathbf{v})$ depends only on the distance $\tau = \|\mathbf{u} - \mathbf{v}\|$ between the observation points and not on the direction, the field is

called *isotropic*. This requirement poses strong restrictions on the covariance function, as we shall see in Chapter 7, where random fields are treated in more detail.

A digital image can be thought of as a random field with discrete parameter $\mathbf{t} \in \mathbf{T} \subset \mathbb{N}^2$. A time-dependent random surface is a field $(s_1, s_2, x(t, s_1, s_2))$ with $\mathbf{t} = (t, s_1, s_2)$, where t represents time and $(s_1, s_2) \in \mathbb{R}^2$ is location.

1.4 Gaussian processes

1.4.1 Normal distributions and Gaussian processes

The Gaussian, or normal, univariate distribution with mean m and variance σ^2 has probability density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/(2\sigma^2)} = \frac{1}{\sigma}\,\phi\left(\frac{x-m}{\sigma}\right),$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ is the standardized normal density function. A distribution concentrated at one point m is regarded as Gaussian with $\sigma^2 = 0$.

Definition 1.5. *A vector $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_p)$ of p random variables is said to have a p-variate Gaussian (normal) distribution if every linear combination of its components $\mathbf{a}\boldsymbol{\xi}' = \sum_k a_k \xi_k$ has a normal distribution. The variables $\xi_1, \ldots, \xi_p$ are said to be "jointly normal."*

With mean vector $\mathbf{m} = E(\boldsymbol{\xi})$ and covariance matrix

$$\boldsymbol{\Sigma} = C(\boldsymbol{\xi}; \boldsymbol{\xi}) = E((\boldsymbol{\xi} - \mathbf{m})'(\boldsymbol{\xi} - \mathbf{m})),$$

the variance of $\mathbf{a}\boldsymbol{\xi}'$ is

$$V(\mathbf{a}\boldsymbol{\xi}') = \mathbf{a}\boldsymbol{\Sigma}\mathbf{a}' \geq 0.$$

The *characteristic function* of a normal variable ξ is (see Appendix A.5)

$$\varphi(s) = E(e^{is\xi}) = e^{ims - \sigma^2 s^2/2},$$

and that of a multi-variate normal $\boldsymbol{\xi}$ is

$$\varphi(\mathbf{s}) = E(e^{i\mathbf{s}\xi'}) = e^{i\mathbf{m}\mathbf{s}' - \mathbf{s}\boldsymbol{\Sigma}\mathbf{s}'/2}. \tag{1.7}$$

It follows from the form of the characteristic function that if two jointly normal variables are uncorrelated, i.e., have zero covariance, then they are also independent. To see this, one can first note that if $\mathbf{\Sigma}$ is diagonal, then the characteristic function factorizes into a product of characteristic functions for independent normals. The uniqueness of characteristic functions does the rest.

If the determinant of $\mathbf{\Sigma}$ is positive, the distribution of $\boldsymbol{\xi}$ is non-singular and has a density

$$f_{\boldsymbol{\xi}}(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}\sqrt{\det \mathbf{\Sigma}}}\, e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})\mathbf{\Sigma}^{-1}(\mathbf{x}-\mathbf{m})'}.$$

If the determinant is zero, the distribution of $\boldsymbol{\xi}$ is concentrated to a linear subspace of $\mathbb{R}^n$ and there exists at least one linear relationship between the components, i.e., there is at least one $\mathbf{a}$ for which $\mathbf{a}\boldsymbol{\xi}'$ is a constant.

Definition 1.6. *A stochastic process $\{x(t), t \in \mathbb{R}\}$ is a Gaussian process if every linear combination $S = \sum_k a_k x(t_k)$ for real a_k and $t_k \in \mathbb{R}$ has a Gaussian distribution.*

Example 1.5 (Existence of a Gaussian process)**.** *What knowledge is needed before we can ascertain the existence of a specific Gaussian process? In view of Kolmogorov's existence Theorem 1.1, we must present a consistent family of finite-dimensional distribution functions, such that every linear combination $\sum_k a_k x(t_k)$ of variables is Gaussian, i.e., that for every selection, the variables $x(t_1), \ldots, x(t_n)$ are jointly normal.*

Since the multi-variate normal distribution is defined by its characteristic function, it is clear from (1.7) *that it suffices to specify a function $m(t)$ and a function $r(s,t)$ such that*

$$\sum_{j,k} a_j a_k r(t_j - t_k) \geq 0, \tag{1.8}$$

for all real a_k and t_k. Then there exists a Gaussian process with $m(t)$ and $r(s,t)$ as mean value and covariance function, respectively. We will meet condition (1.8) *in Chapter 8.*

It is an easy consequence of the definition that the derivative of a Gaussian

process is also Gaussian (when it exists), since it is the limit of the Gaussian variable $z_h = (x(t+h) - x(t))/h$ as $h \to 0$. For a stationary Gaussian process $\{x(t), t \in \mathbb{R}\}$ the mean of z_h is 0 and it has variance $V(z_h) = 2(r(0) - r(h))/h^2$. As we shall prove in Section 2.1.2 this converges to $\omega_2 = \int \omega^2 \, \mathrm{d}F(\omega \leq \infty$. The derivative exists only if this limit is finite.

Also the integral of a Gaussian process is a Gaussian variable; conditions for the existence will be given in Section 2.4.

1.4.2 Conditional normal distributions

The multivariate normal distribution has the very useful property that, conditional on observations of a subset of variables, the unobserved variables are also normal.[4] Further, the conditional expectation is a linear combination of the observations while variances and covariances are independent of the observed values.

Let $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ and $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_m)$ be two jointly Gaussian vectors with mean values

$$E(\boldsymbol{\xi}) = m_{\boldsymbol{\xi}}, \quad E(\boldsymbol{\eta}) = m_{\boldsymbol{\eta}},$$

and with covariance matrix (with $\boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\eta}} = \boldsymbol{\Sigma}'_{\boldsymbol{\eta}\boldsymbol{\xi}}$),

$$\boldsymbol{\Sigma} = C((\boldsymbol{\xi}, \boldsymbol{\eta}); (\boldsymbol{\xi}, \boldsymbol{\eta})) = \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\xi}} & \boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\eta}} \\ \boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\xi}} & \boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}} \end{pmatrix}.$$

