

Article

Fisher, Bayes, and Predictive Inference [†]

Sandy Zabell

Department of Mathematics, Northwestern University, Evanston, IL 60208, USA; zabell@math.northwestern.edu

[†] Dedicato a Eugenio Regazzini, con ammirazione, affetto, e stima.

Abstract: We review historically the position of Sir R.A. Fisher towards Bayesian inference and, particularly, the classical Bayes–Laplace paradigm. We focus on his Fiducial Argument.

Keywords: Bayesian inference; Fisher fiducial argument; inverse probability; uniform distribution

MSC: 62F15

1. Introduction

R.A. Fisher was—famously—an outspoken critic of Bayesian inference. Yet, anyone who reads Fisher’s later work cannot but be struck by Fisher’s evident admiration for Bayes himself. This paper discusses this apparent paradox. There are really three threads to this story. One is to recognize that Fisher was a critic of “inverse probability” (in the narrow sense of the use of uniform priors to capture the notion of ignorance or lack of knowledge about a prior) and not “Bayesian inference” per se (if by the latter, we mean the use of prior distributions summarizing genuine knowledge). Indeed, Fisher appears to have introduced the use of the term “Bayesian”; see [1]. A second thread in our story is Fisher’s discovery of fiducial inference in 1930 as an alternative to the Bayesian approach. Never widely accepted in the statistical community (unlike virtually all of Fisher’s other contributions to statistics), its ultimately inference-based approach meant Fisher came to associate himself more with statisticians such as his fellow Cambridge don Harold Jeffreys (an “objective Bayesian”) rather than his arch-enemy and rival, the frequentist Jerzy Neyman. The final thread is how all this played itself out in the 1950s. Sensing then that he was on the wrong side of history, during the last decade of his life, Fisher made a supreme effort to clarify and recast the logical basis of his approach to statistical inference, most notably in his 1956 book *Statistical Methods and Scientific Inference* [2–5]. This in turn led him to return to Bayes himself and his famous paper [6], for which Fisher had only praise.

2. Fisher on Inverse Probability in the 1920s

With two exceptions [7,8], Fisher’s papers on statistics all date from 1920 on. By this point, he was already a dedicated opponent of the use of inverse probability, writing that Bayes’ attempt «admittedly depended upon an arbitrary assumption, so that the whole method has been widely discredited» [9] (p. 4), and a year later was even more forceful, pointing with disdain to:

«inverse probability, which like an impenetrable jungle arrests progress towards precision of statistical concepts.» [10] (p. 311)

However, it was not always so; this position represented a change in view, as Fisher later acknowledged:

«I may myself say that I learned [inverse probability] at school as an integral part of the subject, and for some years saw no reason to question its validity.» [11] (p. 248)

Indeed, Fisher conceded that not only had he accepted the legitimacy of inverse probability, he had even employed it in his very first paper:



Citation: Zabell, S. Fisher, Bayes, and Predictive Inference. *Mathematics* **2022**, *10*, 1634. <https://doi.org/10.3390/math10101634>

Academic Editors: Emanuele Dolera and Federico Bassetti

Received: 4 March 2022

Accepted: 4 May 2022

Published: 11 May 2022

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

«I must indeed plead guilty in my original statement of the Method of Maximum Likelihood [in 1912] to having based my argument upon the principle of inverse probability; in the same paper, it is true, I emphasized the fact that such inverse probabilities were relative only.» [12] (p. 326)

There has indeed been some discussion and debate about the extent to which Fisher's 1912 paper did employ "the principle of inverse probability" (see [13,14]), but for our purposes, the key point is Fisher's unambiguous statement that, initially, he "saw no reason to question its validity".

2.1. Fisher's Critique of the Uniform Prior

By inverse probability, Fisher really meant the classical assumption and use of uniform priors (rather than Bayesian inference in the more modern sense of the use of a prior summarizing initial information). Fisher had several grounds on which he criticized the use of such uniform priors; one of these was that uniform priors are not scale invariant. For example, if $0 \leq p \leq 1$ is the probability of a success in a sequence of Bernoulli trials, and:

$$\sin \theta = 2p - 1,$$

then the flat prior dp on the p scale corresponds to the prior:

$$\frac{\cos \theta}{2} d\theta$$

on the θ scale (see [10] (p. 325) and [4] (pp. 16–17)). Thus, this formulation of ignorance is scale dependent, even though, in a hypothetical state of "complete ignorance" (whatever that is), there should be no reason to prefer one scale over the other; instead of the parameter p , we "might, so far as cogent evidence is concerned, equally have taken any monotonic function of p " in its place.

Interestingly, Fisher was not the first to raise this objection; much earlier, the German physiologist Johannes von Kries (1853–1928), in his *Principien der Wahrscheinlichkeitsrechnung* (see [15] (p. 31) and also [16]), gave the scientifically more interesting example of

$$\sigma: \text{specific weight}, \quad \omega = \frac{1}{\sigma}: \text{specific volume},$$

noting that a uniform prior for one is not uniform for the other, and thereby concluding "Der Wahrscheinlichkeits-Ansatz ist also unbestimmt" (the probability assumption is thus indeterminate). Nor was this an isolated, special instance: other examples could easily be given, such as in the case of a pendulum, the length and the duration or frequency of oscillation.

2.2. Fisher vs. Pearson on the Correlation Coefficient

Although this objection (and other critiques given in von Kries and later reported by Keynes in his classic 1921 *Treatise on Probability*, see [17]) were certainly cogent, one might wonder at the vehemence with which the young Fisher pressed. However, here, there is a simple explanation, stemming from Fisher's conflicted relations with Karl Pearson (1857–1936), then the towering figure in English statistics, the only person at that point to hold a chair in statistics in England (at University College London).

In 1915, Fisher published his first great paper on the distribution of the sample correlation coefficient, in Pearson's journal *Biometrika*. In the last section of his paper (see [8] (pp. 520–521)), Fisher derived a point estimate for the theoretical correlation coefficient ρ using his earlier method of maximum likelihood (see [7]).

