FISHER, R. A. (1971–74). *Collected Papers of R. A. Fisher* **1–5** (J. H. Bennett, ed.). Univ. Adelaide.
FRY, T. C. (1928). *Probability and Its Engineering Applications*. van Nostrand, New York.
GINGERICH, O. (1973). Copernicus and Tycho. *Scientific American* **229** 86–101.
GOSSET, W. S. (1908). Probable error of a correlation coefficient. *Biometrika* **6** 302–310.
HAILPERIN, T. (1976). *Boole's Logic and Probability*. North-Holland, Amsterdam.
HARDY, G. F. (1889). Letter. *Insurance Record* 457. (Reprinted, *Trans. Faculty Actuaries* **8** 180–181, 1920.)
JEFFREYS, H. (1939). *Theory of Probability*. Clarendon Press, Oxford. (2nd ed., 1948; 3rd ed., 1967.)
JEVONS, W. S. (1877). *The Principles of Science*, 2nd ed. Macmillan, London.
KEYNES, J. M. (1921). *A Treatise on Probability*. Macmillan, London.
KOESTLER, A. (1959). *The Sleepwalkers*. Macmillan, New York.
MILL, J. S. (1843). *A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. John W. Parker, London. (Many later editions.)
MONTMORT, P. R. (1713). *Essai d'analyse sur les jeux de hazards*, 2nd ed. Jacques Quillan, Paris. (1st ed., 1708.)
NEYMAN, J. (1929). Contribution to the theory of certain test criteria. *Bull. Internat. Statist. Inst.* **24** 3–48.
NEYMAN, J. (1934). On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *J. Roy. Statist. Soc.* **97** 558–625.
NEYMAN, J. and PEARSON, E. S. (1928). On the use of interpretation of certain test criteria for purposes of statistical inference. *Biometrika* **20** 175–240, 263–294.
PASSMORE, J. (1968). *A Hundred Years of Philosophy*, 2nd ed. Penguin, New York.
PEARSON, K. (1892). *The Grammar of Science*. Walter Scott, London. (2nd ed., 1900; 3rd ed., 1911.)
PEARSON, K. (1907). On the influence of past experience on future expectation. *Philos. Mag.* (6) **13** 365–378.
PEARSON, K. (1920). The fundamental problem of practical statistics. *Biometrika* **13** 1–16.
PLACKETT, R. L. (1983). Karl Pearson and the chi-squared test. *Internat. Statist. Rev.* **51** 59–72.
PORTER, T. M. (1986). *The Rise of Statistical Thinking: 1820–1900*. Princeton Univ. Press, Princeton, N.J.
PRATT, J. W. (1976). F. Y. Edgeworth and R. A. Fisher on the efficiency of maximum likelihood estimation. *Ann. Statist.* **4** 501–514.
RAMSEY, F. P. (1926). Truth and Probability. In *The Foundations of Mathematics and Other Logical Essays* (R. B. Braithwaite, ed.) 156–198. Routledge and Kegan Paul, London (1931).
REID, C. (1982). *Neyman—From Life*. Springer, New York.
ROOT-BERNSTEIN, R. S. (1983). Mendel and methodology. *History of Science* **21** 275–295.
ROTHMAN, T. (1982). Genius and biographers: the fictionalization of Évariste Galois. *Amer. Math. Monthly* **89** 84–106.
SALMON, W. C. (1981). John Venn's *Logic of Chance*. In *Probabilistic Thinking, Thermodynamics and the Interaction of the History and Philosophy* (J. Hintikka, D. Gruender and E. Agazzi, eds.) **2** 125–138. Reidel, Dordrecht.
SAVAGE, L. J. (1976). On re-reading R. A. Fisher (with discussion). *Ann. Statist.* **3** 441–500.
SEAL, H. L. (1967). The historical development of the Gauss linear model. *Biometrika* **54** 1–24.
SHAFER, G. (1976). *A Mathematical Theory of Evidence*. Princeton Univ. Press, Princeton, N.J.
STIGLER, S. M. (1978). Francis Ysidro Edgeworth, statistician (with discussion). *J. Roy. Statist. Soc. Ser. A* **141** 287–322.
STIGLER, S. M. (1982). Thomas Bayes's Bayesian inference. *J. Roy. Statist. Soc. Ser. A* **145** 250–258.
STIGLER, S. M. (1986). *The History of Statistics: The Measurement of Uncertainty Before 1900*. Harvard Univ. Press, Cambridge, Mass.
TODHUNTER, I. (1865). *A History of the Mathematical Theory of Probability*. Macmillan, London. (Reprinted by Chelsea, New York, 1949.)
VENN, J. (1866). *The Logic of Chance*. Macmillan, London. (2nd ed., 1876; 3rd ed., 1888; reprinted by Chelsea, New York, 1962.)
VON KRIES, J. (1886). *Die Prinzipien der Wahrscheinlichkeitsrechnung. Eine Logische Untersuchung*. Freiburg. (2nd ed., Tübingen, 1927.)
VON WRIGHT, G. H. (1941). *The Logical Problem of Induction*. Finnish Literary Soc., Helsinki. (2nd rev. ed. Macmillan, New York, 1957.)
WHITTAKER, E. T. (1920). On some disputed questions of probability (with discussion). *Trans. Faculty Actuaries* **77** 163–206.
WHITWORTH, W. A. (1897). *DCC Exercises in Choice and Chance*. (Reprinted by Hafner, New York, 1965.)
WHITWORTH, W. A. (1901). *Choice and Chance*, 5th ed. George Bell and Sons, London.
WINSOR, C. P. (1947). Probability and listerism. *Human Biology* **19** 161–169.