If the determinant of the covariance matrix $\boldsymbol{\Sigma}$ is positive, then the distribution of $(\boldsymbol{\xi}, \boldsymbol{\eta})$ has a non-singular density

$$f_{\boldsymbol{\xi}\boldsymbol{\eta}}(\mathbf{x}, \mathbf{y}) = \frac{1}{(2\pi)^{(m+n)/2}\sqrt{\det \boldsymbol{\Sigma}}} e^{-\frac{1}{2}(\mathbf{x} - m_{\boldsymbol{\xi}}, \mathbf{y} - m_{\boldsymbol{\eta}})\boldsymbol{\Sigma}^{-1}(\mathbf{x} - m_{\boldsymbol{\xi}}, \mathbf{y} - m_{\boldsymbol{\eta}})'}.$$

The density of $\boldsymbol{\eta}$ is

$$f_{\boldsymbol{\eta}}(\mathbf{y}) = \frac{1}{(2\pi)^{m/2}\sqrt{\det \boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}}}} e^{-\frac{1}{2}(\mathbf{y} - m_{\boldsymbol{\eta}})\boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}(\mathbf{y} - m_{\boldsymbol{\eta}})'},$$

and the conditional density $f_{\boldsymbol{\xi}|\boldsymbol{\eta}}(\mathbf{x}|\mathbf{y})$, defined as

$$f_{\boldsymbol{\xi}|\boldsymbol{\eta}}(\mathbf{x} \mid \mathbf{y}) = \frac{f_{\boldsymbol{\eta}\boldsymbol{\xi}}(\mathbf{y}, \mathbf{x})}{f_{\boldsymbol{\eta}}(\mathbf{y})}.$$

[4] A summary of conditional distributions is found in Section A.3.3 in Appendix A.

This is again a Gaussian density and its mean and covariance properties can easily be found by some matrix algebra. We summarize the important facts as a theorem.

Theorem 1.3. *The conditional distribution of* $\boldsymbol{\xi}$ *given* $\boldsymbol{\eta} = \mathbf{y}$ *is Gaussian with conditional mean matrix*

$$E(\boldsymbol{\xi} \mid \boldsymbol{\eta} = \mathbf{y}) = \widehat{\boldsymbol{\xi}}(\mathbf{y}) = E(\boldsymbol{\xi}) + C(\boldsymbol{\xi};\boldsymbol{\eta})\boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}(\mathbf{y} - E(\boldsymbol{\eta}))' = m_{\boldsymbol{\xi}} + \boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\eta}}\boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}(\mathbf{y} - m_{\boldsymbol{\eta}})'. \tag{1.9}$$

The conditional covariance is

$$\boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\xi}|\boldsymbol{\eta}} = E((\boldsymbol{\xi} - \widehat{\boldsymbol{\xi}}(\boldsymbol{\eta}))'(\boldsymbol{\xi} - \widehat{\boldsymbol{\xi}}(\boldsymbol{\eta}))) = \boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\xi}} - \boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\eta}}\boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\xi}}. \tag{1.10}$$

In two dimensions the formulas read

$$m_{x|y} = m_x + \sigma_x \rho_{xy} \cdot \frac{y - m_y}{\sigma_y}, \qquad \sigma_{x|y}^2 = \sigma_x^2(1 - \rho_{xy}^2),$$

with $\rho_{xy} = C(x,y)/\sqrt{V(x)V(y)}$; thus the squared correlation ρ_{xy}^2 gives the relative reduction of the variability (uncertainty) in the random variable x gained by observation of y.

Observe the mnemotechnical friendliness of the formulas for conditional mean and covariance. For example, the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\xi}|\boldsymbol{\eta}}$ has dimension $n \times n$ and the configuration on the right hand side of (1.10) is the only way to combine the matrices involved that matches their dimensions – of course, you have to remember the general structure.

1.4.3 Linear prediction and reconstruction

Prediction and reconstruction are two of the most important applications of stationary process theory. Even though these problems are not main topics in this work, we present one of the basic concepts here; in Section 6.6 we will deal with the more philosophical sides of the prediction problem.

Suppose we have observed the outcomes of a set of random variables, $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_m)$, and that we want to give a statement $\widehat{\xi}$ about the outcome of some other variable ξ, either to be observed sometime in the future, or perhaps a missing observation in a time series. These two cases constitute the framework of *prediction* and *reconstruction*, respectively. Also suppose that

we want to make the statement in the best possible way in the mean square sense, i.e., we want $E((\xi - \widehat{\xi})^2)$ to be as small as possible.

It is easy to see (Theorem A.4 in Appendix A) that the best solution in mean square sense is given by the conditional expectation, $\widehat{\xi} = E(\xi \mid \boldsymbol{\eta}) = \varphi(\boldsymbol{\eta})$. On the other hand, if the variables are jointly Gaussian, then we know from formula (1.9) that the conditional expectation of ξ given $\boldsymbol{\eta}$ is linear in $\boldsymbol{\eta}$, so that for Gaussian variables the optimal solution is

$$\widehat{\xi} = E(\xi \mid \boldsymbol{\eta}) = m_\xi + \boldsymbol{\Sigma}_{\xi\boldsymbol{\eta}} \boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1} (\boldsymbol{\eta} - m_{\boldsymbol{\eta}})', \tag{1.11}$$

an expression that depends only on the mean values and the second order moments, i.e., variances and covariances.

Looking at the general case, without assuming normality, we restrict ourselves to solutions that are linear functions of the observed variables. It is clear that the solution that is optimal in the mean square sense can only depend on the mean values and variances/covariances of the variables. It therefore has the same form for all variables with the same first and second order moments. Thus, (1.11) gives the best linear predictor in mean square sense, regardless of the distribution.

1.5 Four historical landmarks

1.5.1 Brownian motion and the Wiener process

There is no other Gaussian process with as wide applicability as the Wiener process. Even if it is a non-stationary process it appears repeatedly in the theory of stationary processes, and we spend this section to describe some of its properties and applications.

Definition 1.7 (Wiener process)**.** *The Wiener process $\{w(t); t \geq 0\}$ is a Gaussian process with $w(0) = 0$ such that $E(w(t)) = 0$, and the variance of the increment $w(t+h) - w(t)$ over any interval $[t, t+h]$, $h > 0$, is proportional to the interval length,*

$$V(w(t+h) - w(t)) = h\,\sigma^2.$$

A Wiener process $\{w(t), t \in \mathbb{R}\}$ over the whole real line is a combination of two independent Wiener processes w_1 and w_2, so that $w(t) = w_1(t)$ for $t \geq 0$, and $w(t) = w_2(-t)$ for $t < 0$.