The MLE can be interpreted as the mode of the posterior distribution of a parameter starting from a uniform prior, although Fisher did not frame the issue this way in his 1912 paper. Two years later, however, a "Cooperative Study" in *Biometrika* with Pearson as a co-author (see [18]) interpreted Fisher as doing precisely this, criticizing his ostensible use

of a uniform prior in this case as contrary to common experience, and going on to discuss several alternative priors (pp. 352–360). Angered, Fisher wrote in response:

«[Pearson’s] comments upon my methods imply such a serious misunderstanding of my meaning that a brief reply is necessary. . .

From this passage a reader, who did not refer my paper, which had appeared in the previous year, and to which the *Cooperative Study* was called an “Appendix”, might imagine that I had used Boole’s ironical phrase “equal distribution of ignorance”, and that I had appealed to “Bayes’ theorem”. I must therefore state that I did neither.»

These words did not appear in Pearson’s journal *Biometrika*, for Pearson had refused to publish Fisher’s response, and marked a clear and acrimonious break in relations between the two. It is easy to appreciate Fisher’s bitterness given his relative youth and professional vulnerability at this early stage in his career. It is nevertheless surprising to see how much this still rankled twenty years after Pearson’s death, when Fisher wrote in his book *Statistical Methods and Scientific Inference* ([2] (pp. 2–3), cited below as SMSI, page references to the third, 1973 edition unless otherwise noted):

«Pearson’s energy was unbounded. In the course of his long life he gained the devoted service of a number of able assistants, some of whom he did not treat particularly well. He was prolific in magnificent, or grandiose, schemes capable of realization perhaps by an army of industrious robots responsive to a magic wand.

The terrible weakness of his mathematical and scientific work flowed from his incapacity in self-criticism, and his unwillingness to admit the possibility that he had anything to learn from others, even in biology, of which he knew very little. His mathematics, consequently, though always vigorous, were usually clumsy, and often misleading. In controversy, to which he was much addicted, he constantly showed himself to be without a sense of justice.»

Another revealing episode was an ill-advised invitation to Fisher to write an entry on Pearson for the *Dictionary of National Biography* shortly after Pearson’s death in 1936; see [19]. After several drafts in which the editors were unable to persuade Fisher to tone down his evident contempt for Pearson, the project was abandoned by mutual agreement.

3. The Fiducial Argument

Fisher’s harsh criticism of inverse probability was ultimately grounded in principle, even if it was in part sharpened by his animus towards Pearson. However, it is difficult to replace something by nothing, and as long as a credible alternative to inverse methods was missing, inverse probability could be expected to survive. As the Harvard mathematician Julian Lowell Coolidge noted, «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. . . we use Bayes’ formula with a sigh, as the only thing available under the circumstances» (see [20] p. 100). However, in 1930, Fisher believed he had in fact found such an alternative.

3.1. The Original Fiducial Argument

In 1930, Fisher wrote a short paper titled “Inverse probability” [21], which was the origin of his “fiducial argument”, ultimately leading to more than three decades of controversy. After revisiting his oft-stated attack on the inverse probability approach, Fisher wrote:

«There are, however, certain cases in which statements in terms of probability can be made with respect to the parameters of the population.»

How did Fisher succeed in this, breaking the “Bayesian omelet without breaking the Bayesian eggs”? From a modern perspective, what Fisher discovered was a general method of constructing confidence intervals in the case of a continuous one-parameter family of probability distributions. Curiously and remarkably, there is (almost) nothing controversial

in this paper. Indeed, according to Charles Stein (who was in no way a Fisherian), there was only a single objectionable passage in the entire paper (personal communication from Charles Stein to SZ, c. 1980; this would have been the reference on p. 534 to “the fiducial distribution of a parameter θ for a given statistic T ”).

Here is a brief *précise* of Fisher’s discovery recast in modern language. Let:

- T be a sample statistic having a continuous distribution;
- θ be a one-dimensional parameter for the distribution of T ;
- $t_p(\theta)$ be the p -th quantile for T under θ :

$$P_\theta[T < t_p(\theta)] = p.$$

If t_p is strictly increasing in θ , then t_p^{-1} is also strictly increasing, and

$$P_\theta[t_p^{-1}(T) < \theta] = p.$$

Therefore, $(t_p^{-1}(T), \infty)$ is a (lower) $100p\%$ confidence interval for θ .

Perhaps Fisher’s clearest statement of what he had in mind here was a letter from Fisher to Tukey, dated 27 April 1955 ([22] pp. 220–222):

«The probability integral of the exact frequency distribution, in finite samples, of an exhaustive statistic is used to form a continuum of probability-statements, of the form

$$Pr\{T < T_p(\theta)\} = P,$$

and using the monotonic property of the functions T_p for all P , this is transformed to the equivalent

$$Pr\{\theta_{1-P}(T) < \theta\} = P,$$

a complete set of probability statements about θ , in terms known for a given sample value T .»

3.2. Example

The following simple example should clarify the basic conceptual issues. Suppose T is an exponentially distributed random variable with parameter $\theta > 0$ and cumulative distribution function:

$$F(t, \theta) := P_\theta(T \leq t) = 1 - e^{-\theta t} \quad (0 < t < \infty).$$

If we let $F(t, \theta) = p$, then the relation $p = 1 - e^{-\theta t}$ implicitly defines functions

$$t_p(\theta) := -\frac{1}{\theta} \log(1 - p), \quad \theta_p(t) := -\frac{1}{t} \log(1 - p).$$

It is immediate that t_p and θ_p are strictly decreasing functions mapping $(0, \infty)$ to itself and inverses of each other:

$$\theta_p(t_p(\theta)) = \theta, \quad t_p(\theta_p(t)) = t.$$

In modern parlance, $t_p(\theta)$ is the *critical value* for a $100(1 - p)\%$ test of significance for the parameter θ in which one rejects θ whenever $T > t_p(\theta)$. Since $T > t_p(\theta) \iff \theta_p(T) < \theta$, the random interval $(0, \theta_p(T)]$ consists of precisely those θ consistent with the observed value of T in the sense that it contains those θ not rejected by the test of significance. Put another way, this random interval contains the true value of θ $100p\%$ of the time; in Neyman’s terminology, it is a $100p\%$ confidence interval for θ conveniently summarizing the corresponding continuum of tests of significance.

