Rejoinder

Sandy Zabell

1. ROBIN L. PLACKETT

Professor Plackett has made a number of distinguished contributions to the history of statistics, and I am grateful for the many interesting questions and issues that he raises in his commentary. Many of these touch on the evolution of Fisher's own thoughts about inference and provide a useful complement to my paper. I was pleased to see that several of the points Professor Plackett makes are consistent with and indeed support the thesis of my paper.

That thesis, in brief, was that: (1) After the criticisms of Boole, Venn and Chrystal, inverse probability and the Bayesian approach, although controversial, remained intellectually respectable until the 1920's. (2) During the 1920's and 1930's these methods fell into disrepute due to the efforts of Fisher and Neyman.

As Professor Plackett notes, Bayesian methods were still part of the curriculum of University College London in 1921; it was instead Rothamsted, Fisher's own institution, that first began to turn out a stream of statisticians opposed to such methods after Fisher's arrival there in 1919; and it was Fisher's textbook *Statistical Methods for Research Workers*, first published in 1925, which played a decisive part in discrediting Bayesian methods during the next quarter century.

Let me turn to a detailed comment on the issues Professor Plackett raises.

1. *The impact of Mendelian genetics on Fisher's view of statistical inference.* I think Professor Plackett makes an important observation when he suggests that genetics may have been of crucial importance in molding Fisher's view of the nature of probability. Fisher was one of those rare individuals who made major contributions to two different fields—statistics and genetics—and it is remarkable how evenly his output was divided between the two areas. Fisher cites genetics in *SMSI* as an area where objective prior probabilities are available, and the last chapter of *SMRW* uses the problem of estimating a linkage parameter to illustrate the key elements in Fisher's theory of statistical inference.

2. *Did Karl Pearson abandon uniform priors?* This is an interesting question, but I do not think it a crucial one. *Laplace* did not advocate the use of uniform priors in all instances (see Stigler, 1986, pages 135–136) and, as noted in the paper, the use of nonuniform priors was discussed by many people during the 19th century. Fisher's attack was not solely on the use of uniform priors, but on the entire Bayesian approach.

I agree with Professor Plackett that the correlation coefficient incident involving Fisher and Pearson is a crucial episode in the history of the subject. It was in large part responsible for the later break between the two, and much of Fisher's subsequent work can be viewed as an attack—implicitly or explicitly—on the Pearsonian edifice: the controversy over the degrees of freedom for chi-squared; the inefficiency of the method of moments vs. maximum likelihood; the subjectivity of Bayesian methods vs. the objectivity of fiducial probability. In his 1917 paper with Soper et al. on the correlation coefficient, a confused Pearson had criticized Fisher for using an inappropriate prior (actually Fisher hadn't used Bayesian methods at all!); it is surely no accident that in his first paper on fiducial inference, Fisher used by way of illustration the bivariate correlation coefficient, rather than the simpler univariate t-statistic.

3. *The role of "Student."* Plackett dismisses Gosset's early excursus into Bayesianism as being at a time when he was under the influence of Pearson, but in many ways that's the point: if you studied statistics at a research level in Britain shortly after the turn of the century, you studied under Pearson and were likely to emerge with a Bayesian perspective. Admittedly Gosset made little actual use of Bayesian statistics in his own work, but then again, on a *practical* level few people after Laplace did. This brings us to another important question discussed by Professor Plackett.

4. *How widely were Bayesian methods employed?* In thinking about this issue, an important distinction needs to be made. During the 19th century, Bayesian methods functioned primarily as a conceptual framework for thinking about the inferential problem, rather than as a working tool in everyday statistical practice. This was partly because most of the common statistical methods then employed could be derived from, and were often thought of as, large-sample approximations to Bayesian solutions employing flat priors.

