Institute of Mathematical Statistics

LECTURE NOTES — MONOGRAPH SERIES

AN OVERVIEW OF THE SYMPOSIUM ON INFERENCE FOR STOCHASTIC PROCESSES

I. V. Basawa
University of Georgia

C. C. Heyde
Columbia University and
Australian National University

R. L. Taylor
University of Georgia

Abstract

Some historical snap-shots of developments in the general area of stochastic processes and statistical inference are given. An overview of the papers appearing in this volume is then presented.

1 Introduction

The Symposium on Inference for Stochastic Processes was held at the University of Georgia from May 10, 2000 to May 12, 2000. The Symposium was cosponsored by the Institute of Mathematical Statistics and was a satellite meeting of the Fifth World Congress of the Bernoulli Society. The major focus of the symposium was to provide a forum for the interchange of information on inference for stochastic processes and related applications. Partial funding by the University of Georgia's "State-of-the-Art" Conference Program, a National Security Agency Grant and a National Science Foundation Grant contributed to the success of the Symposium and is gratefully acknowledged. The Symposium attracted 79 registered participants from several countries including Australia, Belgium, Canada, Denmark, France, Germany, India, Iran, Portugal, Sweden, Taiwan, the United Kingdom and the United States of America. The program consisted of 17 sessions with 49 speakers. This volume represents selected proceedings of the Symposium.

2 Some Historical Perspectives

Before describing the current research and applications of inference for stochastic processes, it is important to consider some historical developments in this subject area. Our most influential ancestors in the subject were very strongly motivated by the scientific needs of their times. They did applied work of the highest quality for which they developed theoretical tools as necessary.

To illustrate, we might start with what is arguably some prehistory, namely with Daniel Bernoulli. He was one of the famous Swiss Bernoulli family and in 1760 he produced the first epidemic model - a deterministic model for tracking the spread of smallpox. This disease was still a major concern at that time and variolation, the first attempt at vaccination, was new and very topical. Bernoulli was a strong advocate of its advantages, for which he provided important quantitative support. Of course, the first stochastic epidemic model came much later, what is now called the chain binomial model, first published by the Russian physician Pyotr En'ko in 1889. He had measles data, and he did careful estimation and fitting.

Next let us move to France in the early 1840's. At the time there was considerable interest in the demographics of aristocratic families. Benoiston de Châteauneuf did considerable statistical work in which he reached conclusions such as, on average, an aristocratic family name dies out in around 300 years. There was interest and concern in these matters, inheritance being a key factor in the operation of society. This motivated I.-J. Bienaymé, then a civil servant, to model the extinction of family names. He developed the branching process model, and was able to state the correct form of the criticality theorem. Galton and Watson took up the same subject in England in 1873-4, but they did not know of Bienaymé's earlier work of 1845, and they did not find the correct form of the criticality theorem. Unfortunately, it is only their names that are generally remembered for the work.

From the start of the 20th Century the practical use of stochastic processes and associated inference began to flourish. In 1900 Louis Bachelier published the theory of Brownian motion, in his thesis, five years before the work of Einstein on the subject. Bachelier's motivation was the modelling of the stock market. But unfortunately his work was distinctly unappreciated till late in his life. Indeed, he did not manage to obtain a tenured university appointment until he was age 47. It was actually Kolmogorov in 1931 who first recognized the importance of his work, but his fame was established only in the 1960's and 1970's. Now we even have a Bachelier Society.

The next of our ancestors from the turn of the century who can be usefully mentioned is the Swede Filip Lundberg. To him we owe the basic theory and practice of collective risk as it is applied by insurance companies, his contributions beginning with his thesis of 1903. His starting point was the description of the total claim by what we would now describe as a compound Poisson process. He went on to spend his working life in the most senior and influential positions in the insurance industry in Sweden.

One exception in terms of practical motivation was the work of Markov and his contributions on Markov chains which date from 1906. Markov's motivation was actually to strike a blow on behalf of the St. Petersburg School (founded by Chebyshev) against Nekrasov, then leader of the rival Moscow School. Nekrasov had unwisely asserted that independence was a necessary condition for the weak law of large numbers, and Markov pounced on this, developing the idea of chain dependence to show that Nekrasov was wrong. Philosophical and religious differences underpinned a bitter enmity between them. But practical use of the Markov chain idea had significantly predated Markov. In 1846 Quetelet used a two-state Markov chain to model the weather type from one day to the next. He noted from the available data that independence did not hold, rain being more likely to be followed by rain etc.

Much of our modern queueing theory comes from the Dane Agner Erlang. He worked for the Copenhagen Telephone Company from 1908 until his death in 1929, and he was involved in all facets of queueing performance. Indeed, he is reputed to have been regularly seen walking the streets of Copenhagen accompanied by a workman with a ladder. He was hunting for network loss sources.