Comment

Robin L. Plackett

Sandy Zabell deserves our thanks for discovering further details of what Boole, Venn and Chrystal wrote on the subject of inverse probability, for explaining why Fisher could not have relied on them to provide consistent arguments against this form of statistical inference and for an analysis of how far Fisher's claims concerning the eclipse of inverse probability are justified. Like everything else connected with Fisher, matters are indeed complex, and Zabell's paper provides a good topic for discussion.

At the height of his career, Fisher was certainly familiar with what mattered in developments of

Robin L. Plackett is Emeritus Professor of Statistics, University of Newcastle upon Tyne. His mailing address is: 57 Highbury, Newcastle upon Tyne NE2 3LN, United Kingdom.

the eighteenth and nineteenth centuries. Churchill Eisenhart pointed out during the 125th Annual Meeting of the American Statistical Association that Fisher was remarkably well read with respect to Laplace and Gauss, and had personally told Eisenhart that he had gotten the idea for his z transformations and the large sample approximation to the percentage points thereof from Laplace's paper of 1781. Fisher's work also contains mentions of Bessel, Helmert and Peters, slightly flawed by his constant references to Peter's formula—but these could be printers' errors. So what are the reasons for the disparity between Fisher's account of inverse probability in Chapter II of *SMSI* and the historical sketch presented by Zabell?

I believe that Fisher's account has to be judged by Fisherian standards and approached with due caution. When Zabell suggests poor eyesight as a factor to be considered, he has perhaps not seen the excellent review of Joan Fisher Box (1978) by William Kruskal (1980), where footnote 1 refers to arguments by George Stigler against facile inferences from personal traits. An important question for discussion is how far Fisher's thinking on the subject of probability was affected by the rediscovery of Mendel's laws, because these embody a concept of experimental probability that is distinguishable from the personalistic concept involved in the "equal distribution of ignorance." Fisher's chapter on the early attempts and their difficulties clearly needs to be reassessed in the light of subsequent research, but the same would apply to most historical works reconsidered 30 years later. Furthermore, *SMSI* was written when Fisher was in his mid-sixties, doubtless in his study with copies of famous books conveniently at hand on nearby shelves. He can be excused for not searching out publications less easily accessible and for any lapses of memory. His own views on statistical inference are discussed in Chapter III onwards, and Fisher may have wanted to get to these in a hurry. However, the imputation of motives is always a hazardous business.

