Lecture Notes in Mathematics

1892

A. Baddeley I. Bárány
R. Schneider W. Weil

Stochastic Geometry

Martina Franca, Italy 2004

Editor: W. Weil

With Additional Contributions by
D. Hug, V. Capasso, E. Villa

Springer

Fondazione
C.I.M.E.

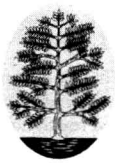
A. Baddeley · I. Bárány
R. Schneider · W. Weil

Stochastic Geometry

Lectures given at the
C.I.M.E. Summer School
held in Martina Franca, Italy,
September 13–18, 2004

With additional contributions by
D. Hug, V. Capasso, E. Villa

Editor: W. Weil



Springer

Fondazione
C.I.M.E.

Authors, Editor and Contributors

Adrian Baddeley
School of Mathematics & Statistics
University of Western Australia
Nedlands WA 6009
Australia
e-mail: adrian@maths.uwa.edu.au

Imre Bárány
Rényi Institute of Mathematics
1364 Budapest Pf. 127
Hungary
e-mail: barany@renyi.hu

and

Mathematics
University College London
Gower Street
London, WC1E 6BT
United Kingdom

Rolf Schneider
Daniel Hug
Mathematisches Institut
Albert-Ludwigs-Universität
Eckerstr. 1
79104 Freiburg i. Br.
Germany
e-mail: rolf.schneider@math.uni-freiburg.de
daniel.hug@math.uni-freiburg.de

Wolfgang Weil
Mathematisches Institut II
Universität Karlsruhe
76128 Karlsruhe
Germany
e-mail: weil@math.uni-karlsruhe.de

Vincenzo Capasso
Elena Villa
Department of Mathematics
University of Milan
via Saldini 50
20133 Milano
Italy
e-mail: vincenzo.capasso@mat.unimi.it
villa@mat.unimi.it

Library of Congress Control Number: 2006931679

Mathematics Subject Classification (2000): Primary 60D05
Secondary 60G55, 62H11, 52A22, 53C65

ISSN print edition: 0075-8434
ISSN electronic edition: 1617-9692
ISBN-10 3-540-38174-0 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-38174-7 Springer Berlin Heidelberg New York

DOI 10.1007/3-540-38174-0

Typesetting by the authors and SPi using a Springer LaTeX package
Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper SPIN: 11815334 41/SPi 5 4 3 2 1 0

Preface

The mathematical treatment of random geometric structures can be traced back to the 18th century (the Buffon needle problem). Subsequent considerations led to the two disciplines *Integral Geometry* and *Geometric Probability*, which are connected with the names of Crofton, Herglotz, Blaschke (to mention only a few) and culminated in the book of Santaló (Integral Geometry and Geometric Probability, 1976). Around this time (the early seventies), the necessity grew to have new and more flexible models for the description of random patterns in Biology, Medicine and Image Analysis. A theory of *Random Sets* was developed independently by D.G. Kendall and Matheron. In connection with Integral Geometry and the already existing theory of Point Processes the new field *Stochastic Geometry* was born. Its rapid development was influenced by applications in *Spatial Statistics* and *Stereology*. Whereas at the beginning emphasis was laid on models based on stationary and isotropic Poisson processes, the recent years showed results of increasing generality, for nonisotropic or even inhomogeneous structures and without the strong independence properties of the Poisson distribution. On the one side, these recent developments in Stochastic Geometry went hand-in-hand with a fresh interest in Integral Geometry, namely local formulae for curvature measures (in the spirit of Federer's Geometric Measure Theory). On the other side, new models of point processes (Gibbs processes, Strauss processes, hardcore and cluster processes) and their effective simulation (Markov Chain Monte Carlo, perfect simulation) tightened the close relation between Stochastic Geometry and Spatial Statistics. A further, very interesting direction is the investigation of spatial-temporal processes (tumor growth, communication networks, crystallization processes). The demand for random geometric models is steadily growing in almost all natural sciences or technical fields.