It is an easy consequence of the definition that the increment $w(t) - w(s)$ is uncorrelated with $w(s)$ for $s < t$; just expand the variance of $x(t)$,

$$\begin{aligned} t\sigma^2 &= V(w(t)) = V(w(s)) + V(w(t) - w(s)) + 2C(w(s), w(t) - w(s)) \\ &= s\sigma^2 + (t-s)\sigma^2 + 2C(w(s), w(t) - w(s)). \end{aligned}$$

It also follows that $C(w(s), w(t)) = \sigma^2 \min(s,t)$. Since the increments over disjoint intervals are jointly normal, by the definition of a normal process, they are also independent.

A characteristic feature of the Wiener process is that its future changes are statistically independent of its actual and previous values. It is intuitively clear that a process with this property cannot be differentiable. The increment over a small time interval from t to $t+h$ is of the order $\sqrt{h}$, which is small enough to make the process continuous, but it is too large to give a differentiable process. In Chapter 2 we will give conditions for continuity and differentiability of sample functions of Gaussian and general processes.

The sample functions of a Wiener process are in fact objects that have *fractal dimension*, and the process is *self similar* in the sense that when magnified with proper scales it retains its statistical geometrical properties. More precisely, for each $a > 0$, the process $\sqrt{a}\, w(t/a)$ has the same distributions as the original process $w(t)$.

The Wiener process is commonly used to model phenomena where the local changes are virtually independent. Symbolically, one can write $\mathrm{d}w(t)$ for the infinitesimal independent increments, or simply use "the derivative" $w'(t)$. In this form, the Wiener process is used as a driving force in many stochastic systems, appearing in Chapter 4. The Brownian motion is a good example of how one can use the Wiener process to get models with more or less physical realism.

The Brownian motion, observed 1827 by the Scottish biologist Robert Brown, is an erratic movement by small particles immersed in a fluid, for example pollen particles in water as in Brown's original experiment. Albert Einstein presented in 1905 a quantitative model for the Brownian movement in his paper *On the movements of small particles in a stationary liquid demanded by the molecular-kinetic theory of heat*, reprinted in [46], based on the assumption that the movements are caused by independent impacts on the particle by the molecules of the surrounding fluid medium.

In Einstein's model the changes in location due to collisions over separate time intervals are supposed to be independent. This requires however that the particles have no mass, which is physically wrong, but sufficiently accurate for microscopic purposes on a human time scale. According to Einstein, the

change in location in any of the three directions (x,y,z) over a time interval of length t is random and normal with mean zero, which is not surprising, since the changes are the results of a very large number of independent collisions. What made Einstein's contribution conclusive was that he derived an expression for the variance in terms of other physical parameters, namely

$$V(x(t)) = V(y(t)) = V(z(t)) = t\,\frac{4RT}{Nf} = t\,\sigma^2, \tag{1.12}$$

where T is the absolute temperature, R is the Boltzmann constant, and N is Avogadro's number, i.e., the number of molecules per mole, and the friction coefficient f depends on the shape and size of the particle and on the viscosity of the fluid. Here, the coordinates are independent Wiener processes.

Observations of the Brownian movement and estimation of its variance makes it possible to calculate any of the factors in σ^2, for example N, from the other ones.

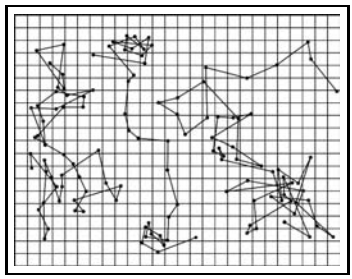


Figure 1.3 *Three particle paths in Perrin's experiment.*

The French physicist J.B. Perrin estimated in a series of experiments 1908-1911 Avogadro's number in this way by observing suspended rubber particles and found an estimate that is correct within about 10%. The figure, reproduced from the book *Les Atomes* by Perrin [96], shows the paths of the motions of three colloidal particles as seen under the microscope. Dots show the positions every 30 seconds. For his experiments and conclusions, Perrin was awarded the Nobel prize in physics, 1926, for his work on the "discontinuous structure of matter."

In a more realistic model, one takes also particle mass and velocity into account. If $v(t)$ denotes the velocity at time t, the fluid offers a resistance from the friction force, which is equal to $fv(t)$, with the friction coefficient as in (1.12). Further, the particle offers a resistance to changes in velocity proportional to its mass m. Finally, one needs to model the independent collisions from the surrounding fluid molecules, and here the Wiener process can again be used, more precisely its increments $\mathrm{d}w(t)$. This gives the *Langevin equation*, named after the French physicist Paul Langevin, for the particle velocity,

$$\mathrm{d}v(t) + \alpha\, v(t)\,\mathrm{d}t = \frac{1}{m}\,\mathrm{d}w(t), \tag{1.13}$$

where $\alpha = f/m$, and $w(t)$ is a standardized Wiener process. It is usually written

$$\frac{\mathrm{d}v(t)}{\mathrm{d}t} + \alpha\, v(t) = \frac{1}{m}\, w'(t). \tag{1.14}$$

We will meet the Langevin equation in Example 3.5 on the Ornstein-Uhlenbeck process in Section 3.4.2.

1.5.2 S.O. Rice and electronic noise

The two papers *Mathematical analysis of random noise*, by Steve O. Rice, appeared in Bell System Technical Journal, 1944–1945 [100]. They represent a landmark in the history of stochastic processes in that they bring together and exhibit the wide applicability in engineering of the spectral formulation of a stationary process as a sum, or asymptotically an integral, of harmonic cosine functions with random amplitudes and phases. Correlation functions and their Fourier transforms had been studied at least since the early 1900s, and Rice's work brought together these results in a systematic way. But it also contained many new results, in particular pertaining to crossing related properties, and on the statistical properties of stationary processes "obtained by passing random noise through physical devices."

Rice used the *spectral representation* of a stationary Gaussian process as a sum over discrete positive frequencies $\omega_k > 0$,

$$x(t) = \sum_k a_k \cos \omega_k t + b_k \sin \omega_k t = \sum_k c_k \cos(\omega_k t + \phi_k), \tag{1.15}$$

where the amplitudes a_k, b_k are normal and independent random variables with mean 0 and $E(a_k^2) = E(b_k^2) = \sigma_k^2$, and ϕ_k uniformly distributed over $(0, 2\pi)$, independent of the amplitudes. As we shall see in Chapter 3, such a process has covariance

$$r(t) = \sum_k \sigma_k^2 \cos \omega_k t.$$

The spectral distributions function is a discrete distribution with point mass $\sigma_k^2/2$ at the symmetrically located frequencies $\pm\omega_k$.

The continuous, integral form of the spectral representation is presented as limiting cases in Rice's work. At about the same time, Harald Cramér and Michel Loève gave a probabilistic formulation of the continuous spectral representation, in a mathematically impeccable way [31, 33]. In Chapter 3 we shall deal in more depth with spectral representation, in different forms.