For any θ , $F(t, \theta)$ is the cumulative distribution function of the random variable T under θ , but equally, given $1 - e^{-\theta t}$ is symmetric in t and θ , for every t , the function

$F(t, \theta)$ is a mathematical distribution function (in the sense that it is a left-continuous increasing function such that $\lim_{\theta \rightarrow 0} F(t, \theta) = 0$ and $\lim_{\theta \rightarrow \infty} F(t, \theta) = 1$). In the statistical model initially described, however ($\{P_\theta, 0 < \theta < \infty\}$), it is not immediately evident what random variable it is the distribution of. Although we are given a family of distributions P_θ on \mathbb{R}^+ , the sample space of possible values of T , we do not have a corresponding family of distributions P_t on $\Theta := (0, \infty)$, the space of possible values of θ , let alone a joint distribution on $\mathbb{R}^+ \times \Theta$.

Fisher recognized this, for he was at pains to explain (in the context of his more complex example involving the sample correlation coefficient) the intimate relationship between the interpretation of a fiducial percentile and the original sampling distribution of the underlying statistic, and that the former only has a meaning when expressed in terms of the latter. Using the concrete example of four pairs of observations drawn from a bivariate normal and using a table that had been computed in this case, Fisher wrote:

«From the table we can read off the 95 per cent. r for any given ρ , or equally the fiducial 5 per cent. ρ for any given r . Thus if a value $r = 0.99$ were obtained from the sample, we should have a fiducial 5 per cent. ρ equal to about 0.765. The value of ρ can then only be less than 0.765 in the event that r has exceeded its 95 per cent. point, an event which is known to occur just once in 20 trials. In this sense ρ has a probability of just 1 in 20 of being less than 0.765.» [21] (p. 534)

Can such an approach be extended to the case of several parameters? Fisher tried to do this after Jerzy Neyman advanced his own multi-parameter theory of confidence intervals ([23]), but the resulting difficulties led to an important shift in Fisher's views, including almost immediately the abandonment of the sampling interpretation of fiducial probabilities; see generally [24]. This led after Fisher's break with Neyman to a surprising detente with the objective Bayesian Sir Harold Jeffreys. On one occasion, Fisher told Jeffreys he agreed with Jeffreys' approach more than the current school of Neyman, and Jeffreys very emphatically replied, «Yes, we are closer in our approach», as related by S. K. Runcorn, who knew Fisher as a Fellow at Fisher's College Gonville and Caius; see [25] (p. 441). This change is evident in Fisher's last book, *Statistical Methods and Scientific Inference* (1956).

3.3. A Short Period of Peaceful Coexistence

Another interesting aspect of Fisher's views in this early period was his willingness to accept that Bayesian and fiducial probabilities can co-exist.

«The fiducial frequency distribution will in general be different numerically from the inverse probability distribution obtained from any particular hypothesis as to a priori probability. Since such an hypothesis may be true, it is obvious that the two distributions must differ not only numerically, but in their logical meaning.

There is . . . no contradiction between the two statements.»

Fisher later changed his mind on this, arguing at the end of his life that both statements had the same logical meaning (referring to probabilities of the same nature), the fiducial distribution corresponding to the case of the complete lack of information, the Bayesian prior representing a state of very specific information. This is discussed below in Section 5.

4. Bayes and Predictive Inference

Fisher's 1956 book *Statistical Methods and Scientific Inference* (SMSI) was an attempt to explain and justify his statistical methods and views. It is not in a class with his other, prewar books: *Statistical Methods for Research Workers* (1925) gave a coherent framework for statistical theory and methods that dominated the field for more than half a century; *The Genetical Theory of Natural Selection* (1930) established Fisher as one of the three leading population geneticists of the first half of the 20th Century; *The Design of Experiments* (1935) created an entirely new field of statistics. Nevertheless, SMSI is filled with interesting ideas

and examples that reward a close reading. Here, we focus primarily on what it has to say about the two (interrelated) topics of Bayes and predictive inference. Page references are to the third (posthumous) 1973 edition. See the note at the beginning of the bibliography.

4.1. Bayes' "Billiard Ball"

Fisher (SMSI, p. 132) distinguished between what Bayes actually wrote in his *Essay*, and the later, uncritical use of uniform priors by other writers:

«In mathematical teaching the mistake is often made of overlooking the fact that Bayes obtained his probabilities *a priori* by an appropriate *experiment*, and that he specifically rejected the alternative of introducing them axiomatically on the ground that this “might not perhaps be looked at by all as reasonable”; moreover, he did not wish to “take into his mathematical reasoning anything that might admit dispute”. This passage (and additional text) was added in the 2nd edition, pp. 127–128, and still further further material was added to this section (5.26, “Observations of two kinds”) in the 3rd edition as well.»

What was Bayes' experimental alternative to the axiomatic assumption of a prior? Bayes envisaged the following procedure:

- Pick a point uniformly on a rectangular table.
- Project this point onto the horizontal axis of the table.
- Posit the position of this point O on the axis to be uniformly distributed.
- Record whether or not subsequent points X selected on the axis in this way lie to the left of O .

This hypothetical process generates a sequence of independent Bernoulli p -trials:

$$X_1, X_2, \dots, X_n, \dots \quad \text{having } p \sim \text{Unif}[0, 1].$$

Proceeding in this way has some immediate, but simple mathematical consequences: if $S_n = X_1 + \dots + X_n$, $q = 1 - p$ and $n = a + b$, then:

1. $P(S_n = a) = \frac{1}{n+1}$, $0 \leq a \leq n$ (the discrete uniform prior);
2. The posterior of p given $S_n = a$: $\frac{(a+b+1)!}{a!b!} p^a q^b dp$;
3. $P(X_{n+1} = 1 | S_n = a) = \frac{a+1}{n+2}$ (Laplace's “rule of succession”);
4. $P(c, d | a, b) = \frac{(a+b+1)!}{a!b!} \cdot \frac{(a+c)!(b+d)!}{(a+b+c+d+1)!} \cdot \frac{(c+d)!}{c!d!}$.