The intention of the Summer School was to present an up-to-date description of important parts of Stochastic Geometry. The course took place in Martina Franca from Monday, September 13, to Friday, September 18, 2004. It was attended by 49 participants (including the lecturers). The main lecturers were Adrian Baddeley (University of Western Australia, Perth), Imre

Bárány (University College, London, and Hungarian Academy of Sciences, Budapest), Rolf Schneider (University of Freiburg, Germany) and Wolfgang Weil (University of Karlsruhe, Germany). Each of them gave four lectures of 90 minutes which we shortly describe, in the following.

Adrian Baddeley spoke on **Spatial Point Processes and their Applications**. He started with an introduction to point processes and marked point processes in $\mathbb{R}^d$ as models for spatial data and described the basic notions (counting measures, intensity, finite-dimensional distributions, capacity functional). He explained the construction of the basic model in spatial statistics, the Poisson process (on general locally compact spaces), and its transformations (thinning and clustering). He then discussed higher order moment measures and related concepts (K function, pair correlation function). In his third lecture, he discussed conditioning of point processes (conditional intensity, Palm distributions) and the important Campbell-Mecke theorem. The Palm distributions lead to G and J functions which are of simple form for Poisson processes. In the last lecture he considered point processes in bounded regions and described methods to fit corresponding models to given data. He illustrated his lectures by computer simulations.

Imre Bárány spoke on **Random Points, Convex Bodies, and Approximation**. He considered the asymptotic behavior of functionals like volume, number of vertices, number of facets, etc. of random convex polytopes arising as convex hulls of n i.i.d. random points in a convex body $K \subset \mathbb{R}^d$. Starting with a short historical introduction (Efron's identity, formulas of Rényi and Sulanke), he emphasized the different limit behavior of expected functionals for smooth bodies K on one side and for polytopes K on the other side. In order to explain this difference, he showed that the asymptotic behavior of the expected missed volume $E(K, n)$ of the random convex hull behaves asymptotically like the volume of a deterministic set, namely the shell between K and the cap body of K (the floating body). This result uses Macbeath regions and the 'economic cap covering theorem' as main tools. The results were extended to the expected number of vertices and of facets. In the third lecture, random approximation (approximation of K by the convex hull of random points) was compared with best approximation (approximation from inside w.r.t. minimal missed volume). It was shown that random approximation is almost as good as best approximation. A further comparison concerned convex hulls of lattice points in K. In the last lecture, for a convex body $K \subset \mathbb{R}^2$, the probability $p(n, K)$ that n random points in K are in convex position was considered and the asymptotic behavior (as $n \to \infty$) was given (extension of the classical Sylvester problem).

The lectures of Rolf Schneider concentrated on **Integral Geometric Tools for Stochastic Geometry**. In the first lecture, the classical results from integral geometry, the principal kinematic formulas and the Crofton formulas were given in their general form, for intrinsic volumes of convex bodies (which were introduced by means of the Steiner formula). Then, Hadwiger's characterization theorem for additive functionals was explained and used to

generalize the integral formulas. In the second lecture, local versions of the integral formulas for support measures (curvature measures) and extensions to sets in the convex ring were discussed. This included a local Steiner formula for convex bodies. Extensions to arbitrary closed sets were mentioned. The third lecture presented translative integral formulas, in local and global versions, and their iterations. The occurring mixed measures and functionals were discussed in more detail and connections to support functions and mixed volumes were outlined. The last lecture studied general notions of k-dimensional area and general Crofton formulas. Relations between hyperplane measures and generalized zonoids were given. It was shown how such relations can be used in stochastic geometry, for example, to give estimates for the intersection intensity of a general (non-stationary) Poisson hyperplane process in $\mathbb{R}^d$.

