Statistical Science
1989, Vol. 4, No. 3, 247–263

R. A. Fisher on the History of Inverse Probability

Sandy Zabell

Abstract. R. A. Fisher's account of the decline of inverse probability methods during the latter half of the 19th century identifies Boole, Venn and Chrystal as the key figures in this change. Careful examination of these and other writings of the period, however, reveals a different and much more complex picture. Contrary to Fisher's account, inverse methods—at least in modified form—remained theoretically respectable until the 1920's, when the work of Fisher and then Neyman caused their eclipse for the next quarter century.

Key words and phrases: R. A. Fisher, inverse probability, history of statistics.

R. A. Fisher was a lifelong critic of inverse probability. In the second chapter of his last book, *Statistical Methods and Scientific Inference* (1956), Fisher traced the history of what he saw as the increasing disaffection with Bayesian methods that arose during the second half of the 19th century. Fisher's account is one of the few that covers this neglected period in the history of probability, in effect taking up where Todhunter (1865) left off, and has often been cited (see, e.g., Passmore, 1968, page 550, n. 7 and page 551, n. 15; de Finetti, 1972, page 159; Shafer, 1976, page 25). The picture portrayed is one of gradual progress, the logical lacunae and misconceptions of the inverse methods being steadily recognized and eventually discredited.

But on reflection Fisher's portrait does not appear entirely plausible. Edgeworth and Pearson, two of the most distinguished statisticians of the generation immediately prior to Fisher's, were both sympathetic to inverse methods; and indeed, as will be discussed later, Bayesian methods were widely taught and employed in England and elsewhere until the 1930's. It was only then that Fisher and Neyman simultaneously administered a nearly lethal blow to Bayesian statistics, one from which it was not to recover until the publication, nearly a quarter of a century later, of Savage's *Foundations of Statistics* in 1954.

How was such a disparity between Fisher's account and historical reality possible? Careful examination of Fisher's own evidence for his claims reveals an interesting story, telling us perhaps in some ways as much about Fisher as it does about the period he discusses.

1. FISHER'S ACCOUNT

Fisher cites three major authorities for the decline in the prestige of inverse methods: Boole, Venn and Chrystal. He had done so repeatedly in earlier papers (Fisher, 1922, pages 311 and 326; 1930, page 531; 1936a, page 248; 1951, page 49), and his account in *Statistical Methods and Scientific Inference* (*SMSI*) is an elaboration on these earlier, fragmentary comments. The following passages give the flavor of his argument:

> The first serious criticism was developed by Boole in his "Laws of Thought" in 1854. . . . Boole's criticism worked its effect only slowly. In the latter half of the nineteenth century the theory of inverse probability was rejected more decisively by Venn and by Chrystal. . . . [Fisher, 1936a, page 248]

> [Venn's criticisms of the Rule of Succession], from a writer of his weight and dignity, had an undoubted effect in shaking the confidence of mathematicians in its mathematical foundation. [*SMSI*, page 25]

> Perhaps the most important result of Venn's criticism was the departure made by Professor G. Chrystal in eliminating from his celebrated textbook of *Algebra* the whole of the traditional material usually presented under the headings of *Inverse Probability* and of the *Theory of Evidence*. [*SMSI*, page 29]

Sandy Zabell is Associate Professor of Mathematics and Statistics at Northwestern University. His mailing address is: Department of Statistics, Northwestern University, 2006 Sheridan Road, Evanston, Illinois 60208.

Fisher did not try to overstate the immediate impact of these criticisms. He noted "the slowness with which the opinions of Boole, Venn and Chrystal were appreciated," and drew attention to the defenses of inverse probability mounted by Edgeworth (1908) and Pearson (1920). Fisher was not always consistent on this point, however. Writing a few years later, he describes the supposed rejection of inverse probability in England as occurring "abruptly and dramatically" (Fisher, 1958, page 273), and uses the phrase "as late as 1908" in referring to Edgeworth's paper. Nevertheless, Fisher's earlier reference to the "decisive criticisms to which [the methods of inverse probability] had been exposed at the hands of Boole, Venn, and Chrystal" (1922, page 326), and his assertion that "[t]hese criticisms appear to be unanswerable, and the theory of inverse probability ... is now almost universally abandoned" (1951, page 49) capture the basic points of his more extended account in *SMSI*: these were the key critics, their criticisms were well-founded and they were largely responsible for the decline and fall of inverse probability.

The reader, however, who turns to Boole, Venn and Chrystal to see what they actually wrote—how accurately Fisher represents their views and to what extent they actually support Fisher's position—will find the result surprising.