Fisher had no problem with this approach because Bayes was assuming the uniform prior on p was the result of a physical mechanism, rather than the claim that it captured the idea of a supposed state of ignorance. The point is that if the distribution of p is uniform, then $P(S_n = a)$ is also uniform, and Bayes argued it was this that captured the idea of the absence of knowledge concerning the probability of the event; see [26,27] and [28] (pp. 743–753) for further discussion. Bayes [6] (p. 393) refers to «an event concerning the probability of which we absolutely know nothing antecedently to any trials made concerning it», making it clear the ignorance in question is about the probability of the event, not the event itself. Similarly, Fisher (SMSI, p. 58) refers to the “absence of knowledge *a priori* of the distribution of θ ” as the precondition for the applicability of the fiducial argument.

4.1.1. Karl Pearson Enters the Fray

However, to virtually all readers of Bayes, the subtlety of this passage from the uniformity of the distribution of the continuous parameter p to that of the discrete parameter a went unappreciated, and thus, there remained the nagging problem of justifying the assumption of uniformity. Against this background, in 1920 Karl Pearson, in his paper “The

fundamental problem of practical statistics” in *Biometrika*, thought he had at last discovered a solution to this conundrum ([29] p. 4):

«It has occurred to me that possibly the bull itself is a chimera, and there may be no need whatever to master it. In short, is it not possible that any continuous distribution of a priori chances would lead us equally well to the Bayes-Laplace result? If this be so, then the main line of attack of its critics fails.»

Then, after a page of dense mathematical argumentation, Pearson triumphantly concluded:

«Thus it would appear that the fundamental formula of Laplace ... in no way depends on the equal distribution of ignorance. It is sufficient to assume any continuous distribution—which may vary from one type of a priori probability problem to a second.»

This bold claim did not go unchallenged, and soon after, both Edgeworth [30] (pp. 82–83) and the mathematician William Burnside [31] wrote to challenge Pearson’s bold claim. In fact, Pearson’s assertion is obviously absurd: one need only use any mathematically convenient prior (such as the Dirichlet) to compute the resulting posterior predictive distribution and see it is *not* uniform. Therefore, what did Pearson actually discover? Suppose:

- X is the position of the initial point O ;
- $F(x) = P(X \leq x)$ is the CDF of X ;
- F is continuous, *but not necessarily uniform*.

Then, as is well known, the distribution of the probability integral is uniform: $p = F(X) \sim \text{Unif}[0, 1]$. That is, the distribution of the chance p is still uniform even though the location of X (the mechanism for generating it) need not be.

Karl Pearson eventually understood this crucial distinction [32,33], but only with some help:

«I owe to Miss Ethel Newbold this insight into the exact relation between the two hypotheses.» [33] (p. 192)

4.1.2. Fisher’s Version

Fisher discusses this example in SMSI (not mentioning Pearson), playfully invoking radioactive decay in place of Bayes’ “billiard ball”. Letting $p = e^{-\xi\theta}$ be the probability (for some fixed $\theta > 0$) that an atomic particle does not decay up to time ξ , Fisher [2] (p. 132) considered the case where ξ is a multiple c of an earlier observation x , $\xi = cx$ (so that when $c = 1$, we are back in Bayes’ original setting):

«In many continental countries this distinction, which Bayes made perfectly clear, has been overlooked, and the axiomatic approach which he rejected has actually been taught as Bayes’ method. The example of this Section exhibits Bayes’ own method, replacing the billiard table by a radioactive source, as an apparatus more suitable for the 20th century.»

4.2. The Rule of Succession

Fisher, as a harsh critic of inverse probability, would presumably not feel kindly towards one of its most famous applications, Laplace’s rule of succession, but somewhat surprisingly, he effectively endorses it. In [2] (Chapter III, Section 5), Fisher contrasts three posterior means of p :

$$\frac{a}{N} + \frac{b-a}{2N^2} - \frac{3(b-a)}{2N^3} + \dots \quad (\text{the fiducial approximation}),$$

$$\frac{a+1}{N+2} = \frac{a}{N} + \frac{b-a}{N^2} - \frac{2(b-a)}{2N^3} + \dots \quad (\text{arising from the flat prior}),$$

$$\frac{a+\frac{1}{2}}{N+1} = \frac{a}{N} + \frac{b-a}{2N^2} - \frac{(b-a)}{2N^3} + \dots \quad (\text{arising from the Jeffreys prior } \frac{1}{\pi\sqrt{pq}} dp).$$

All of these agree to first order, and so, as Fisher notes, “to this extent a particular given distribution *a priori* may be nearly equivalent to complete ignorance *a priori*”. This is one of those places in SMSI where significant changes were made between the first and second editions. In [3] (pp. 68–69), Fisher noted that the fiducial approximate and Jeffreys’ posterior means in fact agree to fourth (N^4) order if allowance is made for the effects of the non-normality of the binomial distribution.

Of course, given that the second and third methods closely approximate Fisher’s fiducial approximation (approximate because the variate is discrete), he could scarcely argue (as Venn did in the second edition of his *Logic of Chance*) that the answer given by the rule of succession was absurd on its face.