Wolfgang Weil, in his lectures on **Random Sets (in Particular Boolean Models)**, built upon the previous lectures of A. Baddeley and R. Schneider. He first gave an introduction to random closed sets and particle processes (point processes of compact sets, marked point processes) and introduced the basic model in stochastic geometry, the Boolean model (the union set of a Poisson particle process). He described the decomposition of the intensity measure of a stationary particle process and used this to introduce the two quantities which characterize a Boolean model (intensity and grain distribution). He also explained the role of the capacity functional (Choquet's theorem) and its explicit form for Boolean models which shows the relation to Steiner's formula. In the second lecture, mean value formulas for additive functionals were discussed. They lead to the notion of density (quermass density, density of area measure, etc.) which was studied then for general random closed sets and particle processes. The principal kinematic and translative formulas were used to obtain explicit formulas for quermass densities of stationary and isotropic Boolean models as well as for non-isotropic Boolean models (with convex or polyconvex grains) in $\mathbb{R}^d$. Statistical consequences were discussed for $d = 2$ and $d = 3$ and ergodic properties were shortly mentioned. The third lecture was concerned with extensions in various directions: densities for directional data and their relation to associated convex bodies (with an application to the mean visible volume of a Boolean model), interpretation of densities as Radon-Nikodym derivatives of associated random measures, density formulas for non-stationary Boolean models. In the final lecture, random closed sets and Boolean models were investigated from outside by means of contact distributions. Recent extensions of this concept were discussed (generalized directed contact distributions) and it was explained that in some cases they suffice to determine the grain distribution of a Boolean model completely. The role of convexity for explicit formulas of contact distributions was discussed and, as the final result, it was explained that the polynomial behavior of the logarithmic linear contact distribution of a stationary and isotropic Boolean model characterizes convexity of the grains.

Since the four lecture series could only cover some parts of stochastic geometry, two additional lectures of 90 minutes were included in the program,

given by D. Hug and V. Capasso. Daniel Hug (University of Freiburg) spoke on **Random Mosaics** as special particle processes. He presented formulas for the different intensities (number and content of faces) for general mosaics and for Voronoi mosaics and then explained a recent solution to Kendall's conjecture concerning the asymptotic shape of large cells in a Poisson Voronoi mosaic. Vincenzo Capasso (University of Milano) spoke on **Crystallization Processes** as spatio-temporal extensions of point processes and Boolean models and emphasized some problems arising from applications.

The participants presented themselves in some short contributions, at one afternoon, as well as in two evening sessions.

The attendance of the lectures was extraordinarily good. Most of the participants had already some background in spatial statistics or stochastic geometry. Nevertheless, the lectures presented during the week provided the audience with a lot of new material for subsequent studies. These lecture notes contain (partially extended) versions of the four main courses (and the two additional lectures) and are also intended as an information of a wider readership about this important field. I thank all the authors for their careful preparation of the manuscripts.

I also take the opportunity, on behalf of all participants, to thank C.I.M.E. for the effective organization of this summer school; in particular, I want to thank Vincenzo Capasso who initiated the idea of a workshop on stochastic geometry. Finally, we were all quite grateful for the kind hospitality of the city of Martina Franca.

Karlsruhe, August 2005 Wolfgang Weil

Contents

Spatial Point Processes and their Applications
Adrian Baddeley 1
1 Point Processes 2
1.1 Point Processes in 1D and 2D 2
1.2 Formulation of Point Processes 3
1.3 Example: Binomial Process 6
1.4 Foundations 7
1.5 Poisson Processes 8
1.6 Distributional Characterisation 12
1.7 Transforming a Point Process 16
1.8 Marked Point Processes 19
1.9 Distances in Point Processes 21
1.10 Estimation from Data 23
1.11 Computer Exercises 24
2 Moments and Summary Statistics 26
2.1 Intensity 26
2.2 Intensity for Marked Point Processes 30
2.3 Second Moment Measures 32
2.4 Second Moments for Stationary Processes 35
2.5 The K-function 38
2.6 Estimation from Data 39
2.7 Exercises 40
3 Conditioning 42
3.1 Motivation 42
3.2 Palm Distribution 44
3.3 Palm Distribution for Stationary Processes 49
3.4 Nearest Neighbour Function 51
3.5 Conditional Intensity 52
3.6 J-function 55
3.7 Exercises 56
4 Modelling and Statistical Inference 57