I agree with Zabell that the history of science as written by scientists can be seriously flawed. He refers to several biographies of Galois, including presumably the one by E. T. Bell in *Men of Mathematics*. This book, while eminently readable, is particularly disappointing for statisticians, because the chapter on Gauss presents an inaccurate history of the method of least squares, and the chapter on Laplace contains only a few uninformative lines on his *Théorie Analytique des Probabilités*. Somewhat nearer to the matter in hand is Bell's chapter on Boole, which has been criticized in a paper by Geoffrey Taylor referenced in Hailperin (2nd ed., 1986).

I doubt whether the eclipse and reappearance of inverse probability took place in quite the way that Zabell has described, and for several reasons.

(1) Karl Pearson was not a consistent supporter of equiprobability, as is suggested. Zabell passes quickly over Fisher's statement in 1912 that his absolute criterion, known later as the method of maximum likelihood, is derived from the principle of inverse probability. Fisher repeated the statement in a *Draft of Note* which he sent to Pearson in 1916, by which time Pearson had decided that the choice of a *uniform* prior distribution is arbitrary. There ensued a long controversy about the supposed use by Fisher of Bayes' theorem in estimating a correlation coefficient. This episode is one of the most significant in the history of inverse probability, and doubtless contributed to the long dispute between these giants of the heroic age. Egon Pearson (1968) gave the details with his usual lucidity and restraint.

(2) The influence of Gosset has been overlooked. His remarks about the posterior distribution of a correlation coefficient were made when his association with Karl Pearson was close. They form a minor part of the paper and could not be followed up until the distribution for $R \neq 0$ had been obtained. Gosset's paper on the probable error of a mean, also published in 1908, makes statements in what Bernard Welch (1958) described as posterior language, but the main thrust of the paper is towards direct probabilities and the calculation of the first table of "Student's" z distribution. This paper profoundly influenced Fisher, and led to his brilliant development of normal sampling theory, culminating in 1928 with the general sampling distribution of the multiple correlation coefficient. Prior distributions were explicitly rejected from Fisher's work, the reaction against inverse probability was strengthened, and the mold of posterior reasoning began to break. Incidentally, Gosset is not a man to be labeled: he possessed what Fisher described in an obituary notice as one of the most original minds in contemporary science.

(3) How widely were Bayesian methods taught and employed? Egon Pearson (1974) remarked that, between 1894 and 1930, University College London was the only place in the United Kingdom where statistical theory and its applications were taught to any depth, and that graduate students from different countries came there year after year. Syllabuses for the first-year and second-year courses in the session 1921–22 show that the treatment of Bayes' theorem, criticisms of Boole, Venn and others, and the "equal distribution of ignorance" occupied less than 6 lectures out of a total of 105. The influence of Fisher at Rothamsted grew steadily after his appointment there in 1919 and was felt by a stream of "voluntary workers," including several whom Gosset recommended. Between 1920 and 1930, these were the centers that dominated the advance of statistical methodology in Britain. Bowley, Edgeworth and Keynes were all men of great distinction, but they were much less influential: Bowley was respected within the tradition of official statistics, the

genius of Keynes was applied elsewhere, and the solitary Edgeworth was at the end of his long career.