4.3. Bayesian Prediction: A Conundrum

As a final example of Fisher’s discussion of predictive inference, let us turn to a curious passage on pp. 116–117 of SMSI. Returning to the fourth and last of the mathematical consequences of Bayes’ postulate listed earlier, Fisher observes:

«It may be noticed that the last factor in the expression developed above [predicting (c, d) from (a, b) assuming a uniform prior],

$$\frac{(c+d)!}{c! d!}$$

stands only for the binomial coefficients forming the last line, or base, of Fermat’s arithmetical triangle; but

$$\sum_{c=0}^{c+d} \frac{(c+d)!}{c! d!} p^c q^d$$

is not the only polynomial in p, q the value of which is constantly equal to unity. »

Fisher then gives the following interesting diagram by way of illustration (reproduced courtesy of University of Adelaide Library, Rare Books and Manuscripts) and comments:

«If, in fact, the triangle is extended to any chosen boundary, as for example in the diagram, the thirteen totals outside the boundary are the coefficients $\omega(c, d)$ of a polynomial

$$\sum \omega(c, d) p^c q^d = p^4 + 4p^6 q + 18p^6 q^2 + \dots$$

of which the value is unity for all values of p .

Then, based on previous experience of a successes out of $a + b$, we may infer the probability of reaching the terminal value (c, d) to be

$$\frac{(a+b+1)!}{a! b!} \cdot \frac{(a+c)!(b+d)!}{(a+b+c+d+1)!} \cdot \omega(c, d),$$

if the subsequent trial were made with these endpoints. »

What is going on here? In ordinary coin-tossing, stopping after n -tosses is a simple stopping rule, which permits us to partition the space of outcomes into those sequences of heads and tails that result in k heads, $0 \leq k \leq n$. However, more generally, suppose the

stopping rule is to stop when, as in Figure 1, one exits the area delineated by the dark line. Then, one can again decompose the sequence of possible outcomes into those resulting in c heads and d tails. The probability of any one such sequence is, as before,

$$\frac{(a+b+1)!}{a!b!} \cdot \frac{(a+c)!(b+d)!}{(a+b+c+d+1)!}$$

but now we have to multiple this single-path probability by the number of such paths $\omega(c,d)$ to obtain the probability of having c heads and d tails when we stop. Since we have to terminate in some (c,d) pair, the sum of their probabilities is one. In effect, Fisher is calculating the consequences of a sequential stopping rule assuming Bayes' postulate.

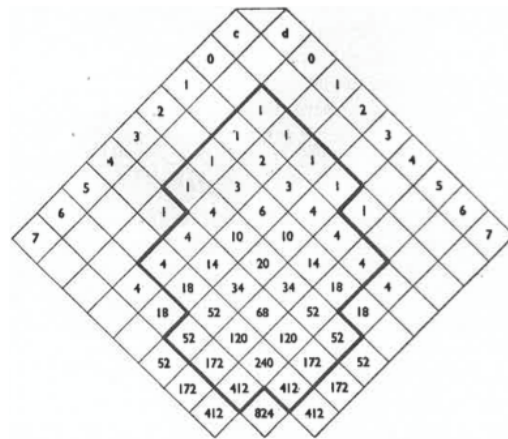


Figure 1. The arithmetical triangle extended.

4.4. Fiducial Prediction

Of course, having first considered Bayesian prediction, it was only to be expected Fisher would also discuss prediction from the fiducial standpoint as well (Section 5.3). Suppose that for $j = 1, 2$ that X_j is the sum of N_j independent exponentials (parameter θ). Then, $X_j \sim \Gamma(N_j, \theta)$, and one has the “pivotal” quantity:

$$\chi^2_{2N_j} = 2\theta X_j.$$

The fiducial distribution of θ given X_1 is $\Gamma(N_1, X_1)$; that is, it has density:

$$\frac{x_1^{N_1}}{(N_1 - 1)!} \theta^{N_1 - 1} e^{-x_1 \theta} d\theta.$$

The predictive distribution of X_2 given X_1 therefore has density:

$$f(X_2|X_1) = \frac{N_1 + N_2 - 1!}{(N_1 - 1)!(N_2 - 1)!} \cdot \frac{X_1^{N_1} X_2^{N_2 - 1}}{(X_1 + X_2)^{N_1 + N_2}}.$$

Suppose that θ has an initial distribution given the Jeffreys' prior $d\theta/\theta$. Then, as Fisher notes, the posterior distribution of θ given X_1 is the same as the fiducial distribution given by Fisher. Therefore, once again, the Fisher and Jeffreys' approaches arrive at similar (here, in fact, identical) conclusions. However, Fisher had much earlier [12] rejected the use of this improper Jeffreys' prior.

Predictive fiducial inference in the case of the normal distribution had been considered much earlier in [34] and was discussed by him at length in [2] (Section 5.4).

4.5. Are Fiducial Probabilities Verifiable?

At the end of his discussion of fiducial prediction for exponential variates, Fisher remarks [2] (p. 118):

«Such fiducial probability statements about future observations are verifiable by subsequent observations to any degree of position required.»

Fisher's comment is readily understood in the case of this particular example, because the ratio $(N_1 X_2) / (N_2 X_1)$ has an F distribution on N_2, N_1 degrees of freedom. For example, if $N_1 = N_2 = 10$, then $P(X_2 / X_1 < 2) = 0.8552$, and the prediction that the observed value of X_2 will be less than twice the observed value of X_1 exactly 86% of the time can be directly verified.

Fisher's more sweeping statement, however, that fiducial probabilities are in general verifiable, made both elsewhere in [2] (pp. 62 and 70) and on other occasions (as in his correspondence with Tukey; see [22] (pp. 221–231)), led to considerable mystification. When asked by Tukey about the case of the Behrens–Fisher distribution, Fisher justified his claim by pointing to the distributions of the two t -statistics entering into the calculation [22] (p. 230)); this suggests that what Fisher had in mind was the use of pivots as building blocks in arriving at the final probability, but hardly answered Tukey's question.

5. The Fiducial Argument Revisited

The fiducial argument as presented in SMSI both differs from and expands on Fisher's earlier prewar views in several important ways and is related to his changing views about Bayes.