There are, however, important exceptions to this general rule regarding the largely theoretical role of inverse probability. Poisson's use of Bayesian methods in his analysis of judicial decision-making is of course familiar; a lesser known, but equally interesting example involves the Tübingen pathologist Carl

Liebermeister. More than half a century before the advent of the "Fisher exact test," Liebermeister proposed a Bayesian approach to the problem of analyzing data resulting from clinical trials with limited numbers of patients (Zabell, 1989). This was a problem which had long vexed medical statisticians, the usual large-sample methods then available presupposing sample sizes often unattainable in practice. Liebermeister's paper on the subject is an impressive one: a solution is rigorously derived assuming independent uniform priors on the two unknown binomial proportions, several examples involving actual clinical data are given, and tables are appended to facilitate practical implementation of the method. (The posterior probability resulting from Liebermeister's solution is virtually indistinguishable from the significance probability resulting from the Fisher exact test. Unknown to Liebermeister and a generation of physicians before him, however, the method—derived for another purpose—appears in Laplace's *Théorie analytique*, pages 374–376 and 384–392; it was either ignored or overlooked because Laplace only gave a large-sample method for evaluating the resulting integral; see Todhunter, 1865, pages 418–420.)

5. *What were the relevant textbooks?* Plackett suggests that many textbooks on probability contained material on inverse methods "suitable only for the examination room," and nominates seven other candidates. In many ways, however, these are not a "representative sample." Apart from Fisher's *SMRW*, four are books on the theory of errors, a subject that developed largely independently of statistical theory for the better part of a century; and the other two are statistics texts written at so elementary a level that they do not go past the binomial and normal distributions and omit most topics of any sophistication in statistical inference (e.g., no mention is made of Pearson's chi-squared test). An important consideration to bear in mind is that the general level of statistical theory available in English prior to World War I was vastly inferior to what was available in German or French, and that in these languages mathematical treatments of statistical inference were largely to be found in textbooks on probability; a good example is the book by Czuber (1903).

6. *The role of personality in the history of science.* I concur with Plackett that one should be cautious in attempting to find interconnections between personal traits and intellectual output. Nevertheless, I am troubled by his implicit suggestion that the study of personality and personal characteristics should play only a limited role in the history of science.

In some instances, of course, such considerations are clearly extraneous. I do not think that a study of Gauss' personality is liable to give us much insight into his proofs of the law of quadratic reciprocity. But on the other hand, when one approaches a subject such as the dispute between Fisher and Neyman, I think it is clear that one cannot possibly understand it without some insight into the personalities of both combatants. By this I do not mean to suggest that Fisher or Neyman were not motivated primarily by questions of serious statistical principle; quite the contrary. But the *form* that dispute took: the level of rancor involved, the abrasive rhetoric, the inability or unwillingness to even try and understand the opposing viewpoint—this obviously *is* a matter of the personalities involved.

2. G. A. BARNARD

As one of Fisher's friends and colleagues, Professor Barnard is in a unique position to tell us about Fisher, and I read Barnard's discussion with great interest. His specific comments on my paper are largely confined to my brief synopsis of the period after Fisher entered the scene, and I welcome the opportunity to expand on what I said there.

In brief, Barnard questions whether Fisher and Neyman "simultaneously" discredited inverse probability, whether inverse probability was ever "in eclipse" and whether it enjoyed a "return" with the publication of Savage's book. Let me discuss each of these issues in turn.

First, Barnard unfortunately misreads me as saying that Fisher and Neyman simultaneously dealt Bayesian statistics a near-fatal blow *in 1930*, rightly questioning this in the case of Neyman. What I wrote was that this happened *in the 1930's*, the decade when Neyman, partly in collaboration with Egon Pearson, developed his theories of testing and estimation. Barnard is quite right that Fisher had earlier landed the first several punches (although I would date this to 1922, rather than 1912 as does Barnard), but what I had in mind when I pointed to the decade of the 1930's as the time of the "near-fatal blow" were Fisher's further attacks on inverse probability (including his papers on fiducial probability and conditional inference), Neyman's entering the fray and the impact on the statistical profession of the apparent agreement between Fisher and Neyman concerning this issue. Although Fisher's earlier attacks had already discredited Bayesian methods in certain circles (for example, Gosset and Egon Pearson), it was only in the ensuing decade that disaffection with inverse probability and Bayesian statistics became widespread. (For example, as late as 1930 one can still find Berkson commenting favorably on Bayesian methods in the first issue of *The Annals of Mathematical Statistics* (Berkson, 1930).)

