STOCHASTIC POINT PROCESSES:

Statistical Analysis, Theory and Applications

Edited by Peter A. W. Lewis

Stochastic Point Processes:

Statistical Analysis, Theory, and Applications

PETER A. W. LEWIS, Editor

Naval Postgraduate School
Monterey, California

WILEY - INTERSCIENCE
a division of John Wiley & Sons, Inc.

New York London Sydney Toronto

Library of Congress Cataloging in Publication Data:

Main entry under title:

Stochastic point processes.

(Wiley series in probability and mathematical statistics)
"Contains the bulk of the papers presented at a conference held at the IBM Center, Yorktown Heights, New York, on August 2-7, 1971."
1. Stochastic processes--Congresses. 2. Analysis of variance--Congresses. I. Lewis, Peter A. W., 1932 ed.
QA274.A1S76 519.2 72-3954
ISBN 0-471-53350-5

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Preface

This volume contains the bulk of the papers presented at a conference held at the IBM Research Center, Yorktown Heights, New York on August 2-6, 1971. The title of the meeting was "Stochastic Point Processes: Statistical Analysis, Theory and Applications."

The idea of holding the meeting grew out of a seminar series and some special lectures on stochastic point processes given in the Mathematics Department at Imperial College, University of London, during the academic year 1969-70, and I wish to express my thanks to the National Institutes of Health for providing me with the fellowship which made my presence at the seminar possible. At the end of that seminar series, it became evident to Professor D. R. Cox and myself that it would be timely to hold a fairly comprehensive conference to pull together the many aspects of point processes, all rapidly developing, and to bring together the applied and theoretical people in the field.

Since the International Statistical Institute (ISI) was to hold its biennial meeting in Washington, D. C., from August 10-20, 1971, a date was set for August 2-6, 1971, for the point process conference and an organizing committee was formed. The committee members were P.A.W. Lewis, Chairman, M. S. Bartlett, D. R. Cox, J. Gani, K. Matthes, P. A. P. Moran, E. Parzen, R. Pyke, W. L. Smith, and D. Vere-Jones. Subsequently the conference became one of the satellite symposia organized by Professor J. Neyman and the International Association for Statistics in the Physical Sciences (IASPS) in conjunction with the ISI meeting in Washington, D. C.. Other satelite symposia were held on statistical problems in the fields of turbulence, metallurgy, pollution, and hydrology.

The conference was started with a small grant from the IBM Mathematical Sciences Department through Drs. R. E. Gomory and S. Winograd. To them and the IBM Research Division, who provided physical facilities and numerous services, I am deeply grateful. The main financial support for the conference was provided by the Office of Naval Research (ONR) and the National Science Foundation (NSF) and thanks are due to Dr. Bruce McDonald of ONR and Dr. William Pell of NSF for all of their help.

On the technical side much help was provided, in addition to the committee members, by Dr. D. J. Daley, Dr. A. J. Lawrance and Professor M. R. Leadbetter. The splendid organization of the conference was due primarily to Mrs. M. Casey of IBM, Dr. A. J. Lawrance (temporarily of IBM) and Dr. W. Turner of IBM. To Mrs. Casey, too, go thanks for constant good cheer and gentle direction under often trying circumstances.

Responsibility for editing this volume is my own. Enormous help was provided, especially with the initial editing at the time of the conference, by Professors Cox and Gani. Subsequent editorial help was also provided by Dr. D. J. Daley, Dr. A. J. Lawrance, Professor D. Vere-Jones and Dr. M. Westcott. The primary aim in editing the volume was to produce the volume as soon as possible after the conference; to do this some liberty had to be taken with editorial niceties. Thus while notational conventions, etc., are substantially the same, certain concessions were made to the very different conventions prevailing in different fields. Also no effort was made to provide absolute consistency in reference to obscure journals and articles.

Thanks for typing of the volume are due to Mrs. Ruth Guthrie and Miss Rosemarie Stampfel.

Finally, the editing of the volume was done at the Naval Postgraduate School, Monterey, California, to whom I am grateful for support.