4.1 Motivation 57
4.2 Parametric Modelling and Inference 58
4.3 Finite Point Processes 61
4.4 Point Process Densities 62
4.5 Conditional Intensity 64
4.6 Finite Gibbs Models 66
4.7 Parameter Estimation 69
4.8 Estimating Equations 70
4.9 Likelihood Devices 72
References 73

Random Polytopes, Convex Bodies, and Approximation
Imre Bárány 77
1 Introduction 77
2 Computing $\mathbb{E}\phi(K_n)$ 79
3 Minimal Caps and a General Result 80
4 The Volume of the Wet Part 82
5 The Economic Cap Covering Theorem 84
6 Macbeath Regions 84
7 Proofs of the Properties of the M-regions 87
8 Proof of the Cap Covering Theorem 89
9 Auxiliary Lemmas from Probability 92
10 Proof of Theorem 3.1 95
11 Proof of Theorem 4.1 96
12 Proof of (4) 98
13 Expectation of $f_k(K_n)$ 101
14 Proof of Lemma 13.2 102
15 Further Results 104
16 Lattice Polytopes 108
17 Approximation 109
18 How It All Began: Segments on the Surface of K 114
References 115

Integral Geometric Tools for Stochastic Geometry
Rolf Schneider 119
Introduction 119
1 From Hitting Probabilities to Kinematic Formulae 120
1.1 A Heuristic Question on Hitting Probabilities 120
1.2 Steiner Formula and Intrinsic Volumes 123
1.3 Hadwiger's Characterization Theorem for Intrinsic Volumes 126
1.4 Integral Geometric Formulae 129
2 Localizations and Extensions 136
2.1 The Kinematic Formula for Curvature Measures 136
2.2 Additive Extension to Polyconvex Sets 143
2.3 Curvature Measures for More General Sets 148

3 Translative Integral Geometry 151
3.1 The Principal Translative Formula for Curvature Measures 152
3.2 Basic Mixed Functionals and Support Functions 157
3.3 Further Topics of Translative Integral Geometry 163
4 Measures on Spaces of Flats 164
4.1 Minkowski Spaces 165
4.2 Projective Finsler Spaces 173
4.3 Nonstationary Hyperplane Processes 177
References 181

Random Sets (in Particular Boolean Models)
Wolfgang Weil 185
Introduction 185
1 Random Sets, Particle Processes and Boolean Models 186
1.1 Random Closed Sets 187
1.2 Particle Processes 190
1.3 Boolean Models 194
2 Mean Values of Additive Functionals 198
2.1 A General Formula for Boolean Models 199
2.2 Mean Values for RACS 204
2.3 Mean Values for Particle Processes 208
2.4 Quermass Densities of Boolean Models 210
2.5 Ergodicity 214
3 Directional Data, Local Densities, Nonstationary Boolean Models 214
3.1 Directional Data and Associated Bodies 215
3.2 Local Densities 222
3.3 Nonstationary Boolean Models 224
3.4 Sections of Boolean Models 230
4 Contact Distributions 231
4.1 Contact Distribution with Structuring Element 232
4.2 Generalized Contact Distributions 237
4.3 Characterization of Convex Grains 241
References 243

Random Mosaics
Daniel Hug 247
1 General Results 247
1.1 Basic Notions 247
1.2 Random Mosaics 248
1.3 Face-Stars 250
1.4 Typical Cell and Zero Cell 252
2 Voronoi and Delaunay Mosaics 253
2.1 Voronoi Mosaics 253
2.2 Delaunay Mosaics 255
3 Hyperplane Mosaics 257

4 Kendall's Conjecture 260
4.1 Large Cells in Poisson Hyperplane Mosaics 261
4.2 Large Cells in Poisson-Voronoi Mosaics 263
4.3 Large Cells in Poisson-Delaunay Mosaics 264
References 265