Besides the previously known "Rice's formula" for the expected number of level crossings, Rice's 1945 paper also analyzed crossing and excursion distributions and investigated the joint occurrence of crossings of a fixed level at two distinct points, necessary for calculation of the variance of the number of crossings. The simple Rice's formula can be generalized to vector processes and to random fields, and has developed into one of the most powerful tools in the statistical analysis of random systems; see [6]. Crossing problems are the topic of Chapter 8.

The flexibility and generality of Rice's methods and examples made correlation and spectral theory fundamental ingredients in communication theory and signal processing for decades to come. An example by Rice himself is the ingenious explanation of the intriguing click noise in analogue FM radio [101].

1.5.3 Gaussian random wave models

Steve Rice's analysis of time dependent stationary processes had, as mentioned, great influence on signal processing in the information sciences. Less well known in the statistical world is the effect his work had in oceanography and naval architecture.

To give the reader a feeling for the dramatic effect Rice's work on random noise had in a quite different field, we cite in extenso (references deleted) the first two paragraphs in Manley St Denis and Willard J. Pierson's paper, *On the motion of ships in confused seas*, which came out 1954 [115]:

> HISTORY
>
> Three years ago the first co-author of the present work collaborated with Weinblum in the writing of a paper entitled "On the motion of ships at sea." In that paper Lord Rayleigh was quoted saying: "The basic law of the seaway is the apparent lack of any law." Having made this quotation, however, the authors then proceed to consider the seaway as being composed of "a regular train of waves defined by simple equations." This artificial substitution of pattern for chaos was dictated by the necessity of reducing the utterly confused reality to a simple form amenable to mathematical treatment.
>
> Yet at the same time and in other fields the challenging study of confusion was being actively pursued. Thus in 1945 Rice was writing on the mathematical analysis of random noise and in 1949 Tukey and Hamming were writing on the properties of stationary time series and their power spectra in connection with colored noise. In the same year Wiener published his now famous book on time series. These works

> were written as contributions to the theory of communication. Nevertheless the fundamental mathematical discipline expounded therein can readily be extended to other fields of scientific endeavor. Thus in 1952 the second co-author, inspired by a contribution of Tukey, was able to apply the foregoing theories to the study of actual ocean waves. As the result of analyses of actual wave records, he succeeded in giving not only a logical explanation as to why waves are irregular, but a statement as well of the laws underlying the behavior of a seaway. There is indeed a basic law of the seaway. Contrary to the obvious inference from the quotation of Lord Rayleigh, the seaway can be described mathematically and precisely, albeit in a statistical way.

If Rice's work had been in the vein of generally accepted ideas in communication theory, the St Denis and Pierson paper represented a complete revolution in common naval practice. Nevertheless, its treatment of irregular water waves as, what is now called, a random field was almost immediately accepted, and set a standard for much of naval architecture. A parallel theory was developed in England where M.S. Longuet-Higgins studied the geometric properties of random wave fields. His long paper [85] from 1957 is still much cited in oceanography, optics, and statistics.

One possible reason for the success of the St Denis-Pierson approach can be that the authors succeeded to formulate and analyze the motions of a ship that moved with constant speed through the random field. The random sea could directly be used as input to a linear (later also non-linear) filter representing the ship. Linear filters are the topic of Chapters 4 and 5.

St Denis and Pierson extended the one-dimensional description of a time dependent process $\{x(t), t \in \mathbb{R}\}$, useful for example to model the waves measured at a single point, to a random field $x(t,(s_1,s_2))$ with time parameter t and location parameter (s_1,s_2). They generalized the sum (1.15) to be a sum of directed waves, with $\boldsymbol{\omega} = (\omega, \kappa_1, \kappa_2)$,

$$\sum_{\boldsymbol{\omega}} A_{\boldsymbol{\omega}} \cos(\omega t - \kappa_1 s_1 - \kappa_2 s_2 + \phi_{\boldsymbol{\omega}}), \tag{1.16}$$

with random amplitudes and phases.

For fixed t each element in (1.16) is a cosine-function in the plane, which is zero along lines $\omega t - \kappa_1 s_1 - \kappa_2 s_2 + \phi_{\boldsymbol{\omega}} = \pi/2 + k\pi$, k integer. The parameters κ_1 and κ_2, called the *wave numbers*, determine the direction and wave length of the waves and ω is the wave *(angular) frequency*, when the sea is observed at a fixed location (s_1,s_2). More on this will follow in Chapter 7.

1.5.4 Detection theory and statistical inference

The first three landmarks illustrated the relation between stochastic model building and physical knowledge, in particular how the concepts of statistical independence and dependence between signal and functions relate to the physical world. About the same time as Rice and St Denis & Pierson advanced physically based stochastic modeling, the statistical inference methodology was placed firmly into a theoretical mathematical framework, as it was by then documented by the classical book by Harald Cramér, *Mathematical Methods of Statistics*, 1945 [32].

A few years later, the connection between the theoretical basis for statistical inference and important engineering questions related to signal detection was elegantly exploited by Ulf Grenander in his PhD thesis from Stockholm, *Stochastic processes and statistical inference* [56]. The classical problem in signal processing of deciding whether a deterministic signal of known shape $s(t)$ is present in an environment of Gaussian dependent, colored as opposed to white, random noise $x(t)$ can be treated as an infinite dimensional decision problem, testing an infinite dimensional statistical hypothesis; see, e.g., the standard work [117] on detection theory.

Suppose one observes a Gaussian stochastic process $x(t)$, $a \leq t \leq b$, with known correlation structure, but with unknown mean value function $m(t)$. If no signal is present, the mean value is 0, but with signal, the mean is equal to a function $s(t)$ of known shape. In statistical terms one has to test the following two hypotheses against each other:

$$H_0: \quad m(t) = 0,$$
$$H_1: m(t) = s(t).$$

Grenander introduced a series of independent Gaussian *observables*,

$$y_k = \int h_k(t)x(t)\,\mathrm{d}t,$$

by choosing the filter functions h_k as solutions to the integral equation

$$\int r(s,t)h_k(t)\,\mathrm{d}t = c_k h_k(s),$$

with $c_1 \geq c_2 \geq \ldots, c_k \to 0$ as $k \to \infty$. Under H_1 the observables will have mean $a_k = \int h_k(t)m(t)\,\mathrm{d}t$ and variance c_k, while under H_0 they will have mean 0, and the same variances. So instead of a continuous problem, we have gotten a denumerable problem, in which one can make a Likelihood-Ratio test of the two alternatives. We will return to this problem in Section 5.3.4.