(4) Textbooks on probability by the mathematicians Whitworth, Burnside and Coolidge gave examples on the probabilities of causes suitable only for the examination room. In view of the title of *SMSI*, there are better reasons for inspecting how inverse probability was treated in textbooks on statistics, or on topics that are statistical in nature. A short list for the period between 1880 and 1930 might include the following books, detailed references for which are scarcely necessary: M. Merriman (1884); A. L. Bowley (1901); G. U. Yule (1911); D. Brunt (1917); E. T. Whittaker and G. Robinson (1924); R. A. Fisher (1925); H. L. Rietz, (1927). The choice of *Statistical Methods for Research Workers* seems appropriate. This book made a fundamental break with tradition (Yates, 1951), and successive editions tolled the death knell of inverse probability for all to hear. The mathematician Neyman seems to have been rather hard of hearing. But Harold Jeffreys firmly rejected the claim and he carried the banner of Bayes and Laplace until the next generation was ready to take over.

ADDITIONAL REFERENCES

KRUSKAL, W. (1980). The significance of Fisher: A review of *R. A. Fisher: The Life of a Scientist*. *J. Amer. Statist. Assoc.* **75** 1019–1030.

PEARSON, E. S. (1968). Some early correspondence between W. S. Gosset, R. A. Fisher and Karl Pearson, with notes and comments. *Biometrika* **55** 445–457.

PEARSON, E. S. (1974). Memories of the impact of Fisher's work in the 1920s. *Internat. Statist. Rev.* **42** 5–8.

WELCH, B. L. (1958). "Student" and small sample theory. *J. Amer. Statist. Assoc.* **53** 777–788.

YATES, F. (1951). The influence of *Statistical Methods for Research Workers* on the development of the science of statistics. *J. Amer. Statist. Assoc.* **46** 19–34.

Comment

G. A. Barnard

In drafting these comments I have had the advantage of seeing Robin Plackett's, with which I broadly agree. Matters are indeed complex.

I beg to differ from Zabell when he writes that in 1930 "Fisher and Neyman *simultaneously* (my stress, G. A. B.) administered a nearly lethal blow to Bayesian statistics, one from which it was not to recover until the publication ... of Savage's *Foundations of Statistics* in 1954." Neyman's continued interest in Bayesian methods in 1929, correctly noted by Zabell, is hardly consistent with his having shared in giving them a near lethal blow the following year. But Fisher's rejection of inverse *probability*, in the sense used here, is already quite clear in the paper of 1912 to which Zabell refers. The most important difference between 'probability' and Fisher's 'likelihood' as a measure of credibility of statistical hypotheses is that 'likelihood' does not obey the addition laws—as Fisher was wont to say, "the likelihood of H or H'" is, like "the height of Peter or Paul," meaningless unless it is specified which of the two is meant. In the final paragraph of his 1912 paper, Fisher specifically says that what he has been calling "probability" is not to be understood as capable of summation over a set of alternative hypotheses. True, he does not put forward the term 'likelihood' until 1921, but the difference of concept is already there in 1912.

Fisher clearly persuaded Egon Pearson, who in turn eventually persuaded Neyman to abandon Bayesian methods, though, unlike Pearson *fils*, Neyman never accepted likelihood as a valid measure of credibility distinct from probability. Neyman's view of the Neyman–Pearson theory had strong "decision" aspects, while Pearson's view was always more flexible.

But "eclipse" does not seem appropriate to describe the state of a theory which, through the 1930's and later continued to have the support, not only of Jeffreys, but also of such other leading users of statistics as Haldane and Gini. In 1940 Deming caused to be published a reprint of Bayes' paper of 1763, and in his introduction E. C. Molina makes it clear that Bayes' ideas continued to demand attention. Frank Yates' contribution to discussion of a paper of mine in 1946 shows Fisher's most distinguished co-worker in statistics agreeing with a Bayesian approach to problems of a certain type. When Maurice Frechet organized a discussion on statistical inference for the 1949 Paris International Congress on the History and Philosophy of Science, it was natural for him to invite

G. A. Barnard is Professor Emeritus, University of Essex. His mailing address is: Mill House, 54, Hurst Green, Brightlingsea, Colchester, Essex CO7 0EH, United Kingdom.