2. BOOLE

Boole, Fisher says, was the first to seriously criticize "Bayes' doctrine" (Fisher, 1936a, page 249; cf. Fisher, 1951, page 49). This was only partially true. Robert Leslie Ellis had a decade earlier formulated a frequentist theory of probability (Ellis, 1844) and criticized the Laplacian approach to inference on a number of grounds including *ex nihilo nihil* (out of nothing, nothing)—i.e., no inference at all is warranted in a situation of complete ignorance. John Stuart Mill had also been, albeit briefly, a critic. In addition, both Jakob Friedrich Fries in Germany and Antoine Augustin Cournot in France had earlier discussed objective or frequentist theories of probability and attacked uncritical applications of inverse probability. (Cournot was less strident than an earlier French tradition represented by Destutt de Tracy, Poinsot and Auguste Comte; see generally Porter (1986, pages 77–88) and Stigler (1986, pages 194–200). Fisher sometimes appears to have been surprisingly unfamiliar with 19th century developments outside of England, and this often gives his historical discussions a somewhat insular flavor. Thus, he also makes no mention of Bertrand, although Bertrand's *Calcul des probabilités* (1st edition, 1889) sharply criticized inverse methods and was without question highly influential.)

Boole's criticisms were a natural outgrowth of his philosophical view that probability is a logical relation between propositions. In this he was very close to De Morgan; both De Morgan's *Formal Logic* (1847) and Boole's *Investigation of the Laws of Thought* (1854) treated probability as a branch of logic. But while De Morgan and others believed that any event possessed—at least in principle—a definite numerical probability relative to a given body of information (see, e.g., De Morgan, 1847, page 178; Donkin, 1851, pages 354–355), Boole argued that, lacking sufficient information, the probabilities of some events were indeterminate.

This was an important point, because a major defense of uniform priors in Boole's day was a challenge to doubters to produce a more plausible alternative: "A person who should dispute the propriety of dividing our belief equally amongst hypotheses about which we are equally ignorant, ought to be refuted by asking him to state which is to be preferred. He must either admit the proposed law, or maintain that there is no law at all" (Donkin, 1851, page 355). The latter is precisely what Boole did. As a result, he was able to criticize previous treatments which attempted to sidestep indeterminacy by hypothesis:

> It has been said, that the principle involved in the above and in similar applications is that of the equal distribution of our knowledge, or rather of our ignorance—the assigning to different states of things of which we know nothing, and upon the very ground that we know nothing, equal degrees of probability. I apprehend, however, that this is an arbitrary method of procedure. [Boole, 1854, page 370]

Boole supported this criticism by making the simple but telling point that in some cases the principle could be applied in more than one way to the same problem, resulting in two or more conflicting probability assignments. For example, Bayes had argued that "in the case of an event concerning the probability of which we absolutely know nothing antecedently to any trials made concerning it ... I have no reason to think that, in a certain number of trials, it should rather happen any one possible number of times than another;" i.e., that

$$P[S_n = k] = 1/(n + 1), \quad k = 0, 1, \ldots, n$$

(where S_n denotes the number of successes in n trials). But, as Boole pointed out, one could equally well argue that all *sequences* of outcomes in n trials should be viewed as equally likely, resulting in an entirely different probability assignment. (Bertrand's paradox (involving random choice of a chord) made the same point for a continuous variate (Bertrand, 1907, pages 4–5). Along the same lines, Fisher was fond of pointing out that uniform priors on continuous parameter spaces were not invariant under all continuous transformations (e.g., Fisher, 1956, page 16).)

This was an important observation, but it did not compel abandonment of the principle of indifference. It did provide a warning that naive application of the principle could lead to paradoxes and inconsistencies, and during the next century many philosophers—notably von Kries, Keynes, Jeffreys, and Carnap—undertook to refine it in an attempt to avoid them (von Kries, 1886; Keynes, 1921; Jeffreys, 1939; Carnap, 1950).

Nor did Boole himself advocate abandonment of the principle. This might not have been apparent to someone reading only *The Laws of Thought*, for there mention of the principle is indeed limited to a discussion of its improper usage. But Boole repeatedly returned to the foundations of probability in his subsequent papers, and Fisher would scarcely have found himself in agreement with Boole's later opinions.

In his last, perhaps most considered thoughts on the subject, Boole wrote that:

> All the procedure of the theory of probabilities is founded on the mental construction of the problem from some hypothesis, either, first, of events known to be independent; or secondly, of events of the connexion of which we are totally ignorant; so that upon the ground of this ignorance, we can again construct a scheme of alternatives all equally probable, and distinguished merely as favouring or not favouring the event of which the probability is sought. In doing this we are not at liberty to proceed arbitrarily. We are subject, first, to the formal *Laws of Thought*, which determine the possible conceivable combinations; secondly, to that principle, more easily conceived than explained, which has been differently expressed as the "principle of non-sufficient reason," the "principle of the equal distribution of knowledge or ignorance," and the "principle of order." We do not know that the distribution of properties in the actual urn is the same as it is conceived to be in the ideal urn of free balls, but the hypothesis that it is so involves an equal distribution of our actual knowledge, and enables us to construct the problem from ultimate hypotheses which reduce it to a calculation of combinations. [Boole, 1862, pages 389–390 of 1952 edition]

Obviously Fisher could never have accepted this view of the nature of probability, or the imprimatur it bestows upon the use of the principle of insufficient reason. (In the third edition of *SMSI*, Fisher added a subsection on Todhunter, who had emphasized that "in Bayes's own problem, we *know* that *a priori* any position of *EF* between *AB* and *CD* is equally likely; or at least we know what amount of assumption is involved in this supposition. In the applications which have been made of Bayes's theorem, and of such results as that which we have taken from Laplace in Art. 551, there has however often been no adequate ground for such knowledge or assumption" (Todhunter, 1865, pages 299–300). Fisher praised Todhunter's emphasis on the necessity for a factual rather than an axiomatic basis for prior probabilities. Nevertheless, because of Todhunter's use of the qualifying phrase "or at least we know what amount of assumption is involved in this supposition," Fisher concluded that "Near as he came to clarifying the situation, Todhunter's name cannot properly be added to those who finally succeeded in extricating the mathematical thought of the mid-nineteenth century from its bewildering difficulties." This suggests that Fisher would have been highly critical of Boole's later remarks.)