5.1. The Logical Status of the Fiducial Distribution

In 1930, Fisher wrote that fiducial and Bayesian posterior distributions differed not only numerically, but also in their logical meaning. A quarter of a century later, Fisher no longer believed this, writing in [2] (p. 59):

«It is essential to introduce the absence of knowledge *a priori* as a distinctive datum in order to demonstrate completely the applicability of the fiducial method of reasoning to the particular and real experimental cases for which it was developed. This last point I failed to perceive when, in 1930, I first put forward the fiducial argument for calculating probabilities. For a time this led me to think that there was a difference in logical content between probability statements derived by different methods of reasoning. They are in reality no grounds for any such distinction. This contradicts Fisher's later statement in [2] (p. 105) that he had simply failed to make clear the need for this “distinctive datum” in his 1930 paper (scarcely credible given essentially the same language appears in two later papers, [12,34]), and then compounded this by going on to criticize Neyman and Pearson for not perceiving its necessity at the time.»

This shift did not occur for some time: in two papers written shortly after, Fisher reprised the fiducial argument, using the examples of estimating the variance σ^2 by s^2 ([12]) and μ using t ([34]) in the case of the normal distribution, in terms virtually identical to those in his 1930 paper. (In his 1935 paper, Fisher added the caveat that “the statistics used contain the whole of the relevant information which the sample provides.” This reflected what he regarded as an important distinction between Neyman's confidence intervals and a fiducial distribution.) Indeed, as late as 1940, in an exchange of letters with the French mathematician Maurice René Fréchet (1878–1973), Fisher sent Fréchet a copy of his 1930 paper as representative of his current views on the fiducial argument and specifically reiterated that “statements of fiducial probability have a logical content different from the more familiar statements of inverse probability” ([22] p. 121).

5.2. Is T Fixed?

The logical status of the fiducial distribution was closely related to the observed value of the statistic T . Initially, Fisher did not view T as fixed. In his 1940 correspondence with Fréchet, Fisher was very clear on this:

«For the population of cases relative to which a fiducial probability is defined, the value of any relevant statistic T is not regarded as fixed. This I have deliberately exerted myself to make clear since my first writings on the subject. . .

I shall be glad to give you all possible support in dissuading mathematicians from thinking that they can obtain a true probability statement logically equivalent to one of the kind aimed at by Bayes' theorem, yet without using the approximate basis of this theorem. Believe me, I have never attempted anything so foolish. The inferences which can be drawn without the aid of Bayes' axiom seem to me of great importance, and quite precisely defined, but are certainly not statements of the distribution of a parameter θ over its possible values in a population defined by random samples selected to give a fixed estimate T .» [22] (p. 124)

Later, Fisher changed his mind on this. Initially, he apparently thought the matter so obvious he did not think through the matter carefully, writing in his 1955 paper that first accepting a symbolic statement such as:

$$Pr\{(\bar{x} - ts) < \mu < (\bar{x} + ts)\} = \alpha,$$

but then, after substituting observed values of \bar{x} and s , proceeding to reject the resulting numerical statement, say:

$$Pr\{92.99 < \mu < 93.01\} = 95 \text{ per cent.},$$

was to violate “the principles of deductive logic”, denying «the syllogistic process of making a substitution in the major premise of terms which the minor premise establishes as equivalent». This was an embarrassing gaffe: it is trivial to think of counterexamples where such a substitution is invalid. For example, if $a < b < c < d$ and X, Y are independent random variables such that

$$P(X = a) = P(X = d) = P(Y = b) = P(Y = c) = \frac{1}{2},$$

then

$$P(Y > X) = \frac{1}{2}, \quad \text{but} \quad P(Y > a) = 1 \quad \text{and} \quad P(Y > d) = 0.$$

Neyman of course pounced on this in a response, and it is surely significant that Fisher never again used this argument. Instead, he conceded (in SMSI, p. 57) there was a legitimate issue:

«The applicability of the probability distribution to the particular unknown value of θ sought by an experimenter, without knowledge *a priori*, on the basis of the particular value of T given by his experiment, has been disputed, and certainly deserves to be examined.»

In doing so, Fisher came up with an interesting and much more defensible (if ultimately still flawed) defense.

5.3. Recognizable Subsets

This invoked the concept of a *recognizable subset*. The simple illustrative example invoked by Fisher was that of tossing a die, and the question was the justification for identifying the relative frequency of, say, an ace in a long sequence of trials with the probability of obtaining an ace in a single toss. Fisher argued [2] (p. 35):

«Before the limiting ratio of the whole set can be accepted as applicable to particular throw, a second condition must be satisfied, namely that no [subset having a different limiting ratio] can be *recognized*. This is a necessary and sufficient condition for the applicability of the limiting ratio of the entire aggregate of possible future throws as the probability of any one particular throw. On this condition we may think of a particular throw, or succession of throws, as a *random* sample from the aggregate, which is in this sense subjectively homogeneous and without recognizable stratification (There is an obvious and close connection here with Richard von Mises's concept of a *kollektiv*, with its twin assumptions of the existence of a limiting frequency (Fisher's first condition) and its invariance under place selection (Fisher's second condition). Note however the two concepts serve very different purposes. von Mises denied the existence of single-case probabilities; the *kollektiv* was designed instead to give a formal mathematical definition of random sequence, and the absence of place selections with differing limiting frequencies was what characterized randomness. For Fisher in contrast the absence of recognizable subsets permitted the identification of the class frequency with the probability of the individual. For Fisher's views on infinite populations and continuous variates as convenient fictions, see [2] (pp. 35 and 114). He was never a frequentist in the von Mises sense).»

Somewhat surprisingly, Fisher went on to write [2] (p. 33):

«This fundamental requirement for the applicability to individual cases of the concept of classical probability shows clearly the role of subjective ignorance, as well as that of objective knowledge in a typical probability statement.»

The knowledge here is that of a “well-defined” population having a known limiting frequency ratio; the ignorance is our “inability to discriminate any of the different sub-aggregates having different limiting frequency ratios”. There are obvious connections here to de Finetti's concept of exchangeability (see [35] for an interesting discussion) and more generally subjective Bayesianism (so much so that Fisher appears to have regretted his wording here, and in the third edition of SMSI, “subjective” and “objective” were changed to “well-specified” and “specific”).