Exercises

1:1. Consider the sample space $\Omega = [0,1]$ with uniform probability P, $P([a,b]) = b-a$, $0 \le a \le b \le 1$. Construct a stochastic process $\mathbf{y} = (x_1, x_2, \ldots)$ on Ω such that the components are independent zero-one variables, with $P(x_k = 0) = P(x_k = 1)$.

1:2. Show that $\mathscr{B}(\mathbb{R}) =$ the class of Borel sets in $\mathbb{R}$ is generated by

a) the open intervals,

b) the closed intervals.

1:3. Prove the claim in Lemma 1.1 that the inverse $x^{-1}(\mathscr{B}) \in \mathscr{F}$.

1:4. A set $A \subset \{1,2,\ldots\}$ is said to have asymptotic density θ if

$$\lim_{n\to\infty} n^{-1}|A \cap \{1,2,\ldots,n\}| = \theta.$$

(Note, $|B|$ denotes the number of elements in B.) Let $\mathscr{A}$ be the family of sets for which the asymptotic density exists. Is $\mathscr{A}$ a field?

1:5. Take $\mathbb{R}^n$ and show that the family $\mathscr{F}_0$ whose elements are unions of finitely many rectangles $(a_i, b_j]$ (with possibly infinite end points) is a field. Then, let T be an interval and convince yourself that the finite dimensional rectangles in $\mathbb{R}^T$ and unions of finitely many such rectangles, is a field.

1:6. Show that the co-ordinate process $x_k(\omega) = \omega_k$ is a stochastic process on $(\mathbb{R}^\infty, \mathscr{B}^\infty)$.

1:7. Let $\mathbb{Z}$ be the integers, and $\mathscr{A}$ the family of subsets A, such that either A or its complement A^c is finite. Let $P(A) = 0$ in the first case and $P(A) = 1$ in the second case. Show that P can not be extended to a probability to $\sigma(\mathscr{A})$, the smallest σ-field that contains $\mathscr{A}$.

1:8. Take $T = [0,1]$, and consider the set of functions which are continuous on the rational numbers, i.e.,

$$C_Q = \{x \in \mathbb{R}^T; x(q) \to x(q_0) \text{ for all rational numbers } q_0\},$$

where the limit is taken as q tends to q_0 through the rational numbers. Show that $C_Q \in \mathscr{B}_T$.

1:9. Prove that the increments of a Wiener process, as defined as in Definition 1.7, are independent and normal.

1:10. Prove the following useful inequality valid for any non-negative, integer-valued random variable N,

$$E(N) - \frac{1}{2}E(N(N-1)) \le P(N > 0) \le E(N).$$

Generalize it to the following inequalities, where

$$\alpha_i = E(N(N-1)\cdot\ldots\cdot(N-i+1))$$

is the i^{th} factorial moment,

$$\frac{1}{k!}\sum_{i=0}^{2n-1}(-1)^i\frac{1}{i!}\alpha_{(k+i)} \le P(N=k) \le \frac{1}{k!}\sum_{i=0}^{2n}(-1)^i\frac{1}{i!}\alpha_{(k+i)}.$$

Bibliography

[1] Åberg, S. (2007); Wave intensities and slopes in Lagrangian seas. *Adv. Appl. Probab.*, **39**, 1020–1035. Cited on p. 273.

[2] Adler, R. (1981): *The geometry of random fields*. John Wiley & Sons, Chichester. Cited on p. 217, 220.

[3] Adler, R.J. and Taylor, J.E. (2007): *Random fields and geometry*. Springer, New York. Cited on p. 60, 217.

[4] Andrews, D.W.K. (1984): Non-strong mixing autoregressive processes. *J. Appl. Probab.*, **21**, 930–934. Cited on p. 205.

[5] Apostol, T.M. (1974): *Mathematical analysis*. 2nd ed., Addison-Wesley, Reading. Cited on p. 299.

[6] Azaïs, J.-M. and Wschebor, M. (2009): *Level sets and extrema of random processes and fields*. John Wiley & Sons, Hoboken. Cited on p. 26, 217, 219, 236, 238, 244, 272, 272.

[7] Bartlett, M.S. (1963): The spectral analysis of point processes. *J. Roy. Stat. Soc. B (Methodological)*, **25**, 264–296. Cited on p. 105.

[8] Belyaev, Yu.K. (1959): Analytic random processes. *Theor. Prob. Appl.*, **4**, 402–409. Cited on p. 47, 51, 197.

[9] Belyaev, Yu.K. (1960): Properties of sample functions of a stationary Gaussian process. *Theor. Prob. Appl.*, **5**, 117–120. Cited on p. 47, 51.

[10] Belyaev, Yu.K. (1961): Continuity and Hölder's conditions for sample functions of stationary Gaussian processes. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, **2**, 22–33. Cited on p. 48.

[11] Belyaev, Yu.K. (1968): On the number of exits across the boundary of a region by a vector stochastic process. *Theor. Prob. Appl.*, **13**, 320–324. Cited on p. 268.

[12] Benedetto, J.J. (1992): Irregular sampling and frames. In *Wavelets: A tutorial in theory and applications*, Ed: C.K. Chui, pp. 445–507. Academic Press, San Diego. Cited on p. 163.

[13] Beran, J. (1994): *Statistics for long-memory processes*. Chapman &

Hall, New York. Cited on p. 142–144.

[14] Billingsley, P. (1999): *Convergence of probability measures*. 2nd ed., John Wiley & Sons, New York. Cited on p. 50, 54.

[15] Billingsley, P. (1995): *Probability and measure*. 3rd ed., John Wiley & Sons, New York. Cited on p. 275, 283.

[16] Bochner, S. (1933): Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse. *Math. Ann.*, **108**, 378–410. Cited on p. 74.

[17] Bolin, D. and Lindgren, F. (2011): Spatial models generated by nested partial differential equations, with an application to global ozone mapping. *Ann. Appl. Stat.*, **5**, 523–550. Cited on p. 228, 233.

[18] Bradley, R.C. (2005): Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys*, **2**, 107–144. Cited on p. 202.

[19] Breiman, L. (1968): *Probability*. Addison-Wesley, Reading. Reprinted 1992 in SIAM Classics in Applied Mathematics, SIAM, Philadelphia. Cited on p. xviii, 80, 106, 126, 252, 275, 296–300.

[20] Brémaud, P. and Massoulié, L. (2002): Power spectra of general shot noises and Hawkes point processes with a random excitation. *Adv. Appl. Probab.*, **34**, 205–222. Cited on p. 141.

[21] Brillinger, D. (1993): The digital rainbow: Some history and applications of numerical spectrum analysis. *Can. J. Stat.*, **21**, 1–19. Cited on p. 97.

