PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Introductory Remarks

D. R. Cox and Valerie S. Isham

Phil. Trans. R. Soc. Lond. A 1990 **332**, 417-418
doi: 10.1098/rsta.1990.0123

Introductory remarks

By D. R. Cox[1] and Valerie S. Isham[2]
[1] *Nuffield College, Oxford OX1 1NF, U.K.*
[2] *Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, U.K.*

Stochastic processes are systems that evolve in time probabilistically; their study is the 'dynamics' of probability theory as contrasted with rather more traditional 'static' problems. The analysis of stochastic processes has as one of its main origins late 19th century statistical physics leading in particular to studies of random walk and brownian motion (Rayleigh 1880; Einstein 1906) and via them to the very influential paper of Chandrasekhar (1943). Other strands emerge from the work of Erlang (1909) on congestion in telephone traffic and from the investigations of the early mathematical epidemiologists and actuarial scientists. There is by now a massive general theory and a wide range of special processes arising from applications in many fields of study, including those mentioned above. A relatively small part of the above work concerns techniques for the analysis of empirical data arising from such systems.

Another strong theme, closely linked with stochastic processes, is, however, that of time series analysis where primary emphasis has been put on empirical spectral analysis, stemming from early interest of meteorologists and others in detecting 'hidden periodicities' (Schuster 1898) and on analogous time-domain methods based on serial correlations, often involving generalizations of the autoregressive processes introduced by Yule (1927). Modern interest in nonlinear time series reinforces the artificiality of drawing any rigid distinction between time series and stochastic processes.

The papers in this issue illustrate just a few of the ideas that arise in analysing empirical data from systems that have strong stochastic aspects. In particular, they illustrate one general trend in current statistical work, namely of basing statistical analysis on substantive models involving very specific subject-matter considerations. Much traditional and indeed current statistical literature emphasizes the analysis of empirical data based on probability models of very wide applicability, but having little or no specific subject-matter base. The models of analysis of variance and regression and their generalizations are usually of this type. Their fruitfulness is beyond question, but the desirability, where feasible, of introducing detailed analysis of the subject-matter content is, we hope, well illustrated by papers presented here.

It is possible to classify data from stochastic processes in many different ways, in particular as follows.

1. By the nature of the variable under analysis, discrete or continuous, scalar or vector-valued. Point processes, used to model occurrences of events in space and/or time involve binary variables but have a quite extensive literature of their own.

2. By whether the data consist essentially of one or a few long realizations, or of a large number of relatively short sections.

3. By whether the process is stationary or closely related to a stationary process, or is evolutionary, i.e. essentially non-stationary.

Phil. Trans. R. Soc. Lond. A (1990) **332**, 417–418
Printed in Great Britain

4. By whether the process occurs in time (as emphasized above) or space or both, there being an essential difference because of there normally being an absence of a special direction in space.

5. By the nature of the dependencies conceived, whether they are internal to the process (self-exciting) or involve external variables and whether they are relatively short term or long term in character.

Although the papers in this issue span a broad range of fields of application and of mathematical and statistical technique, they are a small sample of topics and very far from comprehensive. Important subjects not covered range from speech technology, with its ingenious applications of hidden Markov processes and dynamic time warping (nonlinear data-dependent transformations of the timescale to match signals), signal processing more generally, including non-stationary spectral analysis, the statistical aspects of neural nets, the special problems of large numbers of short time series, as arising for example with panel data in econometric and sociological contexts and, especially, the contrasts and links between deterministic chaos and stochastic systems (Ruelle 1990). Also the papers deal only implicitly with the technical statistical issues of the fitting of relatively complex models and the examination of their adequacy. Nevertheless we hope that the seven papers that follow give the flavour of a vigorous and wide-ranging line of work.

References

Chandrasekhar, S. 1943 Stochastic problems in physics and astronomy. *Rev. mod. Phys.* **15**, 1–89.

Einstein, A. 1906 Zur Theorie der Brownschen Bewegung. *Annln Phys.* IV **19**, 371–381.

Erlang, A. K. 1909 The theory of probabilities and telephone conversations. *Nyt Tidsskift Matematik* B **20**, 33–39.

Rayleigh, Lord (J. W. R. Strutt) 1880 On the resultant of a large number of vibrations of the same pitch and of arbitrary phase. *Phil. Mag.* 5th Ser. **10**, 73–78.

Ruelle, D. 1990 Deterministic chaos: the science and the fiction. *Proc. R. Soc. Lond.* A **427**, 241–248.

Schuster, A. 1898 On the investigation of hidden periodicities with application to a supposed 26-day period of meteorological phenomena. *Terr. Mag. atmos. Elect.* **3**, 13–41.

Yule, G. U. 1927 On a method for investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Phil. Trans. R. Soc. Lond.* A **226**, 267–298.