5.4. Recognizability and Fiducial Inference

In general, whether recognizable subsets exist in any specific setting is ultimately a matter of judgment, but in the specific setting of statistical estimation, their existence is a matter of mathematical fact. For the basic example of the t and the probability of the inequality:

$$\mu < \bar{x} - \frac{1}{\sqrt{N}}ts,$$

Fisher asserted [2] (p. 84):

«The reference set for which this probability statement holds is that of the values of μ , \bar{x} , and s corresponding to the same sample, for all samples of a given size of all normal populations. Since \bar{x} and s are jointly Sufficient for estimation, and knowledge of μ and σ *a priori* is absent, there is no possibility of recognizing any sub-set of cases, within the general set, for which any different value of the probability should hold. The unknown parameter μ has therefore a frequency distribution *a posteriori* defined by Student's distribution (Contrast this with Fisher's discussion on pp. 61–62, where he advances a different, equally lapidary justification (“in the absence of a prior distribution of population values there is no meaning to be attached to the demand for calculating the results of random sampling among populations, and it is just this absence which completes the demonstration”). Savage [36] (p. 476), quoting at length from the passage containing this statement, refers to it as illustrating “Fisher's dogged blindness about it all”).»

No proof of this bold statement on Fisher's part was ever given however; presumably, he thought it so obvious as not to require one. Unfortunately for this approach, however, and somewhat surprisingly, recognizable subsets with variable frequencies do exist in the case of the t ; see [37–39]. One cannot make the Bayesian omelet without breaking the Bayesian eggs.

5.5. Fiducial Reprise

Fisher's original fiducial construct gave a general method for constructing confidence intervals in the continuous one-dimensional case. This indeed had a logical content entirely different from that of a Bayesian prior. However, over time, the fiducial distribution—initially so carefully yoked to a pure sampling theory interpretation—began to suffer from mission creep. In 1934, after Neyman had advanced a more general theory of confidence intervals for the multi-parameter case, Fisher added the requirement that the statistics employed in a fiducial argument had to be exhaustive, even though this property did not enter into the logic of his 1930 paper. Later, Fisher attempted to extend the fiducial argument to the multi-parameter case, but could only accomplish this by approaches in which the original sampling interpretation had to be jettisoned.

At the same time—despite his protestations to Fréchet in 1940—Fisher began to write more and more in conditional rather than unconditional terms, culminating in his appeal in 1956 to the absence of recognizable subsets. However, this meant fiducial distributions could not co-exist with Bayesian posterior distributions, since the existence of a Bayesian prior meant recognizable sets with differing relative frequencies might exist. The basic requirement for using the fiducial argument it turned out was an absence of knowledge about the underlying parameters. However, why the absence of knowledge ensured the absence of such sets was never explained—just asserted—and in the case of the t , in fact, turned out to be false.

5.6. Aftermath

From 1956, when SMSI appeared, until his death in 1962, Fisher's book generated considerable controversy. One instance of particular interest is Lindley's review of the book in the journal *Heredity* [40], because it sheds some further light on Fisher's attitude towards Bayes.

George Barnard later reported [41] (p. 184) that Fisher “had been much upset” by Lindley's review, as well he might: Lindley, who was by then a prominent Bayesian, was highly critical of the book, beginning with what he described as a mathematical criticism, which “demonstrates that an error has been made”, followed by “other criticisms which are far more matters of opinion”. Fisher of course would not have cared about the criticisms that were matters of opinion, but the charge of an outright mathematical error was quite another matter. Fisher had asserted that the concept of probability involved in the fiducial argument was “entirely identical with the classical probability of the early writers, such as Bayes” (SMSI, p. 54). Lindley gave an example to show that this could not be the case.

Here is Lindley's example, interesting precisely because it is so simple. Consider the family of probability densities in x :

$$\frac{\theta^2}{\theta + 1}(x + 1)e^{-x\theta} \quad (x \geq 0, \theta > 0).$$

Given one or more observations drawn from such a population, their sum is a sufficient statistic for θ . In particular, given two independent observations x, y , let θ_{xy} denote the result of using x to generate a fiducial distribution for θ , which one then in turn uses as a Bayesian prior to generate a posterior distribution for θ using the other observation y . Then, a trite calculation [42] shows that the result depends on the order in which the two observations are used: $\theta_{xy} \neq \theta_{yx}$, and both differ from the fiducial distribution ψ generated by using x and y simultaneously. If fiducial probabilities were ordinary probabilities, how could this be? Why should the order matter?

Fisher clearly found Lindley's example vexing. In both a letter to George Barnard dated 27 February 1958 (quoted in [41]) and a paper two years later, Fisher interpreted Lindley as criticizing a particular example in [2] (Chapter V, Section 6) "Observations of two kinds", the version of Bayes' calculation "more suitable for the 20th century" discussed earlier, in which a continuous time is combined with a discrete count (while in Lindley's examples, the two successive observations were of the same type). This was to completely miss Lindley's point, which was not the particular example or the method endorsed by Fisher of combining the two observations, but the much more general issue of the lack of consistency if probabilities generated by the fiducial argument were really probabilities in the classical sense. Fisher added in his letter to Barnard:

«In fact, the more I consider it, the more clearly it would appear that I have been doing almost exactly what Bayes had done in the 18th century. As Lindley purports to be a protagonist of Bayes, it seems that his misunderstanding and confusion goes deeper than anyone could imagine.»

As Barnard notes, in his response to Fisher, his own confusion was "hardly less than Lindley's" and "led to an acrimonious encounter between Fisher and myself at a conference later that year".

Here, Fisher has in some ways come almost full circle, in effect defending his example because of its agreement with Bayes' own approach. Of the many other contemporaneous critiques of Fisher's defense of fiducial inference in SMSI and other papers from this period, three of particular note are [43–45].