Had Boole changed his mind? He claims not, for he added in a footnote:

> . . . I take this opportunity of explaining a passage in the *Laws of Thought*, page 370, relating to certain applications of the principle. Valid objection lies not against the principle itself, but against its application through arbitrary hypotheses, coupled with the assumption that any result thus obtained is necessarily the true one. The application of the principle employed in the text and founded upon the general theorem of development in Logic, I hold to be *not* arbitrary.

The distinction that Boole intends pits the so-called "principle of *insufficient* reason," against what was later described as the "principle of *cogent* reason," i.e., that the probabilities assigned to alternatives should be taken to be equal if the information about those alternatives equally favors each (as Boole puts it, if there is "an equal distribution of our *actual* knowledge"). In any case, it is clear that Boole was not an opponent of the use of some form of the principle, and was opposed instead to what he considered its uncritical application. (As Keynes (1921, page 167) and many others have noted, Boole's writings on probability are also marred by a systematic confusion between two different meanings of independence. Hailperin (1976) provides a helpful guide through the thicket.)

3. VENN

John Venn was a Cambridge logician, best known today for his popularization of "Venn diagrams," and in his own day for his influential textbook *Symbolic Logic* (1st edition, 1881; 2nd edition, 1894). Yet in terms of originality and long-term impact, Venn's most important work is his *Logic of Chance* (1st edition, 1866), which gave the first detailed discussion in English of a frequentist theory of probability, as well as a careful critique of the earlier Laplacean position, including both its use of uniform priors and the consequences that follow from such an assumption. It was Venn's discussion of one of these consequences that Fisher examined in *SMSI*.

3.1 The Rule of Succession

Laplace's "Rule of Succession" states (in brief) that an event which has occurred n times in succession will recur the next time with probability $(n + 1)/(n + 2)$. Venn ridiculed the Rule of Succession, pointing out a variety of cases where it contradicted common sense (rain on three successive days; death caused by administered strychnine on three separate occasions; people answering a false call of fire on three different occasions). While Fisher cited Venn with general approbation, he took issue with him on this particular point. As Fisher was quick to point out, "such a rule can be based on Bayes' theorem only on certain conditions." In particular, the successive trials must be *independent*, which is certainly not the case in two of Venn's examples.

Fisher was in fact highly critical of Venn: Venn "perhaps was not aware that it [the Rule of Succession] had a mathematical basis demonstrated by Laplace;" "there is no doubt that Venn in this chapter uses arguments of a quality which he would scarcely have employed had he regarded the matter as one open to rational debate;" Venn's examples "seem to be little more than rhetorical sallies intended to overwhelm an opponent with ridicule;" and "by his eagerness to dispose of [the Rule of Succession] ... he became uncritical of the quality of the arguments he used."

In order to judge the validity and persuasiveness of Venn's treatment, in the light of Fisher's comments, it is natural to turn to Venn's original discussion, in order to read his arguments in context. The reader who turns to the reprinted edition of *The Logic of Chance*, however, will find to his surprise that although Venn does indeed devote an entire chapter to the Rule of Succession, the passages that Fisher quotes are nowhere to be found!

The solution to this puzzle, however, is not difficult. *The Logic of Chance* went through three editions—1866, 1876, and 1888, the currently available Chelsea reprint being the last of these. Although Fisher does not indicate in *SMSI* which edition he consulted, a comparison of editions reveals that Fisher was quoting from the 2nd edition, a copy of which he may have owned (this edition is cited in Fisher, 1955).

This was not a minor matter, inasmuch as Venn made substantial revisions in both the second and third editions of *The Logic of Chance*. (A comparative study of the three editions of the *Logic*, tracing the evolution of Venn's thought, would be of considerable interest. Salmon (1980) discusses some differences, but largely confines his attention to the first edition.)

In this instance, between the second and third editions Venn made major changes in the chapter on the Rule of Succession, taking out precisely the examples that Fisher so vehemently objected to. It is natural to assume that between editions a colleague or correspondent—very likely Edgeworth, whose help is acknowledged in the preface—voiced criticisms very similar to Fisher's; indeed, Venn's revision addresses precisely the points raised by Fisher: the mathematical assumptions underlying the derivation of the rule, and their possible empirical validity.

Another puzzle is the tenor of Fisher's discussion. Fisher was in a certain sense very "political" in his writings; often quick to attack the opposition, he seldom expressed in print reservations he might express to close friends and allies. That he should sharply criticize an ally like Venn seems strangely inconsistent with his usual practice. In this case, however, a simple explanation suggests itself.

