Statistical Science
1997, Vol. 12, No. 3, 161

Introduction to R. A. Fisher on Inverse Probability and Likelihood

Stephen E. Fienberg

Abstract. When R. A. Fisher studied statistics as a student at Cambridge, the typical way to think about statistical inference was in terms of the method of inverse probability and Bayes' theorem. While others groped for alternatives with systematic structure and desirable alternatives, it remained for Fisher to invent the notion of likelihood and to explore its properties. These two papers trace the emergence of Fisher's thinking on likelihood over a 10-year period.

In 1908, Gosset [4] published his famous paper on the t-distribution, which most statisticians today associate with a Fisherian or modified Neyman–Pearson approach to statistics. But, the approach implicit in Gosset's paper was that of inverse probability and the posterior distribution compute via Bayes' theorem, as a quick perusal of his 1908 paper on the correlation coefficient makes clear [5]. Today we know that this result was in effect derived much earlier by Lüroth in 1876 and then again independently by Edgeworth in 1883, both of whom used inverse probability (e.g., see [3]).

This inverse probability approach represented a standard method for statistical inference, albeit often uncritically, at the time that Fisher began his study of statistics at Cambridge and, as Edwards [2] notes, it was an integral part of the subject as he learned it. However, alternatives to inverse probability already abounded, and frequentist and other non-Bayesian ideas were clearly being explored by many others and otherwise "in the air." Fisher read widely but cited selectively. Thus, even as a student, he might well have been led to read papers and books dealing with such alternatives by his teacher, J. M. F. Stratton, even though he rarely cited relevant precursors or alternatives to his own work (see Aldrich's [1] account of what Fisher studied). Fisher's invention of *likelihood* as a concept distinct from probability, therefore, had to be viewed in terms of his training in the theory of inverse probability as well as in the prevailing criticisms of it.

In the following pair of papers, John Aldrich and A. W. F. Edwards trace the evolution of Fisher's meaning of both the term *inverse probability* and the term *likelihood* over a 10-year period from 1912 to 1922, as he moved away from the inverse method and toward his own approach to inference. Because the practice of citation was perhaps less rigorous in Fisher's day than at present, we must confront the seeming ambiguity of meaning for *inverse probability* that Edwards chronicles, and the conceptual development of *likelihood* that is the primary focus of Aldrich's account. The two are, of course, intimately related, and the authors traverse the same territory exploring the same core papers by Fisher. Their accounts overlap but are complementary, and they sometimes interpret what Fisher wrote quite differently. Anyone who has worked hard at reading Fisher, however, will not be surprised!

These papers may serve as self-contained accounts of Fisher's work for some, or merely as an introduction to some of the most fascinating papers in the history of statistics in the 20th century.

Stephen E. Fienberg is Maurice Falk University Professor of Statistics and Social Science, Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213-3890.

REFERENCES

[1] ALDRICH, J. (1997). R. A. Fisher and the making of maximum likelihood. *Statist. Sci.* **12** 162–176.
[2] EDWARDS, A. W. F. (1997). What did Fisher mean by "inverse probability" in 1912–1922? *Statist. Sci.* **12** 177–184.
[3] PFANZAGL, J. and SHEYNIN, O. (1996). Studies in the history of probability and statistics XLIV: a forerunner of the t-distribution. *Biometrika* **83** 891–898.
[4] STUDENT (1908a). On the probable error of a mean. *Biometrika* **6** 1–25.
[5] STUDENT (1908b). Probable error of a correlation coefficient. *Biometrika* **6** 302–310.