P. A. W. Lewis
Naval Postgraduate School
March, 1972

INTRODUCTION

P. A. W. Lewis
Naval Postgraduate School, Monterey, California

The aim of this introduction is to provide a brief summary of the papers in this volume and to put them in the context of the whole field of point process methodology as it is developing today. The papers do represent a good cross-section of the field, but time and space limitations induced some truncation. Thus, for example, interest in and development of techniques for handling doubly stochastic Poisson process has burgeoned very recently, but it was only possible to devote one paper (Grandell, p. 90) to it completely, although some aspects of the problems involved are touched on in Lewis (p. 1), Lawrance (p. 199), Daley and Vere-Jones (p. 299) and Krickeberg (p. 514). Some truncation was also produced by the limitations imposed by my own interests; as the organization of the conference proceeded, and more particularly during the discussion at the conference, my own knowledge of the field and its interrelationship with other aspects of stochastic process theory grew. However, by that time the program was already set!

It is perhaps worthwhile to start by reproducing the prospectus from the original call for papers for the conference.

"The aim of the conference is to bring together mathematicians and statisticians working in the field of point processes and workers in applied fields such as ecology, neurophysiology, traffic studies, reliability, geography, forestry, epidemiology, physics and geophysics. Consequently, there will be three categories of papers presented at the conference;

i) Invited survey papers on the mathematical theory, statistical analysis and models of univariate point processes, multivariate point processes, multidimensional point processes and line processes;

ii) Invited review papers on the types of problems involving point processes encountered in fields of application such as ecology, neurophysiology, physics, forestry, reliability, traffic, geography, etc.;

iii) A very limited number of contributed and invited papers on new work in the field."

Two problems arose from this call.

One was that the distinction between survey and review papers, never really too clear, became blurred. Moreover many invited survey and/or review papers missed some of the coverage which was asked for when they were solicited, and also overlapped into other areas. This of course is inevitable and desirable in light of the professed aim of the conference to join together some of the disparate and overlapping areas of endeavour.

A second problem was that of nomenclature; this was discussed at some length at the conference. The general name "point process" appears to go back to Wold (1948) and is now fairly widely used; other names are "event processes," "series of events," "streams of events" and "flows of events."

On the other hand, the taxonomy of point processes is not so easy to resolve. The usage "univariate point processes" in this volume refers to events occurring in a one dimensional continuum, usually time, the events being distinguishable only by where they occur, i.e., having no qualitative or quantitative information attached to them. The usage "multidimensional point processes" refers to the same situation where the events occur in a multidimensional continuum. However, "multidimensional point processes" is also used to refer to the situation where different types of events occur (events distinguished by qualitative information), especially in the Russian literature. The tenor of the discussion at the conference was that "multivariate point processes" (Cox and Lewis, 1972) was preferable in this latter situation, since events of two types occur in higher dimensions, e.g., different types of trees in a forest. This latter terminology has been adopted here. It is also perhaps worthwhile to point out that there are two categories of multivariate point processes; in the one the typing arises because of qualitative information and should perhaps be called multitype point processes; in the other the typing is in fact because of physical (time or space) separation of events (e.g., arrivals and departures from a queue; this spring's flowers and last spring's flowers) and should perhaps be called multiple point processes, following Hannan (1970).

Another fairly confused usage is that of "line processes." Bartlett (1967) used this to describe processes in which the events were accompanied by quantitative information. For example, earthquakes not only occur at "points" in time, but also have space and energy information attached to them. This is a special case of a general situation defined by Matthes (1963) and called "marked point processes" (see Daley and Vere-Jones, p. 299). This designation of the situation where events have quantitative information attached to them seems best referred to as marked point processes, even though Matthes' definition is broad enough to cover multitype processes too.

True "line processes" do occur. For example, planes flying in the air and particles in space or bubble chambers, the "lines" usually being of finite length, although it is mathematically convenient to think of infinite lines. These "line processes" when sampled at several instants of time produce multiple, multivariate point processes, the space being usually three dimensional. A survey of line process theory is given by Krickeberg (p. 514), who calls them "hyperplane processes." This perhaps is a better name than "line processes."

A related situation which occurs frequently in physics (Ramakrishnan, p. 533) is that of events occurring in time and space. This is distinguished from multidimensional point processes because the time and space dimensions usually do not occur on an equal footing. Sampling of these spatial point processes by pooling events occurring in a time period produces ordinary point processes, e.g., the records of places of occurrence of traffic accidents on successive days.