Venn's criticisms were not of the inverse rule per se, but its mathematical consequence, the Rule of Succession. Thus, the examples he adduces, to the extent that they discredit the Rule of Succession, also discredit *any* form of inference that gives rise to the Rule of Succession.

And that would include fiducial inference. For in the next chapter of *SMSI*, during a discussion of the application of the fiducial argument to discontinuous data, Fisher notes that:

> An odd consequence of the analysis developed above is that the Rule of Succession derivable from the particular distribution of probability *a priori*
>
> $$\frac{dp}{\pi\sqrt{pq}},$$
>
> namely that the probability of success in the next trial is
>
> $$\frac{a + \frac{1}{2}}{a + b + 1}$$
>
> is justifiable, at least to a remarkably high approximation, in the absence of any knowledge *a priori*; and this although the corresponding complete distribution *a posteriori* is not so justifiable. [Fisher, 1956, page 68]

Thus an attack on the Rule of Succession was actually an indirect attack on the fiducial argument as well and, as such, had to be met. But Fisher was curiously coy about the matter. In his discussion of Venn, no mention is made of the fact that the Rule can be so justified, only that Venn's criticisms were specious. And when Fisher derives the Rule as an approximate consequence of the fiducial argument, no mention is made of Venn's criticisms.

There is no clear evidence whether Fisher was aware of the third edition of Venn's *Logic of Chance*. Certainly, had he seen it, he would have approved of the changes Venn made in the chapter on the Rule of Succession. But Venn made a number of other revisions as well, one of which Fisher would most certainly not have approved.

3.2 Probability and Listerism

In 1879, Dr. Donald MacAlister posed the following question in the pages of the *Educational Times*:

> Of 10 cases treated by Lister's method, 7 did well and 3 suffered from blood-poisoning: of 14 treated with ordinary dressings, 9 did well and 5 had blood poisoning; what are the odds that the success of Lister's method was due to chance?

Due to the small sizes of the samples involved, the large-sample methods then available for analyzing such differences were inapplicable, and the Bayesian solution advocated by MacAlister involved assigning independent uniform priors to the two unknown binomial proportions (see generally Winsor, 1947).

In the 3rd edition of the *Logic of Chance*, Venn included a discussion of MacAlister's question. Consistency required that Venn reject MacAlister's approach, yet Venn was obviously uncomfortable with a position that no inference could be drawn. The result was a surprising reversal. Venn describes the example as illustrating those cases which afforded "[t]he nearest approach to any practical justification for [inverse] judgments," and approves of MacAlister's treatment of it as a 'bag and balls' problem, being "the only reasonable way of treating the problem, if it is to be considered capable of numerical solution at all" (Venn, 1888, pages 186–187). Thus far Fisher might still have had no difficulty. But then Venn went on to add:

> Of course the inevitable assumption has to be made here about the equal prevalence of the different possible kinds of bag—or, as the supporters of the justice of the calculation would put it, of the obligation to assume the equal *a priori* likelihood of each kind—but I think that in this particular example the arbitrariness of the assumption is less than usual. This is because the problem discusses simply a balance between two extremely similar cases, and there is a certain set-off against each other of the objectionable assumptions on each side. Had *one* set of experiments only been proposed, and had we been asked to evaluate the probability of continued repetition of them confirming their verdict, I should have felt all the scruples I have already mentioned. But here we have got two sets of experiments carried on under almost exactly similar circumstances, and there is therefore less arbitrariness in assuming that their unknown conditions are tolerably equally prevalent.

Venn's logic is difficult to follow; the last three sentences seem more a rationalization than a carefully thought-out argument. (This is hardly surprising, since the position Venn now takes is totally incompatible with the one he had previously adopted.) What *is* clear is that Fisher would have rejected it entirely. Todhunter had been excluded from the pantheon of clarification for defending Bayes's postulate when "we know what amount of assumption is involved in this supposition." Fisher's reaction to Venn's apostasy can only be conjectured.

4. CHRYSTAL

Chrystal, Fisher says, "does not discuss the objections to this material [inverse probability and the theory of evidence]." This was only partly true. Although Chrystal did not elaborate in his *Algebra* on his reasons for omitting inverse probability, he did return to the subject 5 years later and present his objections in detail. It was easy to overlook this paper of Chrystal's, for it appeared in the *Transactions of the Actuarial Society of Edinburgh* (1891), a journal not widely available, as anyone who attempts to consult Chrystal's paper will readily find. In his 1891 paper, Chrystal spelled out his views on probability, views that Fisher would have found a serious embarrassment.

Fisher had always been at pains to emphasize that he had no objection to the use of Bayes's theorem, only to its unwarranted application in situations where information justifying the use of a prior was unavailable; in particular, Fisher objected to the principle of insufficient reason to assign priors (see, e.g., *SMSI*, page 20). Chrystal's objections, ironically, were exactly the opposite: he did not object to the use of ignorance priors, but thought that given a prior, Bayes's theorem could generate an incorrect answer! He writes:

> Perhaps the following . . . will make the absurdity of the supposed conclusion of the Inverse Rule still clearer.