[22] Brockwell, P.J. and Davis, R.A. (1991): *Time series: theory and tethods*, 2nd ed., Springer-Verlag, New York. Cited on p. 82, 101, 134, 152, 210, 215.

[23] Bulinskaya, E.V. (1961): On the mean number of crossings of a level by a stationary Gaussian process. *Theor. Prob. Appl.*, **6**, 435–438. Cited on p. 58, 236, 262, 273.

[24] Campbell, N. (1909): The study of discontinuous phenomena. *Proc. Cambr. Phil. Soc.*, **17**, 117–136. Cited on p. 141.

[25] Capiński, M. and Kopp, E. (2004): *Measure, Integral and Probability*. 2nd ed., Springer, London. Cited on p. 275.

[26] Chung, K.L. (2001): *A course in probability theory*. 3rd ed., Academic Press, San Diego. Cited on p. 77, 80, 291.

[27] Chung, K.L. and Williams, R.J. (1990): *Introduction to stochastic integration*. 2nd ed., Birkhäuser, Boston. Cited on p. 61.

[28] Cooley, J.W. (1992): How the FFT gained acceptance. *IEEE Signal Proc. Mag.*, **9**, 10–13. Cited on p. 309.

[29] Cooley, J.W. and Tukey, J.W. (1965): An algorithm for the machine calculation of complex Fourier series. *Math. Comp.*, **19**, 297–301. Cited on p. 309.

[30] Cramér, H. (1939): On the representation of a function by certain Fourier integrals. *Trans. Amer. Math. Soc.*, **46**, 191–201. Cited on p. 74.

[31] Cramér, H. (1942): On harmonic analysis in certain function spaces. *Arkiv Mat. Astron. Fysik*, **28B**, no. 12. Cited on p. 25, 85.

[32] Cramér, H. (1945): *Mathematical methods of statistics*. Princeton University Press, Princeton. Cited on p. 28.

[33] Cramér, H. (1946): On the theory of stochastic processes. *Proc. Tenth Scand. Congr. of Math.*, Copenhagen, pp. 28–39. Cited on p. 25.

[34] Cramér, H. (1962): Décompositions orthogonales de certains processus stochastiques. *Ann. Fac. Sci. Clermont*, **11**, 15–21. Cited on p. 196.

[35] Cramér, H. and Leadbetter, M.R. (1967): *Stationary and related stochastic processes*. John Wiley & Sons, New York. Reprinted by Dover Publications, 2004. Cited on p. xvii, 41, 47–54, 66, 88, 106, 159, 173, 236, 240, 244, 247, 253.

[36] Cressie, N. and Wikle, C.K. (2011): *Statistics for spatio-temporal data*. John Wiley & Sons, Hoboken. Cited on p. 217, 228.

[37] Daley, D.J. and Vere-Jones, D. (2003): *An introduction to the theory of point processes*, Vol I and II, 2nd ed., Springer, New York. Cited on p. 17, 105, 108, 111.

[38] Dedecker, J., Doukhan, P., Lang, G., et al. (2007): *Weak dependence: With examples and applications*, Springer Lecture Notes in Statistics, No. 190. Cited on p. 202.

[39] Diggle, P. (2003): *Statistical analysis of spatial point patterns*. 2nd ed., Edward Arnold, London. Cited on p. 17, 105.

[40] Dobrushin, R.L. (1960): Properties of sample functions of a stationary Gaussian process. *Theor. Probab. Appl.*, **5**, 120–122. Cited on p. 48.

[41] Doob, J.L. (1953): *Stochastic processes*. John Wiley & Sons, New York. Reprinted 1990 in Wiley Classics Library. Cited on p. 40, 198–200.

[42] Doukhan, P. and Louhichi, S. (1999): A new weak dependence condi-

tion and applications to moment inequalities. *Stoch. Proc. Appl.*, **84**, 313–342. Cited on p. 205.

[43] Doukhan, P. and Louhichi, S. (2001): Functional estimation of a density under a new weak dependence condition. *Scand. J. Stat.*, **28**, 325–341. Cited on p. 205.

[44] Durbin, J. (1985): The first-passage density of a continuous Gaussian process to a general boundary. *J. Appl. Probab.*, **22**, 99–122. Cited on p. 261.

[45] Durrett, R. (1991): *Probability: Theory and examples*. 10th ed., Cambridge University Press, Cambridge. Cited on p. 80, 173.

[46] Einstein, A. (1905): *Investigations on the theory of Brownian movement*. Reprinted 1956 by Dover Publications, New York. Cited on p. 23.

[47] Eyer, L. and Bartholdi, P. (1999): Variable stars: Which Nyquist frequency? *Astron. Astrophys. Sup.*, **135**, 1–3. Cited on p. 101.

[48] Feichtinger, H.G. and Gröchenig, K. (1994): Theory and practice of irregular sampling. In *Wavelets: Mathematics and applications*, pp. 305–363. Stud. Adv. Math., CRC, Boca Raton. Cited on p. 163.

[49] Fernique, X. (1964): Continuité des processus Gaussiens. *C.R. Acad. Sci. Paris*, **258**, 6058–6060. Cited on p. 60.

[50] Fernique, X. (1974): Régularité des trajectoires des fonctions aléatoires Gaussiennes. *Lecture Notes in Mathematics*, **480**, Springer-Verlag, New York. Cited on p. 220.

[51] Fisher, R.A. (1929): Tests of significance in harmonic analysis. *P. R. Soc. Lond. A-Conta.*, **125**, 54–59. Cited on p. 97.

[52] Garsia, A.M. (1965): A simple proof of Eberhard Hopf's maximal ergodic theorem. *J. Math. and Mech.*, **14**, 381–382. Cited on p. 184.

[53] Genz, A. (1992): Numerical computation of multivariate normal probabilities. *J. Comput. Graph. Stat.*, **1**, 141–149. Cited on p. 262.

[54] Genz, A. and Kwong, K.-S. (2000): Numerical evaluation of singular multivariate normal distributions. *J. Stat. Comput. Sim.*, **68**, 1–21. Cited on p. 262.

[55] Gradshteyn, I.S. and Ryzhik, I.M. (2007): *Table of integrals, series, and products*. 7th ed., Academic Press, New York. Cited on p. 146, 224, 226.

[56] Grenander, U. (1950): Stochastic processes and statistical inference.

Ark. Mat. **1**, 195-277. Cited on p. 28, 170, 193.

[57] Grimmett, G.R. and Strizaker, D.R. (2001): *Probability and random processes*. 3rd ed., Oxford University Press, Oxford. Cited on p. 275.

[58] Gut, A. (2009): *An intermediate course in probability*. 2nd ed., Springer, Dordrecht. Cited on p. 275.