Another very practical man was the British engineer Harold Hurst. It is to him that we owe the ideas of long-range dependence. These ideas were developed in the 1940's and 1950's when he played a key role in the design of the Aswan High Dam on the river Nile. Hurst had abundant data, and his numerical work convinced him that standard ARMA time series models could not match the data. Much new theory, largely developed by Benoit Mandelbrot, came out of Hurst's pioneering work.

These are some of our innovative predecessors who did much to influence the development of stochastic processes and its associated inference. They had good models and genuine data. They did first-rate science and they are good role models for us. More details can be found in articles in Heyde and Seneta (2001).

The mathematical foundations for the modern theory of statistical inference were laid by R. A. Fisher in the 1920's. Neyman and Pearson developed the theory for hypotheses testing in the 1930's. Subsequently, the work by Wald, LeCam, C. R. Rao and the others unified various developments in the theory of inference. Most of the pioneering work on inference was, however, devoted to the classical framework of independent and identically distributed observations. Grenander (1950) addressed the problem of extending the classical inference theory to stochastic processes. See also Grenander (1981).

Billingsley (1961) discussed inference problems for Markov processes. Hall and Heyde (1980) gave a general treatment of likelihood based inference using martingales. The monograph by Basawa and Prakaso Rao (1980) gave a comprehensive survey of the general area of inference for stochastic processes. See also, Basawa (2001) for a recent review on this topic. The recent monograph by M. M. Rao (2000) gives a rigorous mathematical treatment of the theory of inference for stochastic processes.

3 An Overview

This volume containing the Selected Proceedings of the Symposium on Inference for Stochastic Processes includes twenty referred articles in addition to this overview. These papers are grouped into eight sections. The introductory Section 1 contains the overview article and Chris Heyde's foundational paper on shifting paradigms in inference. Section 2 has five papers on applications to various stochastic models: Barndorff-Nielsen discusses applications of Lévy processes in finance; McCormick and Seymour study extreme value results for a shot-noise model; inference problems for stochastic partial differential equations are discussed by Prakasa Rao; Roussas considers design problems under association and finally, Smith and Taylor present their results on bootstrap confidence intervals for dependent data.

Section 3 contains three papers on time series: El Bantli and Hallin discuss Kolmogorov-Smirnov tests for autoregressive models based on ranks; Bhansali and Kokoszka review long-memory parameter estimation and discuss an extension; the problem of stability of nonlinear time series is studied by Cline and Pu. Two papers on population genetics are included in Section 4: Susmita Datta studies the problem of testing neutrality using multi-generation data; Huggins, Qian and Loesch discuss inference problems for random coefficient models for haplotype effects.

Two papers on semiparametric models are included in Section 5: Greenwood and Haydon discuss semiparametric inference for synchronization of population cycles; Muller, Schick and Wefelmeyer study estimation for semiparametric stochastic processes. Section 6 contains two papers on optimal estimating functions: Durairajan and William discuss nuisance parameter elimination in optimal estimating functions; Park and Basawa study optimal estimating functions for mixed effects nonlinear models with dependent observations.

Section 7 contains three papers on spatial models: Benhenni discusses systematic sampling from a stationary spatial process; Seymour studies variance estimation for the pseudo-likelihood estimator; Vedel Jensen and Nielsen review inhomogeneous spatial Markov point processes. Finally, Section 8 contains two papers on perfect simulation: Loizeaux and McKeague discuss perfect sampling from posterior distributions in a spatial model; Møller reviews perfect simulation in stochastic geometry. Markov Chain Monte Carlo (MCMC) methods and more recently perfect simulation techniques provide a powerful link between stochastic processes and inference. The papers in Sections 7 and 8 illustrate the use of these methods.

Acknowledgements

The editors of these selected proceedings are very grateful to the many referees who carefully reviewed all papers which were submitted for publication in the proceedings. Also, the cooperation of individual authors in producing this volume in the IMS Lecture Notes Series is greatly appreciated.

Special thanks go to Connie Durden for the preparation of this Volume. Ms. Durden worked tirelessly and patiently with the various authors in securing software files of their papers, providing uniformity of margins, spacing and similar editorial changes which greatly enhanced the general appearance of this Volume.

References

Basawa, I. V. (2001). Inference in Stochastic Processes. In *Handbook of Statistics*, Vol 19, 55- 77, Eds.: D. N. Shanbhag, and C. R. Rao. Elsevier, Amsterdam

Basawa, I. V. and B. L. S. Prakasa Rao (1980). *Statistical Inference for Stochastic Processes*, Academic Press, London.

Billingsley, P. (1961). *Statistical Inference for Markov Processes.* Univ. Chicago Press, Chicago.

Grenander, U. (1950). Stochastic Processes and Statistical Inference. *Arkiv für Mat.* 1, 195-277.

Grenander, U. (1981). *Abstract Inference*, Wiley, New York.

Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and its Applications.* Academic Press, New York.

Heyde, C. C. and E. Seneta (Eds.) (2001). *Statisticians of the Centuries*, Springer, New York.

Rao, M. M. (2000). *Stochastic Processes: Inference Theory*, Kluwer, Boston.