On the Evolution Equations of Mean Geometric Densities for a Class of Space and Time Inhomogeneous Stochastic Birth-and-growth Processes
Vincenzo Capasso, Elena Villa 267
1 Introduction 267
2 Birth-and-growth Processes 268
2.1 The Nucleation Process 268
2.2 The Growth Process 270
3 Closed Sets as Distributions 271
3.1 The Deterministic Case 271
3.2 The Stochastic Case 273
4 The Evolution Equation of Mean Densities for the Stochastic Birth-and-growth Process 274
4.1 Hazard Function and Causal Cone 276
References 280

Spatial Point Processes and their Applications

Adrian Baddeley

School of Mathematics & Statistics, University of Western Australia
Nedlands WA 6009, Australia
e-mail: adrian@maths.uwa.edu.au

A spatial point process is a random pattern of points in d-dimensional space (where usually $d = 2$ or $d = 3$ in applications). Spatial point processes are useful as statistical models in the analysis of observed patterns of points, where the points represent the locations of some object of study (e..g. trees in a forest, bird nests, disease cases, or petty crimes). Point processes play a special role in stochastic geometry, as the building blocks of more complicated random set models (such as the Boolean model), and as instructive simple examples of random sets.

These lectures introduce basic concepts of spatial point processes, with a view toward applications, and with a minimum of technical detail. They cover methods for constructing, manipulating and analysing spatial point processes, and for analysing spatial point pattern data. Each lecture ends with a set of practical computer exercises, which the reader can carry out by downloading a free software package.

Lecture 1 ('Point Processes') gives some motivation, defines point processes, explains how to construct point processes, and gives some important examples. Lecture 2 ('Moments') discusses means and higher moments for point processes, especially the intensity measure and the second moment measure, along with derived quantities such as the K-function and the pair correlation function. It covers the important Campbell formula for expectations. Lecture 3 ('Conditioning') explains how to condition on the event that the point process has a point at a specified location. This leads to the concept of the Palm distribution, and the related Campbell-Mecke formula. A dual concept is the conditional intensity, which provides many new results. Lecture 4 ('Modelling and Statistical Inference') covers the formulation of statistical models for point patterns, model-fitting methods, and statistical inference.

1 Point Processes

In this first lecture, we motivate and define point processes, construct examples (especially the **Poisson process** [28]), and analyse important properties of the Poisson process. There are different ways to mathematically construct and characterise a point process (using finite-dimensional distributions, vacancy probabilities, capacity functional, or generating function). An easier way to construct a point process is by transforming an existing point process (by thinning, superposition, or clustering) [43]. Finally we show how to use existing software to generate simulated realisations of many spatial point processes using these techniques, and analyse them using vacancy probabilities (or 'empty space functions').

1.1 Point Processes in 1D and 2D

A **point process** in one dimension ('time') is a useful model for the sequence of random times when a particular event occurs. For example, the random times when a hospital receives emergency calls may be modelled as a point process. Each emergency call happens at an instant, or point, of time. There will be a random number of such calls in any period of time, and they will occur at random instants of time.

Fig. 1. A point process in time.

A **spatial point process** is a useful model for a random pattern of points in d-dimensional space, where $d \geq 2$. For example, if we make a map of the locations of all the people who called the emergency service during a particular day, this map constitutes a random pattern of points in two dimensions. There will be a random number of such points, and their locations are also random.

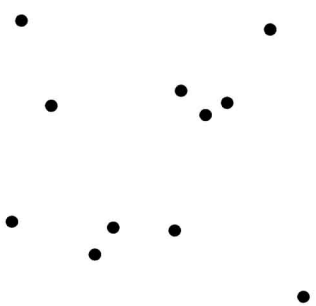


Fig. 2. A point process in two dimensions.

We may also record both the locations and the times of the emergency calls. This may be regarded as a point process in three dimensions (space $\times$ time), or alternatively, as a point process in two dimensions where each point (caller location) is labelled or **marked** by a number (the time of the call).