[59] Hahn, S.L. (1996): *Hilbert transforms in signal processing*. Artech House, Norwood. Cited on p. 157.

[60] Hawkes, A.G. (1971): Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, **58**, 83–90. Cited on p. 116.

[61] Hurst, H.E. (1951): Long-term storage capacity of reservoirs. *Trans. ASCE*, **116**, 770–799. Cited on p. 144.

[62] Ibragimov, I.A. and Linnik, Yu.V. (1971): *Independent and stationary sequences of random variables*. Wolters-Noordhoff, Groningen. Cited on p. 202–205.

[63] Ibragimov, I.A. and Rozanov, Y.A. (1978): *Gaussian random processes*. Springer-Verlag, New York. Cited on p. 198, 204.

[64] Isserlis, L. (1918): On a formula for the product-moment correlation of any order of a normal frequency distribution in any number of variables. *Biometrika*, **12**, 134–139. Cited on p. 152.

[65] Jordan, D.W. and Smith, P. (2007): *Nonlinear ordinary differential equations*. 4th ed., Oxford University Press, Oxford. Cited on p. 133, 134, 138.

[66] Kac, M. and Slepian, D. (1959): Large excursions of Gaussian processes. *Ann. Math. Stat.*, **30**, 1215–1228. Cited on p. 253.

[67] Kadane, J.B. and O'Hagan, A. (1995): Using finitely additive probability: Uniform distributions on the natural numbers. *J. Am. Stat. Assoc.*, **90**, 626–631. Cited on p. 277.

[68] Karhunen, K. (1947): Über lineare Methoden in der Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae, Ser. A*, **37**, 1–79. Cited on p. 166.

[69] Karlin, S. and Taylor, H.M. (1981): *A second course in stochastic processes*. Academic Press, New York. Cited on p. 275.

[70] Kinsman, B. (1965): *Wind waves – their generation and propagation on the ocean surface*. Dover Editions 1984, 2002. Dover, Mineola. Cited on p. 232.

[71] Kolmogorov, A. (1933): *Grundbegriffe der Wahrscheinlichkeitsrech-*

nung. Springer-Verlag, Berlin. Cited on p. 6, 285.

[72] Lamperti, J. (1962): Semi-stable stochastic processes. *T. Am. Math. Soc.*, **104**, 62–78. Cited on p. 126.

[73] Lasota, A. and Mackey, M.C. (1994): *Chaos, fractals, and noise; stochastic aspects of dynamics*. 2nd ed., Springer-Verlag, New York. Cited on p. 173, 180.

[74] Leadbetter, M.R. (1974): On extreme values in stationary sequences. *Z. Wahrscheinlichkeit*, **28**, 289–303. Cited on p. 205, 246.

[75] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983): *Extremes and related properties of random sequences and processes*. Springer-Verlag, New York. Cited on p. 205, 235, 237, 244–248, 254.

[76] Lindgren, F., Rue, H. and Lindström, J. (2011): An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach (with discussion). *J. R. Statist. Soc. B*, **73**, 423–498. Cited on p. 217, 228.

[77] Lindgren, G. (1970): Some properties of a normal process near a local maximum. *Ann. Math. Stat.*, **41**, 1870–1883. Cited on p. 254, 260.

[78] Lindgren, G. (1972): Local maxima of Gaussian fields. *Ark. Mat.*, **10**, 195–218. Cited on p. 273.

[79] Lindgren, G. (1975): Prediction from a random time point. *Ann. Probab.*, **3**, 412–423. Cited on p. 253.

[80] Lindgren, G. (1980): Model processes in nonlinear prediction with application to detection and alarm. *Ann. Probab.*, **8**, 775–792. Cited on p. 268.

[81] Lindgren, G. (2012): A detailed statistical representation of optical vortices in random wave fields. *J. Opt.*, **14**, 035704. Cited on p. 273.

[82] Lindgren, G. and Rychlik, I. (1991): Slepian models and regression approximations in crossing and extreme value theory. *Int. Stat. Rev.*, **59**, 195–225. Cited on p. 254, 263.

[83] Loève, M. (1948): Fonctions aléatoire du second ordre, appendix to P. Lévy, *Processus stochastiques et mouvement Brownien*. Gautier-Villard, Paris. Cited on p. 85.

[84] Loève, M. (1962): *Probability theory*. Van Nostrand Reinhold, Princeton. Cited on p. 166.

[85] Longuet-Higgins, M.S. (1957): The statistical analysis of a random moving surface. *P. Tr. R. Soc. S-A*, **249**, 321–387. Cited on p. 27, 272.

[86] Mandelbrot, B. (1965): Une classe de processus stochastiques homothetiques a soi; application a loi climatologique de H.E. Hurst. *Comptes Rendus (Paris)*, **260**, 3274–3277. Cited on p. 144.

[87] Maruyama, G. (1949): The harmonic analysis of stationary stochastic processes. *Mem. Fac. Sci. Kyusyu Univ.* **A4**, 45–106. Cited on p. 193.

[88] Matérn, B. (1960): Spatial variation. *Meddelande från statens skogsforskningsinstitut*, **39(5)**. Cited on p. 226.

[89] Matheron, G. (1973): The intrinsic random functions and their applications. *Adv. Appl. Probab.*, **5**, 439–468. Cited on p. 143, 146, 217.

[90] Mercadier, C. (2005): MAGP toolbox. Available at http://math.univ-lyon1.fr/∼mercadier/MAGP/. Cited on p. 262.

[91] Micheson, A.A. and Stratton, S.W. (1898): A new harmonic analyser. *American Journal of Science*, **5**, 1–13. Cited on p. 97

[92] Miller, K.S. (1974): *Complex stochastic processes*. Addison-Wesley, Reading. Cited on p. 74.

[93] Mittal, Y. and Ylvisaker, D. (1975): Limit distributions for the maxima of stationary Gaussian processes. *Stoch. Proc. Appl.*, **3**, 1–18. Cited on p. 246.

[94] Øksendal, B. (1998): *Stochastic differential equations, an introduction with applications*. 5th ed., Springer, Berlin. Cited on p. 136.

[95] Percival, D.B. and Walden, A.T. (1993): *Spectral analysis for physical applications*. Cambridge University Press, Cambridge. Cited on p. 101.

[96] Perrin, J.B. (1913): *Les atomes*, Félix Alcan, Paris. Cited on p. 24

[97] Petersen, K. (1983): *Ergodic theory*. Cambridge University Press, Cambridge. Cited on p. 173.

[98] von Plato, J. (1994): *Creating modern probability*. Cambridge University Press, Cambridge. Cited on p. 180.

