Statistical Science
1988, Vol. 3, No. 2, 147–148

In This Issue

Of the many broad areas of statistics that have undergone major development during the past 10 or 15 years, perhaps the area that has been developed most rapidly, at least in the United States, is that of statistics and the law. Statisticians are now routinely hired as consultants and expert witnesses by both sides in a wide variety of civil and criminal court cases, as well as in administrative proceedings conducted by governmental regulatory agencies. In particular, litigation involving employment discrimination has provided a rich new source of interesting methodological problems (as well as consulting income) for statisticians. In these cases, an employee or a group of employees sues their employer on the grounds that they were discriminated against with regard to their hiring, firing, job placement, promotion, salary or other working conditions, on the basis of their age, race, sex or some other characteristic that the employer is not legally allowed to consider.

It is inevitable that in this adversarial setting the statistical methods that have been used, including such standard methods as linear regression, have become controversial. Each side hires its own statistical expert and the judge or jury must evaluate the different statistical analyses presented by the experts, analyses that usually lead to contradictory conclusions. Arthur P. Dempster's article, "Employment Discrimination and Statistical Inference" is based on his Neyman Lecture for the Institute of Mathematical Statistics which was presented at the Joint Statistical Meetings in Chicago in August 1986. In this article, Dempster emphasizes that reaching conclusions about the effects of employment discrimination is a type of causal inference. Therefore, statistical estimates of these effects must be adjusted by knowledge about the employer's causal mechanisms and decision processes that comes from sources outside the usual statistical data. Dempster is critical of the econometric analyses that have been used in the past in these cases, and he develops his own model in detail. It is evident from the discussion that his views are themselves not free of controversy. Contributors to the wide-ranging discussion include economists and econometricians (Franklin M. Fisher, Arthur S. Goldberger, Gail Blattenberger and John Geweke) as well as statisticians (Harry V. Roberts, Delores A. Conway, Joseph L. Gastwirth, Paul W. Holland and Stephen E. Fienberg).

* * *

In his discussion of Nancy Reid's comprehensive review of saddlepoint methods in statistics, Robert E. Kass writes: "The world of asymptotics is beautiful and mysterious . . . Each time we consider a sample of size n to be part of an infinite sequence of observations, we are faced with an irony: limits do not depend on the first n values, yet they are able to inform us about the behavior of the sample. Our finite world seems tied to asymptopia, but how?" In a fundamental paper published in 1954, H. E. Daniels developed a strikingly accurate technique of asymptotic analysis generally called the saddlepoint approximation. Since then, particularly with the additional stimulation of a discussion paper by O. E. Barndorff-Nielsen and D. R. Cox published in 1979, saddlepoint methods have been developed for many problems of statistical inference. Reid writes that the purpose of her article is "to summarize these recent developments and show how they are all related to Daniels' original saddlepoint expansion." A valuable bonus is the annotated bibliography incorporated with the references at the end of the article. The discussants, who generally praise the usefulness and clarity of Reid's account and then add their own dimensions, include both H. E. Daniels and O. E. Barndorff-Nielsen, as well as the following other researchers of asymptopia: Phillip Hougaard, D. V. Hinkley and S. Wang, Luke Tierney, Robert E. Kass and Ann F. S. Mitchell. (Incidentally, H. E. Daniels makes another appearance in this issue in the interview with George Barnard, in connection with another activity of his that is ostensibly unrelated to his work on saddlepoint methods.)

* * *

It is an understatement to say that standard statistical methods based on linear models are widely used in applied work. These methods are taught in every elementary course in applied statistics, and they pervade virtually every area of applied research in which statistics is used. However, it is less-widely appreciated among the users of statistical methods that the standard analyses of these linear models depend on technical assumptions whose failure, when they are not appropriate for the data being analyzed, can lead to poor performance. In his review article, "Rank-based Robust Analysis of Linear Models," David Draper describes four overlapping approaches for dealing with such possible violations of the assumptions: (0) The *do-nothing* approach, in which the classical analysis is applied to the data without question.

(1) The *data-analytic* approach, in which the data are altered by setting aside outliers or transforming the observed variables, and the classicial analysis is then applied. (2) The *model expansion* approach, in which the parametric model is broadened, e.g., to the class of generalized linear models. (3) The *robust* approach, in which nonclassical methods of analysis are used that are relatively insensitive to violations of the assumptions, so the original data can be analyzed without transformation or additional modeling.

Draper goes on to state that the development of robust methods in linear models has proceeded roughly along three parallel lines: generalized maximum-likelihood type, or *M-methods*; rank-based, or *R-methods*; and methods based on linear combinations of order statistics, or *L-methods*. His article provides a careful exposition and review of R-methods. The discussants include A. H. Welsh, Roger Koenker and Stephen Portnoy, T. P. Hettmansperger and J. C. Aubuchon, Peter J. Bickel and R. Douglas Martin. They describe other promising robust approaches and defend the merits of their favorite.

* * *

A highlight of this issue is an interview with the famous British statistician, George A. Barnard. Most recently, in our interview with George Box in *Statistical Science*, Vol. 2, No. 3 (August 1987), Box described Barnard as "a wonderful person" who "had a tremendous influence" on him. Barnard is well known for his work in industrial statistics, his development of sequential analysis in Britain during World War II, and his path-breaking research on likelihood-based methods, pivotals and other aspects of the foundations of statistics. In this interview he discusses his long career and comments on his interactions with R. A. Fisher, Egon Pearson and Jerzy Neyman, among others. Barnard retired from his academic position in 1981, and now lives in Brightlingsea, Essex, where he continues his full-scale statistical activity.