> A bag contains three balls, each of which is either white or black, all possible numbers of white being equally likely. Two at once are drawn at random and prove to be white; what is the chance that all the balls are white?
>
> Any one who knows the definition of mathematical probability, and who considers this question apart from the Inverse Rule, will not hesitate for a moment to say that the chance is ½; that is to say, that the third ball is just as likely to be white as black. For there are four possible constitutions of the bag:
>
> | | 1^0 | 2^0 | 3^0 | 4^0 |
> |---|---|---|---|---|
> | W | 3 | 2 | 1 | 0 |
> | B | 0 | 1 | 2 | 3 |
>
> each of which, we are told, occurs equally often in the long run, and among those cases there are two (1^0 and 2^0) in which there are two white balls, and among these the case in which there are three white occurs in the long-run just as often as the case in which there are only two.

Chrystal then goes on to correctly calculate that, in contrast, the "application of the Inverse Rules" leads to posterior odds of 3 to 1 in favor of the third ball being white, and concludes:

> No one would say that if you simply put two white balls into a bag containing one of unknown colour, equally likely to be black or white, that this action raised the odds that the unknown ball is white from even to 3 to 1. It appears, however, from the Inverse Rule that if we find out that the two white balls are in the bag, not by putting them in, but by taking them out, it makes all the difference.

Indeed it does. Chrystal's error is exactly the point of the closely related *Bertrand box paradox* (Bertrand, 1907, pages 2–3).

In the light of this fundamental misunderstanding, Chrystal's objections to inverse probability can scarcely be described as intellectually devastating. He was merely one of many (e.g., D'Alembert and Mill) whose intellectual attainments in other areas led him to uncritically accept his own untutored probabilistic intuitions. As Jevons once noted, "It is curious how often the most acute and powerful intellects have gone astray in the calculation of probabilities" (Jevons, 1877, page 213). (In 1893, shortly after Chrystal read his paper before the Actuarial Society of Edinburgh, John Govan read a paper before the same body, pointing out the errors and confusions in Chrystal's paper. It went unpublished, however, until 1920, when the eminent mathematician E. T. Whittaker read a similar exposé before the London Faculty of Actuaries (Whittaker, 1920).)

The conclusion to this episode in the history of the history of statistics is somewhat bizarre. Of his trinity of authorities—Boole, Venn and Chrystal—Fisher thought Boole was an opponent of inverse methods, but Boole was not; Venn was an opponent, but only in part; and Chrystal was an unqualified opponent, but on grounds Fisher would have found repugnant, had he known of them.

5. INVERSE PROBABILITY FROM 1880 TO 1930

What was the actual impact of these critics? Contrary to what Fisher suggests, they did not eliminate inverse methods. Edgeworth and Pearson, perhaps the two most prominent English statisticians of the generation immediately preceding Fisher's, both remained sympathetic to Bayesian methods. Moreover, we have the testimony of Fisher himself that he had "learned it at school as an integral part of the subject, and for some years saw no reason to question its validity" (Fisher, 1936a, page 248). Indeed, he had to "plead guilty in my original statement of the Method of Maximum Likelihood [Fisher, 1912] to having based my argument upon the principle of inverse probability . . ." (Fisher, 1922, page 326).

The real effect of Boole, Venn, and Chrystal and other critics appears rather to have been to cause the exponents of inverse methods to hedge their claims for the theory. For example, William Allen Whitworth, the author of a popular 19th century textbook *Choice and Chance*, dealt with objections to the rule of succession by conceding that expressions such as "entirely unknown" in its formulation were "vague." He proposed that they be replaced in the rule by the explicit hypothesis that "all possible probabilities [are] equally likely," and noted that:

> Though the cases are very rare in which the radical assumption of the Rule of Succession is strictly justified, the rule may be taken to afford a rough and ready estimate in many cases in which the assumption is approximately justified. [Whitworth, 1901, page 193]

This defense essentially originates with Edgeworth, who was an important defender of inverse methods throughout this period (see Stigler, 1978, page 296; 1986, page 310). In 1884, at the beginning of his career, Edgeworth wrote a review of Venn's *Logic*, entitled "The Philosophy of Chance," which appeared in the English philosophical journal *Mind*. (Nearly 40 years later, in the twilight of his career, Edgeworth would return to the same subject with an article of the same title in the same journal, this time reviewing Keynes's *Treatise*.) Edgeworth took an empirical and pragmatic view of the subject, and, as noted earlier, may well have been responsible for many of the changes Venn made in the third edition of *The Logic of Chance*.

The defenses mounted by Edgeworth and others fell into three broad categories. They were: (1) *The Bayes-Laplace postulate of equiprobability corresponds, at least approximately, to experience.* Karl Pearson found this argument particularly persuasive, and adopted it in his influential *Grammar of Science* (1st edition, 1892) and later articles (Pearson, 1907; Pearson, 1920, page 4). (2) *Other priors.* Another move was to concede that experience might indeed point to other priors. Both the actuary G. F. Hardy (1889) and the mathematician Whitworth (1897, pages 224–225) proposed the class of beta priors as suitable for this purpose. Others, such as Gosset (1908) and Bachelier (1912), suggested the use of polynomial priors. (3) *The suppression of a priori probabilities* (Edgeworth, 1922, page 264). A third and final defense was that when large samples were involved the particular prior employed did not matter. This had been noted as early as 1843 by both Cournot (1843, Section 95, page 170) and Mill (1843, Book 3, Chapter 18, Section 6), and had been extended by Edgeworth to parameters other than binomial proportions (Edgeworth, 1884b, page 204). A related development was Poincaré's method of arbitrary functions; see, e.g., Borel (1965, Chapter 9).