It is when Professor Barnard goes on to question, however, whether inverse probability subsequently went into eclipse that I fear we part company.

Webster's *New Collegiate Dictionary* gives as one of the meanings of eclipse, "a falling into obscurity or decline: disgrace." Eclipse does not mean extinction, and the meager examples Barnard cites merely underlines how feebly the Bayesian flame burned during this period: the reprinting of Bayes' essay, the occasional conference invitation to de Finetti, a comment by Yates when discussing an RSS paper (rather than a paper by Yates himself). Jeffreys' 1939 book is an important exception, but it is the exception that proves the rule: as Jeffreys himself writes in the 2nd, 1948 edition of his book:

> Most of the present books on statistics, and of the longer papers in journals, include a careful disclaimer that the authors propose to use inverse probability, and emphasize its lack of logical foundation, which is supposed to have been repeatedly pointed out [Jeffreys, 1948, page 372].

I would certainly describe this as a state of eclipse, one which did not exist in 1929.

Since Professor Barnard does not think that Bayesian statistics was ever in eclipse, not surprisingly he questions whether it enjoyed a "return" after the appearance of Savage's book in 1954, pointing to the earlier work of Ramsey and de Finetti. In doing so, I think Barnard fails to distinguish between the appearance of a work and its later influence. Ramsey and de Finetti both made extremely important contributions to the foundations of subjective probability, but I think it is fair to say that their *influence* on the statistical profession was very limited until after 1954. Obviously Savage was not the only actor in this revival: Jack Good wrote many fine papers during this period, Raiffa and Schlaifer's book was widely influential, and Lindley's conversion was certainly important. Nevertheless, the publication of Savage's *Foundations of Statistics* marks an important watershed, at least in the United States. (Bayesian statistics might have quoted Dryden: "I was a little eclipsed, but I'll cheer up.")

As for my actual discussion of Fisher's text, Professor Barnard says that Fisher's discussion stands up well when we view it as portraying events as Fisher saw them; the question is, how accurate a portrayal is it? And the answer must remain: not very. What I think Professor Barnard really means to say is that Fisher's chapter discusses those earlier authors who were most influential on Fisher himself, and here I think he is quite right.

It remains for me to thank both Professors Plackett and Barnard for their thoughtful comments and the insights they offer about Fisher. Although my paper primarily dealt with the history of inverse probability prior to the time Fisher entered the scene, much of their discussion—not surprisingly, given the force of his personality and thought—concerns Fisher himself. I view Fisher, as I assume do most of my colleagues, as the greatest member of our profession. He transformed statistical science into a rich and powerful theory capable of handling a remarkable range of practical problems. Reading him is not easy, but it is always rewarding. His style is dense, but every sentence has been thought through and polished. Reading Fisher with critical respect was the genesis of my paper, and encouraging others to do the same was its purpose.

ADDITIONAL REFERENCES

BERKSON, J. (1930). Bayes' theorem. *Ann. Math. Statist.* **1** 42–56.

CZUBER, E. (1903). *Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik, und Lebensversicherung.* Teubner, Leipzig.

SOPER, H. E., YOUNG, A. W., CAVE, B. M., LEE, A., and PEARSON, K. (1917). A cooperative study. On the distribution of the correlation coefficient in small samples. Appendix II to the papers of 'Student' and R. A. Fisher. *Biometrika* **11** 328–413.

ZABELL, S. L. (1989). Carl Liebermeister and the statistical analysis of clinical trials. Manuscript in preparation.