Analysis and discussion of these "line processes" or "hyperplane processes" and "spatial point processes" by mathematicians and statisticians has barely begun and they are clearly a broad enough category that a detailed attempt at naming would be futile. Moreover taxonomies are at best leaky.

To conclude the discussion it should be pointed out that these line processes occur in metallurgy, and the interested reader might also refer to the papers presented at the IASPS Satellite Symposium on Metallurgy which will appear in the journal "Advances in Applied Probability," in 1972.

We now describe briefly the papers in the volume and their interrelationships. The volume contains all of the papers presented at the conference except three. The paper by Milne and Westcott (p. 257) is a summary of a long paper which will appear in "Advances in Applied Probability" in 1972, as will the paper presented at the conference by J. E. Moyal entitled, "Number operators and population processes in quantum field theories." Dr. Moyol graciously offered to submit his paper to "Advances in Applied Probability" when it became evident that this present volume was far too long to be practicable. The paper by W. L. Smith entitled "Queues; a broad view" which was presented at the conference will appear elsewhere. The papers by Grigelionis and Ambartzumian were scheduled for the conference, but the authors were unable to present them in person.

The organization of the papers in the volume is not exactly the same as the session organization at the conference. This has been done partly because of editorial convenience and partly because the overlaps and abutments of the papers were clearer in hindsight. Thus there is an implied classification of the papers by their appearance in Statistical Analysis, Models, Theory or Applications and the order in which they appear in those four categories. Again, the classification is leaky and only partially valid. The paper by Neyman and Scott, for instance, describes a long term effort to apply statistics and probability to the analysis of the dispersion of stars in space. However, it also describes some important results on cluster models and the statistical analysis of cluster point processes.

Starting with the section entitled Statistical Analysis, the summary paper by Lewis (p. 1) attempts to survey results obtained on the statistical analysis of univariate point processes since the publication of the monograph by Cox and Lewis (1966). There is particular emphasis on new results, both theoretical and tabular, on tests for renewal processes and on (time) trend analysis in non-homogeneous Poisson processes. Both cyclical and evolutionary trends are considered, and the methodology is illustrated by results obtained on the analysis of a very long series of arrivals at an intensive care unit. Cox (p. 55) gives a general model for the analysis of trends in Poisson processes and extends it to time-dependent (modulated) renewal processes. The methodology given for handling these situations is then used to derive formally a test for dependence in a bivariate point process which was given in Cox and Lewis (1972). The paper by Brown (p. 67) supplies much of the theory needed to rigorize the models and methodology for non-homogeneous Poisson processes given in the papers by Lewis (p. 1) and by Cox (p. 55). Other aspects of the analysis of these processes are also considered.

Methods based on least squares theory for making inferences in the important class of doubly stochastic point processes are given by Grandell (p. 90). Analysis of doubly stochastic Poisson processes is presently a very active field, with interest in applications in physics, engineering and medicine. References to the recent work by, in particular, Snyder, Karp and Clark is given in Lewis (Section 4.3). Other recent work is by Macchi (1971). Krickeberg (p. 514) refers to these doubly stochastic Poisson processes as "Cox processes;" they were first discussed in Cox (1955).

The fifth paper in this section is a survey paper by Holgate (p. 122) on the use of distance methods for analyzing spatial patterns; the fields of application are very numerous.

Bartlett (p. 136) takes up two important problems in multivariate and multidimensional point processes which are only just beginning to receive attention. One problem is the analysis of dependencies between point processes and continuous stochastic processes; in Bartlett's paper these are the times at which vehicles pass a point on a road and their velocities. In addition, typing by which lane the vehicle is in is added and spectral methods are suggested for the analysis of data. A similar problem in multivariate two-dimensional point processes is also discussed. The papers by Gazis and Szeto (p. 151) and Newell and Sparks

(p. 166) give analyses of two specific situations involving marked multivariate point processes occurring in traffic studies.

Finally the paper by Sachs (p. 175) in this section on Statistical Analysis is an attempt to inform statisticians of the very complex and large scale point process problems which are faced and analyzed by physicists quite routinely. Both the types of problems and the techniques of analysis are described.