These were creditable arguments and, given the *imprimatur* of Edgeworth and Pearson, it is not surprising to find acceptance of prior probabilities at least initially even among statisticians of Fisher's own generation. Gosset's ["Student"] discussion of the issue in his classic 1908 paper on the "Probable error of a correlation coefficient" is a good example. Gosset describes the estimation problem for the correlation coefficient as that of determining "the probability that R [the population correlation coefficient] for the population from which the sample is drawn shall lie between any given limits" (Gosset, 1908, page 302). He then adds:

> It is clear that in order to solve this problem we must know two things: (1) the distribution of values of r [the sample correlation coefficient] derived from samples of a population which has a given R, and (2) the *a priori* probability that R for the population lies between any given limits. Now (2) can hardly ever be known, so that some arbitrary assumption must in general be made ... I may suggest two more or less obvious distributions. The first is that any value is equally likely between $+1$ and -1, and the second that the probability that x is the value is proportional to $1 - x^2$: this I think is more in accordance with ordinary experience: the distribution of *a priori* probability would then be expressed by the equation $y = (3/4)(1 - x^2)$.

Gosset's discussion clearly reflects a change in climate; "some arbitrary assumption must in general be made;" and a nonuniform prior seems "more in accordance with ordinary experience." Nevertheless, his basic view of estimation is clearly Bayesian. Nor were the references to prior probabilities in the statistical literature of this period mere lip-service: Edgeworth's important 1908 papers on maximum likelihood were based in part on them, and Neyman himself later employed prior probabilities in some of his earlier papers (Neyman and Pearson, 1928; Neyman, 1929). (Neyman had originally hoped to have Pearson's name appear as a co-author on the second paper, but by this time Pearson was unwilling to have his name associated in print with prior probabilities (Reid, 1982, pages 82–85).)

Acceptance of inverse methods continued into the 1920's, when they received a powerful assist from the work of Frank Ramsey (1926). Indeed, Fisher would appear to be the *first* British statistician of any standing to publicly attack Bayesian methods. The remarkably hostile reaction to his 1935 *JRSS* discussion paper (Fisher, 1935) may reflect in large part the antagonism of the Bayesian old-guard to the *nouvelle statistique*. Writing as late as 1934, Neyman could state that "until recently" it had been assumed that the problem of statistical estimation in sampling from a population required "knowledge of probabilities a priori" (Neyman, 1934).

Nearly half a century elapsed between the appearance of the first edition of Chrystal's *Algebra* (1886) and Fisher's attacks on inverse probability. During that period inverse methods were debated, claims for the theory qualified, and caution in its use advised, but the theory itself was never totally abandoned, and there is no evidence whatever for what Fisher described on one occasion as an abrupt and dramatic change. Textbooks continued to cover the subject (e.g., Coolidge, 1925; Burnside, 1928; Fry, 1928), questions on it continued to appear on actuarial examinations (A. Fisher, 1915, page 56), respected statisticians continued to employ it (Bowley, 1926). Fisher suggests that the most important result of Venn's criticism had been Chrystal's omission of inverse probability from his *Algebra*. Surely more to the point is that virtually every textbook in probability written in English during the period 1886–1930 *includes* the topic, as well as most texts in French and German. Indeed, it is difficult to find exceptions—apart from Bertrand—at least among texts of the first rank. Writing in 1921, Keynes could state that "the reaction against the traditional teaching during the past hundred years has not possessed sufficient force to displace the established doctrine, and the Principle of Indifference is still very widely accepted in an unqualified form" (Keynes, 1921, page 84).

Fisher was, in fact, being too modest when he ascribed the demise of inverse probability to Boole, Venn and Chrystal. The two most important persons in that undertaking were none other than Fisher himself and Neyman. (Thus for Egon Pearson, the inverse probability approach "had been forever discredited by Fisher in his 1922 paper ..." (Reid, 1982, page 79).) Human nature being what it is, no matter how cogent or convincing the arguments of the opponents of inverse probability were, until a credible alternative to the Bayesian methodology was provided, any attempt to demolish the edifice of inverse probability was doomed to failure (see, e.g., Pearson, 1920, page 3).

The Harvard mathematician Julian Lowell Coolidge was perhaps merely being more candid than most when he wrote (1925, page 100):

> Why not, then, reject the formula outright? Because, defective as it is, Bayes' formula is the only thing we have to answer certain important questions which do arise in the calculus of probability.... Therefore we use Bayes' formula with a sigh, as the only thing available under the circumstances:
>
> 'Steyning tuk him for the reason the thief tuk the hot stove—bekaze there was nothing else that season.' [Kipling, *Captains Courageous*, Chapter 6]

6. DISCUSSION

Paradoxically, the history of science when written by scientists themselves is sometimes seriously flawed. A typology of possible reasons for this suggests two general categories, involving sins of omission and sins of commission.

First and foremost, there may be simply a lack of interest, resources, time or training. A common manifestation of this is the uncritical copying of earlier, secondary, often highly flawed accounts without consulting original sources. Everyone "knows," for example, that during the Middle Ages the Ptolemaic model of the solar system was modified by the addition of epicycle upon epicycle to artificially force agreement with increasingly accurate experimental data. But in reality, nothing of the kind occurred: the original Ptolemaic model of one deferent and one epicycle provided a remarkably good fit to the observational data available prior to the time of Tycho Brahe; indeed, given the mathematical sophistication of Ptolemy's original system, more simplified models were typically employed throughout the Middle Ages, not more complex ones (see, e.g., Gingerich, 1973, page 95). But this misconception fits popular prejudices about the science of the Middle Ages (see, e.g., Arthur Koestler's *The Sleepwalkers*, 1959) and so is repeated from one misinformed source to another. It does not occur to someone to check the authenticity of such a story, any more than it would occur to him to check whether Einstein was responsible for the special theory of relativity, or whether Watson and Crick discovered the structure of DNA.

Even when a person has first-hand knowledge of the events about which he is writing, the passage of time may lead to a subtle erosion in the accuracy with which those events are remembered. A notable example is Karl Pearson's historical account of correlation (Seal, 1967; Plackett, 1983). As Stigler notes, Pearson's commentary "reflects well neither upon Pearson nor the general trustworthiness of the latter recollections of great scientists" (Stigler, 1986, page 344, n. 11).

Under the rubric of sins of commission may be placed an interrelated complex of causes including subconscious bias, dogmatism, sensationalism and deliberate distortion. Everyone "knows," for example, that the night before he was fatally wounded in a duel, the unfortunate Évariste Galois stayed up feverishly writing down a sketch of his theory of equations so that it would not be lost to posterity. In reality Galois had published an outline of his results months earlier, and although he did write further details down the night before the fatal duel, there was not the urgency often depicted. Reality does not make nearly as good a story as the piquant version in circulation. As Rothman (1982) discusses, this is not an isolated incident in Galois's biography: several of the best known accounts of Galois's life (those of Bell, Hoyle and Infeld) are marred by serious inaccuracies which occur because of—rather than in spite of—the ability of their authors to appreciate the technical achievements of Galois; "the misfortune is that the biographers have been scientists" (Rothman, 1982, page 104). Similarly, Stigler (1982) argues that many accounts of Bayes's original paper are seriously inaccurate; here foundational biases often led statisticians of the stature of Pearson, Fisher and Jeffreys to misread into Bayes their own viewpoints.

Fisher's account of the history of inverse probability is marred for reasons falling into both of these general categories. Due perhaps in part to poor eyesight, Fisher was never very scholarly in documenting previous work; this was to prove vexatious years later when Neyman and others would criticize him for not adequately acknowledging Edgeworth's earlier contributions to maximum likelihood (Savage, 1976, pages 447–448; Pratt, 1976).

Nevertheless, throughout his life Fisher had a serious interest in historical matters. Leafing through Todhunter, he was quick to note the Bernoulli-

Montmort correspondence about the optimal strategy in the game of "le Her," and realized (a decade before the work of von Neumann and Morgenstern on game theory) that a randomized strategy was appropriate (Fisher, 1934). (On the other hand, had Fisher referred to Montmort's book he would have discovered an extract of a letter from Waldegrave to Montmort discussing the possibility of randomized strategies! (Montmort, 1713, pages 409–412).) He was often fond of using an historical data set as the perfect pedagogical foil; the entire third chapter of Fisher's *Design of Experiments*, for example, is centered about an analysis of Darwin's data on cross and self-fertilized plants. Occasionally, the result might even suggest a radical historical reassessment, as in his article on whether Mendel fudged his data (Fisher, 1936a; Root-Bernstein, 1983).

And what Fisher was acquainted with, he often knew very well indeed. As Savage (1976, page 447) notes, Fisher "was well read in the statistical literature of his past," and Fisher's writings display a detailed knowledge of Bayes, Boole, Venn, Todhunter and Keynes. But it is a common failing to read into the words of the past the thoughts of the present, and to view the evolution of history as the progressive triumph of one's own viewpoint. This Fisher appears to have done.

ACKNOWLEDGMENTS

The author expresses his thanks to Persi Diaconis and Paul Meier for a number of helpful comments and suggestions during the preparation of the paper, to Elisabeth Vodola for supplying a copy of Chrystal's 1891 paper, and to an anonymous referee for a careful reading of the manuscript.

REFERENCES

BACHELIER, L. (1912). *Calcul des probabilités* **1**. Gauthier-Villars, Paris.

BERTRAND, J. (1889). *Calcul des probabilités*. Gauthier-Villars, Paris. (2nd ed., 1907, reprinted by Chelsea, New York.)

BOOLE, G. (1854). *An Investigation of the Laws of Thought*. Walton and Maberly, London. (Reprinted by Dover, New York, 1976.)