[99] Rényi, A. (1967): Remarks on the Poisson process. *Stud. Sci. Math. Hung.*, **5**, 119–123. Also in Lecture Notes in Mathematics 1967, **31**, 280–286. Cited on p. 108.

[100] Rice, S.O. (1944, 1945): Mathematical analysis of random noise. *Bell Syst. Tech. J.*, **23**, 282–332, and **24**, 46–156. Reprinted in: Wax, N. (1954): *Selected papers on noise and stochastic processes*. Dover Publ., New York. Cited on p. 25, 26, 69, 102, 141, 237.

[101] Rice, S.O. (1963): Noise in FM-receivers. In: *Time Series Analysis*, Ed:

M. Rosenblatt, Ch. 25, pp. 395–422. Wiley, New York. Cited on p. 26.

[102] Rosenblatt, M. (1956): A central limit theorem and a strong mixing condition. *P. Natl Acad. Sci. USA*, **42**, 43–47. Cited on p. 202.

[103] Rosiński, J. and Zak, T. (1997): The equivalence of ergodicity and weak mixing for infinitely divisible processes. *J. Theor. Probab.*, **10**, 73–86. Cited on p. 203.

[104] Royden, H.L. (1988): *Real Analysis*. 3rd ed., Prentice Hall, Englewood Cliffs. Cited on p. 198, 301.

[105] Rozanov, Yu.A. (1982): *Markov random fields*. Springer-Verlag, New York. Cited on p. 228.

[106] Rue, H. and Held, L. (2005): *Gaussian Markov random fields: Theory and applications*. Chapman & Hall, Boca Raton. Cited on p. 147, 217.

[107] Rychlik, I. (1987): A note on Durbin's formula for the first-passage density. *Stat. Probab. Lett.*, **5**, 425–428. Cited on p. 261.

[108] Rychlik, I. (2000): On some reliability applications of Rice's formula for the intensity of level crossings. *Extremes*, **3**, 331–348. Cited on p. 237, 239.

[109] Samorodnitsky, G. (2007): *Long range dependence*. now Publishers Inc, Delft. Cited on p. 142–144.

[110] Sato, K.-I. (1999): *Lévy processes and infinitely divisible distributions*. Cambridge University Press, Cambridge. Cited on p. 126.

[111] Schottky, W. (1912): Über spontane Stromschwankungen in verschiedenen Elektrizitätsleitern. *Annalen der Physik*, **362**, 541–567. Cited on p. 141.

[112] Schuster, A. (1898): On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. *Terrestrial Magnetism and Atmospheric Electricity*, **3**, 13–41. Cited on p. 97.

[113] Slepian, D. (1963): On the zeros of Gaussian noise. In: *Time Series Analysis*, Ed: M. Rosenblatt, pp. 104–115. Wiley, New York. Cited on p. 235, 254.

[114] Slutsky, E. (1937): Alcune propozitione sulla teoria della funzioni aleatorie. *Giorn. Ist. Ital. Attuari*, **8**, 193–199. Cited on p. 41, 60.

[115] St Denis, M. and Pierson, W.J. (1954): On the motion of ships in con-

fused seas. *Transactions, Soc. Naval Architects and Marine Engineers*, **61** (1953), pp. 280–357. Cited on p. 26, 272.

[116] Stein, M.L. (1999): *Interpolation of spatial data; Some theory for kriging*. Springer, New York. Cited on p. 145, 217, 227, 234.

[117] van Trees, H.L. (1968): *Detection, estimation, and modulation theory. Part I*. John Wiley & Sons, New York. Cited on p. 28, 170.

[118] The WAFO group (2011): WAFO – a MATLAB Toolbox for Analysis of Random Waves and Loads; Tutorial for WAFO version 2.5. March 2011.
Available at http://www.maths.lth.se/matstat/wafo/. Cited on p. 262, 264, 266.

[119] Whittle, P. (1954): On stationary processes in the plane. *Biometrika*, **41**, 434–449. Cited on p. 227.

[120] Whittle, P. (1963): Stochastic processes in several dimensions. *Bull. Intern. Statist. Inst.*, **40:2**, 974–994. Cited on p. 227.

[121] Williams, D. (1991): *Probability with martingales*. Cambridge University Press, Cambridge. Cited on p. 283, 291.

[122] Wold, H. (1954): *A study in the analysis of time series*. 2nd ed., Almquist and Wiksell, Stockholm. Cited on p. 196.

[123] Wong, E. and Hajek, B. (1985): *Stochastic processes in engineering systems*. Springer-Verlag, New York. Cited on p. 168, 172.

[124] Yaglom, A.M. (1962): *An introduction to the theory of stationary random functions*. Prentice Hall, Englewood Cliffs. Cited on p. xvii, 65, 88, 103, 198.

[125] Yaglom, A.M. (1987): *Correlation theory of stationary and related random functions, Vol I and II*. Springer-Verlag, New York. Cited on p. xviii, 217, 218.

This page intentionally left blank

This page intentionally left blank

Statistics

"… Lindgren's book offers both an introduction for graduate students and valuable insights for established researchers into the theory of stationary processes and its application in the engineering and physical sciences. His approach is rigorous but with more focus on the big picture than on detailed mathematical proofs. Strong points of the book are its coverage of ergodic theory, spectral representations for continuous- and discrete-time stationary processes, basic linear filtering, the Karhunen–Loève expansion, and zero crossings. …recommended both as a classroom text and for individual study."
—Don Percival, Senior Principal Mathematician, University of Washington, Seattle, USA

"…the book is authoritative and stimulating, a worthy champion of the tradition of Cramer and Leadbetter, admired by the author (and many others). It is a rich, inspiring book, full of good sense and clarity, an outstanding text in this important field."
—Clive Anderson, University of Sheffield, UK

Intended for a second course in stationary processes, **Stationary Stochastic Processes: Theory and Applications** presents the theory behind the field's widely scattered applications in engineering and science. In particular, the book:

- Reviews sample function properties and spectral representations for stationary processes and fields, including a portion on stationary point processes
- Presents and illustrates the fundamental correlation and spectral methods for stochastic processes and random fields
- Explains how basic theory is used in special applications like detection theory and signal processing, spatial statistics, and reliability
- Motivates mathematical theory from a statistical model-building viewpoint
- Introduces a selection of special topics, including extreme value theory, filter theory, long-range dependence, and point processes
- Provides more than 100 exercises with hints to solutions and selected full solutions

K15489

ISBN: 978-1-4665-5779-6
90000
9 781466 557796

CRC Press
Taylor & Francis Group
an **informa** business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
711 Third Avenue
New York, NY 10017
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK