

Edited by P. Bickel, P. Diggle, S. Fienberg, U. Gather,
I. Olkin, S. Zeger

Adrian Baddeley
Pablo Gregori
Jorge Mateu
Radu Stoica
Dietrich Stoyan
(Editors)

Case Studies in Spatial Point Process Modeling

With 107 Figures

Adrian Baddeley
Department of Mathematics
University of Western
Australia
Nedlands 6907 Australia
adrian@maths.uwa.edu.au

Pablo Gregori
Department of
Mathematics
Universitat Jaume I of
Castellon
Castellon 12071 Spain
gregori@mat.uji.es

Jorge Mateu
Department of
Mathematics
Universitat Jaume I of
Castellon
Castellon 12071 Spain
mateu@mat.uji.es

Radu Stoica
INRA - Biometrie,
Domaine St. Paul,
Site Agroparc
84914 Avignon, Cedex 9,
France
Radu.Stoica@avignon.inra.fr

Dietrich Stoyan
Institut für Stochastik
Prüferstraße 9
TU Bergakademie Freiberg
D-09596 Freiberg
Germany
stoyan@orion.hrz.tu-freiberg.de

Library of Congress Control Number: 2005931123

ISBN-10: 0-387-28311-0

ISBN-13: 978-0387-28311-1

Printed on acid-free paper.

© 2006 Springer Science+Business Media Inc.

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media Inc., 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use of general descriptive names, trade names, trademarks, etc., in this publication, even if the former are not especially identified, is not to be taken as a sign that such names, as understood by the Trade Marks and Merchandise Marks Act, may accordingly be used freely by anyone.

Camera ready copy provided by the editors.

Printed in the United States of America. (SBA)

9 8 7 6 5 4 3 2 1

springeronline.com

Preface

The week before Easter 2004 a conference on spatial point process modelling and its applications was held in Benicàssim (Castellón, Spain). The organizers targeted two aims. The first goal was to bring together most of the known people to guarantee the high scientific quality of the meeting to foster the theoretical and practical use of spatial point processes. The second one consisted of enabling young researchers to present their work and to obtain a valuable feed-back coming from the reknown specialists in the domain. The contributions of all the participants were published in the proceedings book of the conference.

The majority of the contributions in this book represents the reviewed version of the papers presented during the conference. In order to offer the reader a larger spectrum of this domain, authors that could not attend the conference were also invited to contribute.

The book is constituted by 16 chapters divided in three parts and gathering 44 authors coming from 13 different countries.

The first part of the volume – represented by its two first contributions – is dedicated to basic notions and tools for understanding and manipulating spatial point processes.

In the first contribution, D. Stoyan presents a general overview of the theoretical foundations for spatial point process. The author defines a point process and a marked point process, and describes the construction of the first and second order moment measures, which leads to the nowadays well known summary statistics such as the K -function, L -function or the pair-correlation function. The Poisson point process plays an important role, since in practice it is often used as null model for hypothesis testing and as reference model for the construction of realistic models for point patterns.

The second contribution, written by A.J. Baddeley and R. Turner, enters directly in the “flesh” of the problem presenting the concrete use of spatial point processes for modelling spatial point patterns, via the `spatstat` package – a software library for the R language. Four main points can be tackled by this package: basic manipulation of point patterns, exploratory data analysis,

parametric model-fitting and simulation of spatial point processes. The very important issue of model validation is also addressed. The contribution contains also the necessary mathematical details and/or literature references in order to avoid the use of this software as a “black box”. Two complete case studies are presented at the end of the contribution.

There is no serious practical application without a rigorous theoretical development. Therefore the second part of the book is more oriented towards theoretical and methodological advances in spatial point processes theory. Topics of this part of the book contain *analytical properties of the Poisson process* (presented in the contribution by S. Zuyev), *Bayesian analysis of Markov point processes* (by K. K. Berthelsen and J. Møller), *statistics for locally scaled point processes* (by M. Prokšová, U. Hahn and E. B. Vedel Jensen), *nonparametric testing of distribution functions in germ-grain models* (by Z. Pawlas and L. Heinrich), and *principal component analysis applied to point processes through a simulation study* (by J. Illian, E. Benson, J. Crawford and H. Staines). Remarkable is the fact, that almost all these contributions show direct applications of the presented development.

The third part of this volume is entirely dedicated to concrete, precise case studies, that are solved within the point processes theory. The presented applications are of big impact: *material science* (by F. Ballani), *human epidemiology* (by M. A. Martínez-Beneito *et al.*), *social sciences* (by N.A.C. Cressie, O. Perrin and C. Thomas-Agnan), *animal epidemiology* (by Webster *et al.* and P.J. Diggle, S. J. Eglen and J. B. Troy), *biology* (by F. Fleischer *et al.* and by A. Stein and N. Georgiadis), and *seismology* (by J. Zhuang, Y. Ogata and D. Vere-Jones and by A. Veen and F.P. Schoenberg). In their contributions, the authors show skill and cleverness in using, combining and continuously evolving the point processes tools in order to answer the proposed questions.

We hope the reader will enjoy reading the book and will find it instructive and inspiring for going a step further in this very open research field.

The Editors are grateful to all the authors that made possible finishing the book within an acceptable time scheduling. A word of thanks is given to Springer-Verlag and, in particular, to John Kimmel for creating the opportunity of making this project real.

Castellón (Spain)
May 2005

Adrian Baddeley
Pablo Gregori
Jorge Mateu
Radu Stoica
Dietrich Stoyan
Editors

Contents

Preface	V
List of Contributors	XV

Part I Basic Notions and Manipulation of Spatial Point Processes

Fundamentals of Point Process Statistics

<i>Dietrich Stoyan</i>	3
1 Basic Notions of the Theory of Spatial Point Processes	3
2 Marked Point Processes	5
3 The Second-order Moment Measure	6
4 Introduction to Statistics for Planar Point Processes	9
5 The Homogeneous Poisson Point Process	17
6 Other Point Processes Models	20
References	21

Modelling Spatial Point Patterns in R

<i>Adrian Baddeley, Rolf Turner</i>	23
1 Introduction	23
2 The spatstat Package	23
2.1 Scope	24
2.2 Data Types in spatstat	24
3 Data Analysis in spatstat	25
3.1 Data Input	25
3.2 Initial Inspection of Data	26
3.3 Exploratory Data Analysis	27
4 Point Process Models	27
4.1 Formulation of Models	28
4.2 Scope of Models	29
4.3 Model-fitting Algorithm	29

5	Model-fitting in spatstat	30
5.1	Overview	30
5.2	Spatial Trend Terms	31
5.3	Interaction Terms	32
6	Fitted Models	33
6.1	Interpretation of Fitted Coefficients	34
6.2	Invalid Models	35
6.3	Predicting and Plotting a Fitted Model	36
7	Models with Covariates	37
7.1	Covariates in a List of Images	38
7.2	Covariates in a Data Frame	40
8	Fitting Models to Multitype Point Patterns	41
8.1	Conditional Intensity	41
8.2	Multitype Models	42
9	Irregular Parameters	44
10	Model Validation	45
10.1	Residuals and Diagnostics	46
10.2	Formal Inference	46
11	Harkness-Isham Ants' Nests Data	50
11.1	Description of Data	50
11.2	Exploratory Analysis	51
11.3	Modelling	54
11.4	Formal Inference	57
11.5	Incorporation of Covariates	59
12	The Queensland Copper Data	62
12.1	Data and Previous Analyses	62
12.2	Analysis	63
12.3	North-South Effect	69
	References	70

Part II Theoretical and Methodological Advances in Spatial Point Processes

Strong Markov Property of Poisson Processes and Slivnyak Formula

<i>Sergei Zuyev</i>	77
1 Filtrations and Stopping Sets	77
2 Slivnyak Theorem for Locally Defined Processes	81
References	84

Bayesian Analysis of Markov Point Processes

<i>Kasper K. Berthelsen, Jesper Møller</i>	85
1 Introduction	85
2 Auxiliary Variable Method	88
3 The Strauss Process	89

3.1	Specification of Auxiliary Point Processes	89
3.2	Results for the Auxiliary Variable Method	91
4	Concluding Remarks	93
	References	96

Statistics for Locally Scaled Point Processes

	<i>Michaela Prokešová, Ute Hahn, Eva B. Vedel Jensen</i>	99
1	Introduction	99
2	Locally Scaled Point Processes	101
3	Simultaneous Maximum (Pseudo)likelihood Estimation of Scaling Function and Template Parameters	104
4	Two Step Maximum Likelihood Estimation of Scaling Parameters Prior to Template Parameters	106
4.1	Estimation of Scaling Parameters, Using the Poisson Likelihood	107
4.2	Statistical Properties of $\hat{\theta}_0$	109
4.3	Estimation of the Template Parameters	110
5	Simulation Study	111
6	Two Step Inference Where Scaling Function Is Estimated Using Other (non ML) Methods	114
7	Model Validation	115
7.1	The K -function	116
7.2	The Q^2 Statistic	117
8	Data Analysis	118
	References	122

Nonparametric Testing of Distribution Functions in Germ-grain Models

	<i>Zbyněk Pawlas, Lothar Heinrich</i>	125
1	Introduction	125
2	Empirical Distribution Functions	127
3	Weak Convergence of an Empirical Multiparameter Process	128
4	Centres of Pores in a Ceramic Coating	129
5	Conclusion	132
	References	133

Principal Component Analysis for Spatial Point Processes – Assessing the Appropriateness of the Approach in an Ecological Context

	<i>Janine Illian, Erica Benson, John Crawford, Harry Staines</i>	135
1	Introduction	135
2	Methods	136
2.1	Functional Data Analysis	136
2.2	Functional Principal Component Analysis of Second-Order Summary Statistics	137
3	Feasibility Study	138
4	Erroneous Data	140

4.1	Inaccurate Location	142
4.2	Data Collected on a Grid	142
4.3	Wrong Species Recorded	144
5	Application	146
6	Discussion	148
	References	149

Part III Practical Applications of Spatial Point Processes

On Modelling of Refractory Castables by Marked Gibbs and Gibbsian-like Processes

<i>Felix Ballani</i>	153
1 Introduction	153
2 The Data	154
2.1 Production and Data Extraction	155
2.2 Statistics	155
3 The Stationary Marked Gibbs Process	157
4 The Random-Shift Model	159
4.1 Definition	159
4.2 Simulation	160
4.3 Fitting	162
5 Further Models	164
5.1 Dense Random Packings	164
5.2 The Ordered RSA Model	164
6 Conclusions	165
References	165

Source Detection in an Outbreak of Legionnaire's Disease

<i>Miguel A. Martínez-Beneito, Juan J. Abellán, Antonio López-Quílez, Hermelinda Vanaclocha, Óscar Zurriaga, Guillermo Jorques, José Fenollar</i>		169
1	Introduction	169
2	Case Study Description	170
3	Methods	172
4	Results	176
5	Discussion	177
References		181

Doctors' Prescribing Patterns in the Midi-Pyrénées rRegion of France: Point-process Aggregation

<i>Noel A.C. Cressie, Olivier Perrin, Christine Thomas-Agnan</i>		183
1	Introduction	183
2	Spatial Regression Model	185
3	Weighted Regression Analysis	188
4	Weighted Spatial Regression Analysis	191

5	Discussion and conclusions	194
	References	195

Strain-typing Transmissible Spongiform Encephalopathies Using Replicated Spatial Data

<i>Simon Webster, Peter J. Diggle, Helen E. Clough, Robert B. Green, Nigel P. French</i>		197
1	Introduction	197
2	Data Collection	199
	2.1 Sample Material	199
	2.2 Extraction of Data from the Sample Material	201
3	Methods	202
	3.1 Vacuole Counts	202
	3.2 Vacuole Sizes	203
	3.3 Spatial Distribution	203
4	Results	206
	4.1 Vacuole Counts	206
	4.2 Vacuole Sizes	208
	4.3 Spatial Distribution of Vacuoles	210
5	Discussion	212
	References	213

Modelling the Bivariate Spatial Distribution of Amacrine Cells

<i>Peter J. Diggle, Stephen J. Eglen, John B. Troy</i>		215
1	Introduction	215
	1.1 Biological Background	215
2	Pairwise Interaction Point Processes	218
	2.1 Univariate Pairwise Interaction Point Processes	218
	2.2 Bivariate Pairwise Interaction Point Processes	221
3	Monte Carlo Likelihood Inference	222
4	Analysis of the Amacrine Data	223
	4.1 Exploratory Analysis	223
	4.2 Structural Hypotheses for the Amacrine Data	225
	4.3 Non-parametric Estimation	225
	4.4 Univariate Parametric Analysis	226
	4.5 Bivariate Analysis	228
5	Conclusions	230
	5.1 Statistical Summary	230
	5.2 Biological Implications	230
	References	232

Analysis of Spatial Point Patterns in Microscopic and Macroscopic Biological Image Data

<i>Frank Fleischer, Michael Beil, Marian Kazda, Volker Schmidt</i>	235
1 Introduction	235
1.1 Analysis of Centromeric Heterochromatin Structures	236
1.2 Planar Sections of Root Systems in Tree Stands	237
1.3 Detection of Structural Differences	238
2 Image Data	238
2.1 NB4 Cell Nuclei	238
2.2 Profile Walls	239
3 Statistical Methods and Results	241
3.1 3D Point Patterns of Chromocenters in NB4 Cells	241
3.2 2D Point Patterns in Planar Sections of Root Systems	244
3.3 Model Fitting for the Root Data	247
4 Discussion	250
4.1 NB4 Cell Nuclei	250
4.2 Planar Sections of Root Systems	252
References	257

Spatial Marked Point Patterns for Herd Dispersion in a Savanna Wildlife Herbivore Community in Kenya

<i>Alfred Stein, Nick Georgiadis</i>	261
1 Introduction	261
2 Materials and Methods	262
2.1 Distribution Data	262
2.2 Measures of Dispersion	264
3 Results	266
3.1 Spatial Point Patterns	266
4 Discussion	269
5 Conclusions	271
References	272

Diagnostic Analysis of Space-Time Branching Processes for Earthquakes

<i>Jiancang Zhuang, Yosihiko Ogata, David Vere-Jones</i>	275
1 Introduction	275
2 The Space-time ETAS Model	276
3 Data and Preliminary Results	280
4 Residual Analysis for the Whole Process	282
5 Verifying Stationarity of the Background	284
6 Verifying Formulation of Branching Structures by Stochastic Reconstruction	286
7 Conclusion	289
References	290

**Assessing Spatial Point Process Models Using Weighted
 K -functions: Analysis of California Earthquakes**

<i>Alejandro Veen, Frederic Paik Schoenberg</i>	293
1 Introduction	293
2 The Weighted K -function	294
2.1 Ripley's K -function and Variants	294
2.2 Definition and Distribution of the Weighted K -function	295
3 Application	298
3.1 Data Set	298
3.2 Analysis	299
4 Concluding Remarks	303
References	304

List of Contributors

1. **Juan J. Abellán** Small Area Health Statistics Unit, Imperial College of London, UK
2. **Adrian Baddeley** School of Mathematics and Statistics, University of Western Australia, Crawley 6009 WA, Australia
3. **Felix Ballani** Institut für Stochastik, Technische Universität Bergakademie Freiberg, Prüferstr. 9, 09596 Freiberg, Germany
4. **Michael Beil** Department of Internal Medicine I, University Hospital Ulm, D-89070 Ulm, Germany
5. **Erica Benson** SIMBIOS, University of Abertay Dundee, Scotland, UK
6. **Kasper K. Berthelsen** Aalborg University, Department of Mathematical Sciences, F. Bajers Vej 7G, DK-9220 Aalborg, Denmark
7. **Helen E. Clough** Department of Veterinary Clinical Sciences, University of Liverpool, UK
8. **John Crawford** SIMBIOS, University of Abertay Dundee, Scotland, UK
9. **Noel A.C. Cressie** Department of Statistics, The Ohio State University, Columbus, OH 43210, USA
10. **Peter J. Diggle** Department of Statistics, Lancaster University, UK
11. **Stephen J. Eglen** University of Cambridge, UK
12. **José Fenollar** Centro de Salud Pública de Alcoi, Generalitat Valenciana, Spain
13. **Frank Fleischer** Department of Applied Information Processing and Department of Stochastics, University of Ulm, D-89069 Ulm, Germany
14. **Nigel P. French** Institute of Veterinary, Animal and Biomedical Sciences, Massey University, New Zealand.
15. **Nick Georgiadis** Mpala Research Centre, PO Box 555, Nanyuki, Kenya
16. **Robert B. Green** Department of Neuropathology, Veterinary Laboratories Agency, UK
17. **Ute Hahn** University of Augsburg, Department of Applied Stochastics, 86135 Augsburg, Germany
18. **Lothar Heinrich** University of Augsburg, Institute of Mathematics. Universitätsstr. 14, D-86135 Augsburg, Germany

19. **Janine Illian** SIMBIOS, University of Abertay Dundee, Scotland, UK
20. **Eva B. Vedel Jensen** University of Aarhus, The T.N. Thiele Centre of Applied Mathematics in Natural Science, Department of Mathematical Sciences, Ny Munkegade, 8600 Aarhus C, Denmark
21. **Guillermo Jorques** Centro de Salud Pública de Alcoi, Generalitat Valenciana, Spain
22. **Marian Kazda** Department of Systematic Botany and Ecology, University of Ulm, D-89069 Ulm, Germany
23. **Antonio López-Quílez** Departament d'Estadística i Investigació Operativa, Universitat de València, Spain
24. **Miguel A. Martínez-Beneito** Dirección General de Salud Pública, Generalitat Valenciana, Spain
25. **Jesper Møller** Aalborg University, Department of Mathematical Sciences, F. Bajers Vej 7G, DK-9220 Aalborg, Denmark
26. **Yosihiko Ogata** Institute of Statistical Mathematics, Japan
27. **Zbyněk Pawlas** Charles University, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics. Sokolovská 83, 186 75 Praha 8, Czech Republic
28. **Olivier Perrin** GREMAQ – Université Toulouse 1 (and LERNA-INRA), 21 allée de Brienne, 31000 Toulouse, France
29. **Michaela Prokešová** Charles University, Department of Probability, Sokolovská 83, 18675 Praha 8, Czech Republic
30. **Volker Schmidt** Department of Stochastics, University of Ulm, D-89069 Ulm, Germany
31. **Frederic Paik Schoenberg** UCLA Department of Statistics, 8125, Math Sciences Building, Box 951554, Los Angeles, CA 90095-1554, USA
32. **Harry Staines** SIMBIOS, University of Abertay Dundee, Scotland, UK
33. **Alfred Stein** International Institute for Geo-Information Science and Earth Observation (ITC), PO Box 6, 7500 AA Enschede, The Netherlands
34. **Dietrich Stoyan** Institut für Stochastik, Technische Universität Bergakademie Freiberg, Agricolastr. 1, 09596 Freiberg, Germany
35. **Christine Thomas-Agnan** GREMAQ – Université Toulouse 1 (and LSP – Université Toulouse 3), 21 allée de Brienne, 31000 Toulouse, France
36. **John B. Troy** Northwestern University, USA
37. **Rolf Turner** Department of Mathematics and Statistics, University of New Brunswick, Fredericton, N.B., E3B 5A3 Canada
38. **Hermelinda Vanaclocha** Dirección General de Salud Pública, Generalitat Valenciana, Spain
39. **Alejandro Veen** UCLA Department of Statistics, 8125 Math Sciences Building, Box 951554, Los Angeles, CA 90095-1554, USA
40. **David Vere-Jones** Victoria University of Wellington, New Zealand
41. **Simon Webster** Department of Veterinary Clinical Sciences, University of Liverpool, UK
42. **Jiancang Zhuang** Institute of Statistical Mathematics, Japan

43. **Óscar Zurriaga** Dirección General de Salud Pública, Generalitat Valenciana, Spain
44. **Sergei Zuyev** Department of Statistics and Modelling Science, University of Strathclyde, Glasgow, G1 1XH, UK

Fundamentals of Point Process Statistics

Dietrich Stoyan

Institut für Stochastik, Technische Universität Bergakademie Freiberg,
Agricolastr. 1, 09596 Freiberg, Germany, stoyan@orion.hrz.tu-freiberg.de

Summary. Point processes are mathematical models for irregular or random point patterns. A short introduction to the theory of point processes and their statistics, emphasizing connections between the presented theory and the use done by several authors and contributions appearing in this book is presented.

Key words: Marked point processes, Second-order characteristics, Spatial point processes overview, Statistical inference

1 Basic Notions of the Theory of Spatial Point Processes

The following text is a short introduction to the theory of point processes and their statistics, written mainly in order to make this book self-contained. For more information the reader is referred to the text [2] and the books [7, 16, 20, 21].

Point processes are mathematical models for irregular or random point patterns. The mathematical definition of a point process on \mathbb{R}^d is as a random variable N taking values in the measurable space $[\mathbb{N}, \mathcal{N}]$, where \mathbb{N} is the family of all sequences $\{x_n\}$ of points of \mathbb{R}^d satisfying the local finiteness condition, which means that each bounded subset of \mathbb{R}^d contains only a finite number of points. In this book only simple point processes are considered, i.e. $x_i \neq x_j$ if $i \neq j$.

The order of the points x_n is without interest, only the set $\{x_n\}$ matters. Thus the x_n are dummy variables and have no particular interpretation; for example x_1 need not be the point closest to the origin o .

The σ -algebra \mathcal{N} is defined as the smallest σ -algebra of subsets of \mathbb{N} to make measurable all mappings $\varphi \mapsto \varphi(B)$, for B running through the bounded Borel sets.

The reader should note that the term “process” does not imply a dynamic evolution over time and therefore the phrase “random point field” would be a more exact term; it is used in [21]. *Spatio-temporal point processes* explicitly

involving temporal as well as spatial dispersion of points constitute a separate theory, see the paper by Zhuang et al. in this volume.

The *distribution* of a point process N is determined by the probabilities

$$\mathbf{P}(N \in Y) \quad \text{for } Y \in \mathcal{N}.$$

The *finite-dimensional distributions* are of particular importance. These are probabilities of the form

$$\mathbf{P}(N(B_1) = n_1, \dots, N(B_k) = n_k)$$

where B_1, \dots, B_k are bounded Borel sets and n_1, \dots, n_k non-negative integers. Here $N(B_i)$ is the number of points of N in B_i . The distribution of N on $[\mathbb{N}, \mathcal{N}]$ is uniquely determined by the system of all these values for $k = 1, 2, \dots$. A still smaller subsystem is that of the *void probabilities*

$$v_B = \mathbf{P}(N(B) = 0) = \mathbf{P}(N \cap B = \emptyset)$$

for Borel sets B . Here N denotes the set of all points of the point process, the so-called support. If the point process is simple, as assumed here, then the distribution of N is already determined by the system of values of v_K as K ranges through the compact sets.

Let B be a convex compact Borel set in \mathbb{R}^d with o being an inner point of B . The *contact distribution function* H_B with respect to the *test set* B is defined by

$$H_B(r) = 1 - \mathbf{P}(N(rB) = 0) \quad \text{for } r \geq 0. \quad (1)$$

In the special case of $rB = b(o, r)$ = sphere of radius r centred at o the contact distribution function is denoted as $F(r)$ or $H_s(r)$ and called the *spherical contact distribution function* or *empty space distribution function*. It can be interpreted as the distribution function of the random distance from the origin of \mathbb{R}^d to the closest point of N . The function $H_B(r)$ is of a similar nature, but the metric is given by B .

A point process N is said to be *stationary* if its characteristics are invariant under translation: the processes $N = \{x_n\}$ and $N_x = \{x_n + x\}$ have the same distribution for all x in \mathbb{R}^d . So

$$\mathbf{P}(N \in Y) = \mathbf{P}(N_x \in Y) \quad (2)$$

for all Y in \mathcal{N} and all x in \mathbb{R}^d . If we put $Y_x = \{\varphi \in \mathcal{N} : \varphi_{-x} \in Y\}$ for $Y \in \mathcal{N}$ then equation (2) can be rewritten as

$$\mathbf{P}(N \in Y) = \mathbf{P}(N \in Y_{-x}).$$

The notion of *isotropy* is entirely analogous: N is isotropic if its characteristics are invariant under rotation. Stationarity and isotropy together yield *motion-invariance*. The assumption of stationarity simplifies drastically the statistics

of point patterns and therefore many papers in this book assume at least stationarity.

The *intensity measure* Λ of N is a characteristic analogous to the mean of a real-valued random variable. Its definition is

$$\Lambda(B) = \mathbf{E}(N(B)) \quad \text{for Borel } B. \quad (3)$$

So $\Lambda(B)$ is the mean number of points in B . If N is stationary then the intensity measure simplifies; it is a multiple of Lebesgue measure ν_d , i.e.

$$\Lambda(B) = \lambda \nu_d(B) \quad (4)$$

for some (possibly infinite) non-negative constant λ , which is called the *intensity* of N it can be interpreted as the mean number of points of N per unit volume.

The point-related counterpart to $F(r)$ or $H_s(r)$ in the stationary case is the *nearest neighbour distance distribution function* $G(r)$ or $D(r)$, i.e. the d.f. of the distance from the typical point of N to its nearest neighbour.

Note that the application of $F(r)$ and $G(r)$ in the characterization of point processes is different. This is particularly important for cluster processes. In such cases $G(r)$ mainly describes distributional aspects in the clusters, while $F(r)$ characterizes particularly the empty space between the clusters. This different behaviour also explains the success of the *J-function* introduced by [12] defined as

$$J(r) = \frac{1 - G(r)}{1 - F(r)} \quad \text{for } r \geq 0. \quad (5)$$

2 Marked Point Processes

A point process is made into a *marked point process* by attaching a characteristic (the *mark*) to each point of the process. Thus a marked point process on \mathbb{R}^d is a random sequence $M = \{[x_n; m_n]\}$ from which the points x_n together constitute a point process (not marked) in \mathbb{R}^d and the m_n are the marks corresponding to the x_n . The marks m_n may have a complicated structure. They belong to a given *space of marks* \mathbb{M} which is assumed to be a Polish space. The Borel σ -algebra of \mathbb{M} is denoted by \mathcal{M} . Specific examples of marked points are:

- For x the centre of a particle, m the volume of the particle;
- For x the position of a tree, m the stem diameter of the tree;
- For x the centre of an atom, m the type of the atom;
- For x the location (suitably defined) of a convex compact set, m the centred (shifted to origin) set itself.

Point process statistics often uses constructed marks. Examples are:

- m = distance to the nearest neighbour of x ;
- m = number of points within distance r from x .

The marks can be continuous variables, as in the first two examples, indicators of types as in the third example (in which case the terms “multivariate point process” or “multitype point process” are often used, in the case of two marks the term “bivariate point processes”) or actually very complicated indeed, as in the last example which occurs in the marked point process interpretation of germ-grain models (see [20]).

There is a particular feature of marked point processes: Euclidean motions of marked point processes are defined as transforms which move the points but leave the marks unchanged. So M_x , the translate of M by x , is given by

$$M_x = \{[x_1 + x; m_1], [x_2 + x; m_2], \dots\}.$$

Rotations act on marked point processes by rotating the points but *not* altering the marks.