The section on Models starts with a survey paper by Lawrance (p. 199) of models for stationary series of events. There is a table of contents on pages 199 and 200. The coverage is primarily of processes (such as semi-Markov processes) with some kind of Markovian dependency, generally defined on the intervals between events; doubly stochastic Poisson processes of several kinds, and cluster processes. The emphasis is on aspects, such as arbitrary event and arbitrary time initial conditions and marginal distributions of interevent intervals, which are useful in modeling and statistical analysis. The paper by Milne and Westcott (p. 257) summarizes results the authors have obtained on a process, defined by Newman, through its characteristic functional and called a Gauss-Poisson process. The authors have shown it to be a special case of the Bartlett-Lewis cluster process (see the paper by Lawrance, p. 237 for a description of the Bartlett-Lewis process). Hawkes' paper (p. 261) extends results on the mutually exciting point processes (which he introduced recently) by attaching auxiliary variables to the points. Count spectra and cross spectra are given.

The generalization of the (univariate) two state semi-Markov process considered by Ekholm (p. 272) allows for other than geometrically distributed runs of intervals of each type. Since the process is considered to be univariate, the type of event or interval is not observable; however this ingenious generalization of the physical mechanism of the model is intuitively appealing. Ekholm has applied the model in neurophysiological contexts. Finally Boswell and Patil (p. 285) discuss several processes which can be considered to be multivariate point processes of finite duration.

The section entitled Theory starts on page 299 with a comprehensive survey paper by Daley and Vere-Jones. This was originally a survey paper on the theory of univariate point processes but its scope is much wider, as can be seen from the detailed table of contents on pages 299 and 300. It overlaps into models, particularly in discussing some of the operations such as clustering which generates one of the most important theoretical and practical point process models. The paper by Matthes (p. 384) compliments that of Daley and Vere-Jones, going into much greater detail on results of the author and his coworkers on infinitely divisible point processes and the important sub-class of cluster point processes. Many of these results will appear in a forthcoming book by Kerstan, Matthes and Mecke (1972). Prediction theory for univariate point processes, both for the interval and counting processes, is considered by Jowett and Vere-Jones (p. 405). Results of prediction theory for generalized random functions are used, although (as the authors point out) the solutions are not entirely satisfactory. Much of this is because of the distinctive role of intervals between events in univariate point processes. Ad hoc methods are considered for certain cluster processes. To conclude this section on univariate point processes, Leadbetter (p. 436) discusses the important class of point processes consisting of times at which a stochastic process crosses a level or curve. This updates results given by Cramer and Leadbetter (1967).

The theory of multidimensional point processes is considered by Fisher (p. 468). Much of this theory developed separately from the theory of univariate point processes, even though there were points of contact, as in the early work of Bartlett. Fisher's concluding remarks (p. 505) are of great interest. He notes that there is an essential difference between multivariate and univariate point processes, this being the presence or lack of ordering, i.e. the fact that in univariate point processes intervals between events are so important. Moreover, because of the wide use of multivariate point processes in fields such as ecology, forestry, etc., he notes that there has been more development of statistical analysis techniques than of theory. This is the reverse of the situation for univariate point processes. I would add that much remains to be done in statistical theory for multidimensional point processes, especially since in most fields of application, stationarity would appear to me to be an unjustifiable artifice.

The theory of line process is considered by Krickeberg (p. 514), who calls them hyperplane processes, and in summary by Papangelou (p. 522) while Ramakrishnan (p. 533) discusses the space-time point process which were considered at an early stage in physics and attacked, very often using the idea of product densities, by Bhaba, Bartlett, and Moyal. It seems certain that there will be further developments in this area, and migration of some of the ideas current in univariate point process theory and (static) multivariate point process theory to this more difficult and complex area.

The last group of papers in this section on Theory concerns the superposition of point processes. The survey paper by Çinlar covers finite superpositions of univariate point processes (an area of great importance and very active at present) and the known and fairly complete results on convergence to Poisson limits. In Sections 6 and 7 important work on rates of convergence to the Poisson limit and on the "error" involved in approximating a finite superposition by a Poisson process is covered. The discussion is mainly in terms of counting distributions; for results using spectra and intensity functions see Cox and Lewis (1966, Chapter 8).

Generalizations of superposition results to processes with marks and multivariate point processes are given by Szász (p. 607) and Grigelionis (p. 616) respectively, while Ambartzumian (p. 626) considers superpositions of multidimensional point processes and new techniques for analyzing such superpositions using Palm distributions.