Spatial point processes can be used directly, to model and analyse data which take the form of a point pattern, such as maps of the locations of trees or bird nests ('statistical ecology' [16, 29]); the positions of stars and galaxies ('astrostatistics' [1]); the locations of point-like defects in a silicon crystal wafer (materials science [34]); the locations of neurons in brain tissue; or the home addresses of individuals diagnosed with a rare disease ('spatial epidemiology' [19]). Spatial point processes also serve as a basic model in random set theory [42] and image analysis [41]. For general surveys of applications of spatial point processes, see [16, 42, 43]. For general theory see [15].

1.2 Formulation of Point Processes

There are some differences between the theory of one-dimensional and higher-dimensional point processes, because one-dimensional time has a natural ordering which is absent in higher dimensions.

A one-dimensional point process can be handled mathematically in many different ways. We may study the **arrival times** $T_1 < T_2 < \ldots$ where T_i is the time at which the ith point (emergency call) arrives. Using these random variables is the most direct way to handle the point pattern, but their use is complicated by the fact that they are strongly dependent, since $T_i < T_{i+1}$.

Fig. 3. Arrival times T_i.

Alternatively we may study the **inter-arrival times** $S_i = T_{i+1} - T_i$. These have the advantage that, for some special models (Poisson and renewal processes), the random variables $S_1, S_2, \ldots$ are independent.

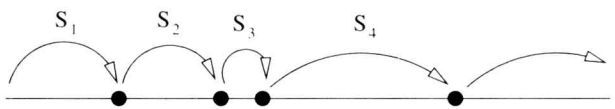


Fig. 4. Inter-arrival times S_i.

Alternatively it is common (especially in connection with martingale theory) to formulate a point process in terms of the cumulative **counting process**

$$N_t = \text{number of points arriving up to time } t$$
$$= \sum_{i=1}^{\infty} \mathbf{1}\{T_i \le t\},$$

for all $t \ge 0$, where $\mathbf{1}\{\ldots\}$ denotes the indicator function, equal to 1 if the statement "..." is true, and equal to 0 otherwise. This device has the advantage of converting the process to a random function of continuous time t, but has the disadvantage that the values N_t for different t are highly dependent.

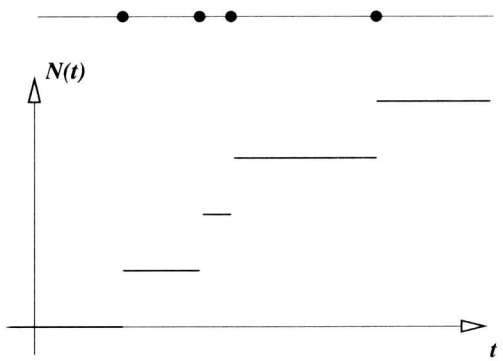


Fig. 5. The counting process N_t associated with a point process.

Alternatively one may use the interval counts

$$N(a, b] = N_b - N_a$$

for $0 \le a \le b$ which count the number of points arriving in the interval $(a, b]$. For some special processes (Poisson and independent-increments processes) the interval counts for **disjoint** intervals are stochastically independent.

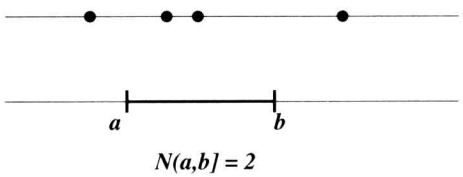


Fig. 6. Interval count $N(a, b]$ for a point process.

In higher dimensions, there is no natural ordering of the points, so that there is no *natural* analogue of the inter-arrival times S_i nor of the counting process N_t. Instead, the most useful way to handle a spatial point process is to generalise the interval counts $N(a, b]$ to the region counts

$$N(B) = \text{number of points falling in } B$$

defined for each bounded closed set $B \subset \mathbb{R}^d$.

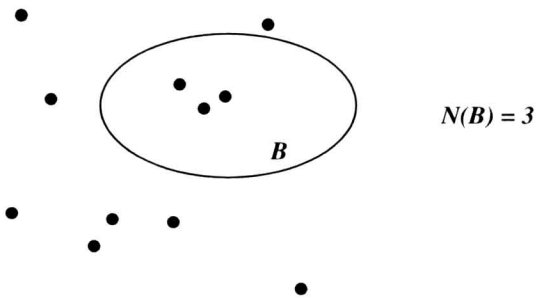


Fig. 7. Counting variables $N(B)$ for a spatial point process.