A marked point process M is said to be *stationary* if for all x the translated process M_x has the same distribution as M . It is *motion-invariant* if for all Euclidean motions m the process mM has the same distribution as M .

The definition of the *intensity measure* Λ of a marked point process M is analogous to that of the intensity measure of M when M is interpreted as a non-marked point process:

$$\Lambda(B \times L) = \mathbf{E}(M(B \times L)) .$$

When M is stationary

$$\Lambda = \lambda \times \nu_d \times P_M, \tag{6}$$

where P_M denotes the mark distribution.

3 The Second-order Moment Measure

In the classical theory of random variables the moments (particularly mean and variance) are important tools of statistics. Point process theory has analogues to these. However, numerical means and variances must be replaced by the more complicated moment measures.

The second-order factorial moment measure of the point process N is the measure $\alpha^{(2)}$ on \mathbb{R}^{2d} defined by

$$\int_{\mathbb{R}^{2d}} f(x_1, x_2) \alpha^{(2)}(d(x_1, x_2)) = \mathbf{E} \left(\sum_{x_1, x_2 \in N}^{\neq} f(x_1, x_2) \right) \tag{7}$$

where f is any non-negative measurable function on \mathbb{R}^{2d} . The sum in (7) is extended over all pairs of different points; this is indicated by the symbol \sum^{\neq} .

It is

$$\mathbf{E}(N(B_1)N(B_2)) = \alpha^{(2)}(B_1 \times B_2) + \Lambda(B_1 \cap B_2)$$

and

$$\mathbf{var}(N(B)) = \alpha^{(2)}(B \times B) + \Lambda(B) - (\Lambda(B))^2.$$

If N is stationary then $\alpha^{(2)}$ is translation invariant in an extended sense:

$$\alpha^{(2)}(B_1 \times B_2) = \alpha^{(2)}((B_1 + x) \times (B_2 + x))$$

for all x in \mathbb{R}^d .

Suppose that $\alpha^{(2)}$ is locally finite and absolutely continuous with respect to Lebesgue measure ν_{2d} . Then $\alpha^{(2)}$ has a density $\varrho^{(2)}$, the *second-order product density*:

$$\alpha^{(2)}(B_1 \times B_2) = \int_{B_1} \int_{B_2} \varrho^{(2)}(x_1, x_2) dx_1 dx_2. \quad (8)$$

Moreover, for any non-negative bounded measurable function f

$$\mathbf{E} \left(\sum_{x_1, x_2 \in N}^{\neq} f(x_1, x_2) \right) = \int \int f(x_1, x_2) \varrho^{(2)}(x_1, x_2) dx_1 dx_2.$$

The product density has an intuitive interpretation, which probably accounts for its historical precedence over the product measure and the K -function introduced below. (Note that there are also n^{th} order product densities and moment measures.) Suppose that C_1 and C_2 are disjoint spheres with centres x_1 and x_2 and infinitesimal volumes dV_1 and dV_2 . Then $\varrho^{(2)}(x_1, x_2) dV_1 dV_2$ is the probability that there is each a point of N in C_1 and C_2 . If N is stationary then $\varrho^{(2)}$ depends only on the difference of its arguments and if furthermore N is motion-invariant then it depends only on the distance r between x_1 and x_2 and it is simply written as $\varrho^{(2)}(r)$. The *pair correlation function* $g(r)$ results by normalization:

$$g(r) = \varrho^{(2)}(r) / \lambda^2. \quad (9)$$

Without using the product density $\varrho^{(2)}$, the second factorial moment measure can be expressed by the *second reduced moment measure* \mathcal{K} as

$$\begin{aligned} \alpha^{(2)}(B_1 \times B_2) &= \lambda^2 \int_{B_1} \mathcal{K}(B_2 - x) dx \\ &= \lambda^2 \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbf{1}_{B_1}(x) \mathbf{1}_{B_2}(x+h) \mathcal{K}(dh) dx. \end{aligned} \quad (10)$$