The Applications section starts with a paper by Neyman and Scott (p. 646), mentioned earlier, on cluster processes, generally multidimensional, and the applications which these processes have had in several fields. Of particular interest is the description of the authors own work, carried out over a long period of time, on the spatial distribution of galaxies and the interrelationship of their results with a developing cosmological theory. A similar application is reported by Marcus (p. 682).

Perhaps of all the fields of application to which point processes (and other stochastic processes) are central, none will be more important and as rapidly developing in the future as the spike train analyses of neurophysiology. This is the subject of two fascinating papers, one by Knight (p. 732) and the other by Stein (p. 700). Knight describes some specific and recent analyses while Stein's paper is a survey. An interesting development reported there from Stein's co-workers is a technique to allow the Fast Fourier Transform (Cooley-Tukey algorithm) to be used for rapid computation of the Bartlett (count) spectrum of a point process.

Stein's presentation featured a display on an oscilloscope of spike trains originating in the speakers hand; this unfortunately could not be presented in this volume.

Epidemiology is an important and old field concerned with point processes. There should be more contact between the work reported by Gani (p. 757) and other work in point processes. In particular, Gani discusses research attempting to analyze the interaction between the time of occurrence and place of occurrence. This problem of space-time interaction, a problem in the analysis of marked point processes, is also important (perhaps crucial), in the analysis of earthquakes. (This application of point processes was not presented at the conference, but has recently been surveyed by Vere-Jones (1970).)

Gaver (p. 775) details point process problems in reliability, a field with a lore of its own and somewhat separate from the mainstream of point process theory. Of particular interest in Gaver's paper is his descriptions of dependency structures in reliability situations and the description of models for correlated positive random variables. This has points of contact with the tests for renewal processes discussed in Lewis (p. 1).

Two important sources of spatial point processes, forestry and photographic science, are the subject of the papers by Warren (p. 802) and Hamilton, Lawton and Trabka (p. 818).

Forestry problems generated much of the statistical work reported on by Holgate (p. 122) and Warren's paper gives an ample description of these problems. Reading of both of these papers will reinforce the conclusions of Fisher (p. 468), and of Warren in his Summation (p. 811) that there is a large gap to be filled between the available theory and the needs of model builders and data analysts. In addition Warren points out the need for interpreters of the theory to practioners.

The paper by Hamilton, Lawton and Trabka is highly recommended as a source of new and important problems; of interest too are the descriptions of problems solved or partially solved by photographic scientists and well known in other fields such as queueing theory.

The last paper by Dacey (p. 869) describes another, perhaps tenuous field of application of point processes, the study of map distributions. Perhaps modesty detered Dacey from referring to some of his own contributions, some of which are referenced in the book by King (1969) on statistical geography.

REFERENCES

Bartlett, M. S. (1967). The spectral analysis of line processes. Proc. 5th Berkeley Symposium 3, 133-54.

Cox, D. R. (1955). Some statistical methods connected with series of events. J. R. Statist. Soc. B 17, 129-64.

Cox, D. R. & Lewis, P. A. W. (1972). *The Statistical Analysis of Series of Events*. London: Methuen and New York: Barnes & Noble.

Cox, D. R. & Lewis, P. A. W. (1972). Multivariate point processes. To appear in Proc. 6th Berkeley Symposium.

Cramer, H. & Leadbetter, M. R. (1967). *Stationary and Related Stochastic Processes*. New York: Wiley.

Hannan, E. J. (1970). *Multiple Time Series*. New York: Wiley.

Kerstan, J., Matthes, K. & Mecke, J. (1972). *Unbegrenzt teilbare Punktprozesse*. Berlin: Akademie-Verlag (in preparation).

King, L. J. (1969). *Statistical Analysis in Geography*. Englewood Cliffs, N.J.: Prentice-Hall.

Macchi, O. (1971). Stochastic point processes and multicoincidences. IEEE Trans. Information Theory IT-17, 2-6.

Matthes, K. (1963). Stationare zufallige Punktfolgen, I. Jber. Deutsch. Math-Verein 66, 66-79.

Vere-Jones, D. (1970). Stochastic models for earthquake occurrence. J. R. Statist. Soc. B 32, 1-62.

Wold, H. (1948). Sur les processus stationnaires ponctuels. Colloques Int. Cent. Natn. Rech. Scient. 13, 75-86.

Contents

MODELS

THEORY