Rather surprisingly, it is often sufficient to study a point process using only the **vacancy indicators**

$$\begin{aligned} V(B) &= \mathbf{1}\{N(B) = 0\} \\ &= \mathbf{1}\{\text{there are no points falling in } B\}. \end{aligned}$$

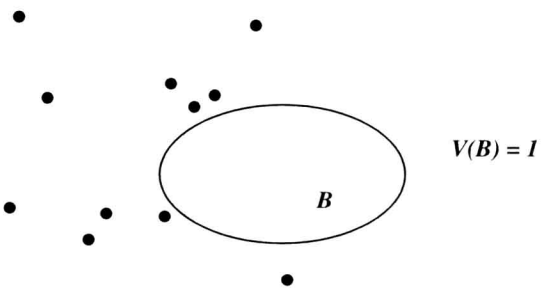


Fig. 8. Vacancy indicators $V(B)$ for a spatial point process.

The counting variables $N(B)$ are natural for exploring additive properties of a point process. For example, suppose we have two point processes, of 'red' and 'blue' points respectively, and we superimpose them (forming a single point process by discarding the colours). If $N_{\text{red}}(B)$ and $N_{\text{blue}}(B)$ are the counting variables for red and blue points respectively, then the counting variable for the superimposed process is $N(B) = N_{\text{red}}(B) + N_{\text{blue}}(B)$.

The vacancy indicators $V(B)$ are natural for exploring geometric and 'multiplicative' properties of a point process. If $V_{\text{red}}(B)$ and $V_{\text{blue}}(B)$ are the vacancy indicators for two point processes, then the vacancy indicator for the superimposed process is $V(B) = V_{\text{red}}(B)\, V_{\text{blue}}(B)$.

1.3 Example: Binomial Process

To take a very simple example, let us place a fixed number n of points at random locations inside a bounded region $W \subset \mathbb{R}^2$. Let $X_1, \ldots, X_n$ be i.i.d. (independent and identically distributed) random points which are uniformly distributed in W. Hence the probability density of each X_i is

$$f(x) = \begin{cases} 1/\lambda_2(W) & \text{if } x \in W \\ 0 & \text{otherwise} \end{cases}$$

where $\lambda_2(W)$ denotes the area of W. A realisation of this process is shown in Figure 9.

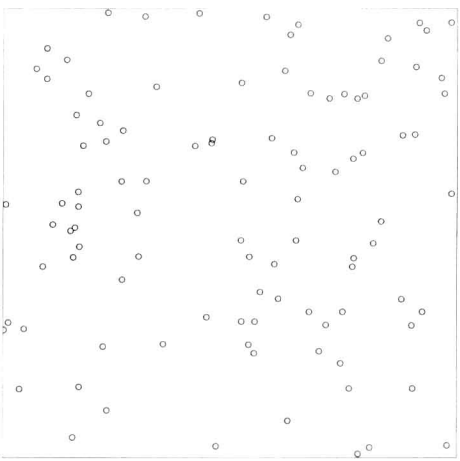


Fig. 9. Realisation of a binomial point process with $n = 100$ in the unit square.

Since each random point X_i is uniformly distributed in W, we have for any bounded set B in $\mathbb{R}^2$

$$\begin{aligned} \mathbb{P}(X_i \in B) &= \int_B f(x)\, \mathrm{d}x \\ &= \frac{\lambda_2(B \cap W)}{\lambda_2(W)}. \end{aligned}$$

The variables $N(B)$ and $V(B)$ may be represented explicitly as

$$N(B) = \sum_{i=1}^{n} \mathbf{1}\{X_i \in B\}$$

$$V(B) = \min_{i=1}^{n} \mathbf{1}\{X_i \notin B\}$$