Statistical Science
2008, Vol. 23, No. 1, 65–68
DOI: 10.1214/07-STS246B
Main article DOI: 10.1214/07-STS246

Comment: The 2005 Neyman Lecture: Dynamic Indeterminism in Science

Hans R. Künsch

1. INTRODUCTION

Professor Brillinger is to be congratulated on this paper which is both a contribution to the history of statistics and an introduction to statistical modeling using stochastic processes, a topic that continues to be of great relevance in theory and applications. It is interesting to see how many ideas have been formulated already at an early stage. In particular, I like the idea of "synthetic data" to judge the adequacy of a fitted model. With time series or spatial data, one typically needs only a few replicates to assess visually the differences between real and synthetic data, and so this is really a powerful tool.

2. COMMENTS ON THE DATA EXAMPLES

If I understand the description of the data behind Figure 4 correctly, the rainfall has been averaged over 53 seeding days. I would expect the wind speeds to vary from day to day, so I would use a hierarchical model for the wind speeds v_j with a variance component within the same day and a variance component between days. The variation between days would lead to some variation of the time of the peak, and averaging would smear it out. Hence the sharp peak in Figure 4 is even more surprising. The only possibility I see for a model that produces a similar peak as in the actual data, is to assume a decaying intensity for the process of rain particles in Ticino.

In the two population dynamics examples, no full probabilistic model is constructed. Only the conditional mean values and not the distribution of the fluctuations are considered. Least squares methods are used for fitting. Moreover, the reproduction process is not part of the model, although the reader is referred to Guttorp (1980) for a treatment of births in the second example. From a pragmatic point, it can certainly be advantageous to focus on those parts that are of primary concern without making assumptions on other processes in the system. If only the population above a certain threshold age are of interest and if information about the number of individuals reaching the threshold age is available, then one does not need to model the births. On the other hand, as I will argue below, there is also the point of view that in order to understand a system, all relevant processes should be included.

As an interesting complement to Example 7, I would like to mention the paper Jonsen, Mills Flemming and Myers (2005) which also analyses seal movement data. They use a discrete-time integrated random walk for the animal movements, with interpolation to accommodate irregular observation times, and t-distributions for the observation errors. With such a model, they can use state-space methodology to fit the model to the data, without having to exclude suspicious observations. Including a drift component to the integrated random walk is possible, but would make the analysis more complicated.

3. DETERMINISM AND INDETERMINISM

Even 50 years after Neyman's work discussed in this paper, many fields of science are still dominated by deterministic models, at least in the area of environmental modeling where I have most experience. The reasons for this dominance are that scientists are interested in models that

- take as much knowledge about the underlying processes into account as possible,
- contribute to the understanding of these processes,
- are transferable to similar systems,
- allow prediction of the same system under different driving conditions than those observed,
- have parameters with a clear subject matter interpretation.

Some of these reservations can be made against the analysis of the weather modification experiment described in the paper. No attempt is made to connect the data with physical knowledge about atmospheric processes in the alpine region on experimental days, and even if the model gave a satisfactory fit it would not

Professor, Seminar für Statistik, Department of Mathematics, ETH Zurich, CH-8092 Zurich, Switzerland (e-mail: kuensch@stat.math.ethz.ch).