BOOLE, G. (1862). On the theory of probabilities. *Philos. Trans. Roy. Soc. London* **152** 225–252. (Reprinted in G. BOOLE, *Collected Logical Works* **1**. *Studies in Logic and Probability* (R. Rhees, ed.) 386–424. Open Court Publishing Co., La Salle, Ill., 1952.)

BOREL, É. (1965). *Elements of the Theory of Probability*. Prentice-Hall, Englewood Cliffs, N.J.

BOWLEY, A. L. (1926). Measurement of the precision of index-numbers attained in sampling. *Bull. Internat. Statist. Inst.* **22** 6–62.

BOX, J. F. (1978). *R. A. Fisher: The Life of a Scientist*. Wiley, New York.

BURNSIDE, W. (1928). *The Theory of Probability*. Cambridge Univ. Press, New York. (Reprinted by Dover, New York, 1959.)

CARNAP, R. (1950). *Logical Foundations of Probability*. Univ. Chicago Press, Chicago. (2nd ed., 1962).

CHRYSTAL, G. (1886). *Algebra*. Adam and Charles Black, London.

CHRYSTAL, G. (1891). On some fundamental principles in the theory of probability. *Trans. Actuarial Soc. Edinburgh* (*N. S.*) **2** 421–439.

COOLIDGE, J. L. (1925). *An Introduction to Mathematical Probability*. Oxford Univ. Press. (Reprinted by Dover, New York, 1962.)

COURNOT, A. A. (1843). Exposition de la théorie des chances et des probabilités. Librairie de L. Hachette, Paris.

DE FINETTI, B. (1972). *Probability, Induction, and Statistics: The Art of Guessing*. Wiley, New York.

DE MORGAN, A. (1847). *Formal Logic: or, the Calculus of Inference, Necessary and Probable*. Taylor and Walton, London. (Reprinted by The Open Court Co., London, 1926.)

DONKIN, W. F. (1851). On certain questions relating to the theory of probabilities. *Philos. Mag.* (*4*) **1** 353–368, 458–466.

EDGEWORTH, F. Y. (1884a). The philosophy of chance. *Mind* **9** 222–235.

EDGEWORTH, F. Y. (1884b). *A priori* probabilities. *Philos. Mag.* (*5*) **18** 204–210.

EDGEWORTH, F. Y. (1908). On the probable errors of frequency constants. *J. Roy. Statist. Soc.* **71** 381–397, 499–512, 651–678. Addendum **72** (1909), 81–90.

EDGEWORTH, F. Y. (1922). The philosophy of chance. *Mind* **31** 257–283.

ELLIS, R. L. (1844). On the foundations of the theory of probabilities. *Trans. Cambridge Philos. Soc.* **8** 1–6. (Reprinted in *The Mathematical and Other Writings of Robert Leslie Ellis M. A.* (W. Walton, ed.). Deighton and Bell, Cambridge, 1863.)

FIENBERG, S. E. and HINKLEY, D. V. (eds.) (1980). *R. A. Fisher: An Appreciation. Lecture Notes in Statist* **1**. Springer, New York.

FISHER, A. (1915). *The Mathematical Theory of Probabilities and its Application to Frequency Curves and Statistical Methods* **1**. *Mathematical Probabilities and Homograde Statistics*, 2nd ed. Macmillan, New York, 1923.

FISHER, R. A. (1912). On an absolute criterion for fitting frequency curves. *Messenger Math.* **41** 155–160. (*Collected Papers* 1.)

FISHER, R. A. (1921). On the "probable error" or a coefficient of correlation deduced from a small sample. *Metron* **1** 3–32. (*Collected Papers* 14; contains Fisher's first critical comment on inverse probability.)

FISHER, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philos. Trans. Roy. Soc. London Ser. A* **222** 309–368. (*Collected Papers* 18.)

FISHER, R. A. (1930). Inverse probability. *Proc. Cambridge Philos. Soc.* **26** 528–535. (*Collected Papers* 84.)

FISHER, R. A. (1934). Randomisation, and an old enigma of card play. *Math. Gaz.* **18** 294–297. (*Collected Papers* 111.)

FISHER, R. A. (1935). The logic of inductive inference. *J. Roy. Statist. Soc.* **98** 39–54. (*Collected Papers* 124.)

FISHER, R. A. (1936a). Uncertain inference. *Proc. Amer. Acad. Arts Sci.* **71** 245–258. (*Collected Papers* 137.)

FISHER, R. A. (1936b). Has Mendel's work been rediscovered? *Ann. Science* **1** 115–137. (*Collected Papers* 144.)

FISHER, R. A. (1951). Statistics. In *Scientific Thought in the Twentieth Century* (A. E. Heath, ed.) 31–55. Watts, London. (*Collected Papers* 242.)

FISHER, R. A. (1955). Statistical methods and scientific induction. *J. Roy. Statist. Soc. Ser. B* **17** 69–78. (*Collected Papers* 261.)

FISHER, R. A. (1956). *Statistical Methods and Scientific Inference*. Hafner, New York. (2nd ed., 1959; 3rd ed., 1973; page references are to the 3rd ed.)

FISHER, R. A. (1958). The nature of probability. *Centennial Review* **2** 261–274. (*Collected Papers* 272.)

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.