6. Conclusions

From relatively early on in his career, Fisher was a dedicated opponent of "inverse probability". This was certainly a matter of principle with him, but his vehemence may also have been spurred on by what he regarded as Karl Pearson's springing on him an unwarranted criticism of an important part of his famous 1915 paper on the correlation coefficient. However, it is difficult to replace something by nothing, and Fisher would certainly have recognized the unsatisfactory state of affairs arising from abandoning the Bayesian position without having an adequate alternative ready at hand.

Fisher later thought he had discovered such an alternative in 1930 in the guise of fiducial inference. However, the happy accident that in the case of a single parameter, fiducial percentiles can be spliced together to form a distribution function in the purely mathematical sense (that is, a function increasing from 0 to 1) led him to regard this construct as a viable and principled probabilistic replacement for a Bayesian posterior distribution. His inability to extend this construction to the multi-parameter setting in a way that won general acceptance never caused him to waver from this view. See [46] for an outstanding discussion of the complexities that arise in the case of the *Behrens–Fisher problem*, estimating a difference of means drawn from two normal populations having possibly different variances.

However, Fisher's evolving view of fiducial inference, one which downplayed its sampling theory origins, led him to view the Bayesian and fiducial approaches as viable alternatives based on qualitatively different states of knowledge, and this appears to have led him to view Bayes' original paper with increasing appreciation (Bayes' "mathematical contributions to the *Philosophical Transactions* show him to have been in the first rank of independent thinkers, very well qualified to attempt the really revolutionary task opened out by his posthumous paper" (SMSI, p. 8)). His last book on statistical methods and scientific inference, although controversial in many of its pronouncements, makes for fascinating reading, in part because of its discussion of prediction, contrasting and developing the Bayesian and fiducial approaches.

Note: There were three editions of *Statistical Methods and Scientific Inference*: 1956, 1959, and 1973 [2–5]. Page references in this paper are to the 3rd edition, as being the most available (in the 1990 Bennett reprint). Where there are differences between editions in a passage being cited, this is noted. Sometimes the changes between editions consisted

of additional material being inserted (as in Fisher’s discussion of the rule of succession in the 2nd edition or Todhunter in the 3rd); and sometimes subtle changes that can only be easily identified by a change to a lighter typeface (for an interesting example of the latter, see [28] p. 380). In the 3rd, posthumous edition the changes were based on material Fisher “had entered in his interleaved copy of the book for this purpose sometime before his death” (p. v).

Funding: This research received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under grant agreement No 817257.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fienberg, S.E. When did Bayesian inference become “Bayesian”? *Bayesian Anal.* **2006**, *1*, 1–40. [\[CrossRef\]](#)
2. Fisher, R.A. *Statistical Methods and Scientific Inference*; Oliver & Boyd: Edinburgh, UK, 1956.
3. Fisher, R.A. *Statistical Methods and Scientific Inference*, 2nd ed.; Oliver and Boyd: Edinburgh, UK, 1959.
4. Fisher, R.A. *Statistical Methods and Scientific Inference*, 3rd (posthumous) ed.; Collier Macmillan: London, UK, 1973.
5. Fisher, R.A. *Statistical Methods and Scientific Inference*, 3rd ed.; Reprinted in *Statistical Methods, Experimental Design, and Scientific Inference*; Bennett, J.H., Ed.; Oxford University Press: Oxford, UK, 1990.
6. Bayes, T. An essay towards solving a problem in the Doctrine of Chances. *Philos. Trans. R. Soc. Lond.* **1763**, *53*, 370–418. [\[CrossRef\]](#)
7. Fisher, R.A. On an absolute criterion for fitting frequency curves. *Messenger Math.* **1912**, *41*, 155–160.
8. Fisher, R.A. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika* **1915**, *10*, 507–521. [\[CrossRef\]](#)
9. Fisher, R.A. On the “probable error” of a coefficient of correlation deduced from a small sample. *Metron* **1921**, *1*, 3–32.
10. Fisher, R.A. On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc. A* **1922**, *222*, 309–368.
11. Fisher, R.A. Uncertain inference. *Proc. Am. Acad. Arts Sci.* **1936**, *71*, 245–258. [\[CrossRef\]](#)
12. Fisher, R.A. The concepts of inverse probability and fiducial probability referring to unknown parameters. *Proc. R. Soc. A* **1933**, *139*, 343–348.
13. Aldrich, J.R.A. Fisher and the making of maximum likelihood 1912–1922. *Stat. Sci.* **1997**, *12*, 162–176. [\[CrossRef\]](#)
14. Edwards, A.W.F. What did Fisher mean by “inverse probability” in 1912–1922? *Stat. Sci.* **1997**, *12*, 177–184. [\[CrossRef\]](#)
15. von Kries, J. *Die Principien der Wahrscheinlichkeitsrechnung, Eine Logische Untersuchung*; Mohr: Tübingen, Germany, 1886.
16. von Kries, J. *Die Principien der Wahrscheinlichkeitsrechnung, Eine Logische Untersuchung*, 2nd ed.; Mohr: Tübingen, Germany, 1927.
17. Keynes, J.M. *A Treatise on Probability*; Macmillan: London, UK, 1921.
18. Soper, H.E.; Young, A.W.; Cave, B.M.; Lee, A.; Pearson, K. On the distribution of the correlation coefficient in small samples. Appendix II to the papers of “Student” and R. A. Fisher. A cooperative study. *Biometrika* **1917**, *11*, 328–413. [\[CrossRef\]](#)
19. Edwards, A.W.F. R.A. Fisher on Karl Pearson. *Notes Rec. R. Soc. Lond.* **1994**, *48*, 97–106.
20. Coolidge, J.L. *An Introduction To Mathematical Probability*; Clarendon Press: Oxford, UK, 1925.
21. Fisher, R.A. Inverse probability. *Proc. Camb. Philos. Soc.* **1930**, *26*, 528–535. [\[CrossRef\]](#)
22. Bennett, J.H. *Statistical Inference and Analysis: Selected Correspondence of R. A. Fisher*; Clarendon Press: Oxford, UK, 1990.
23. Neyman, J. On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *J. R. Stat. Soc.* **1934**, *97*, 558–625. [\[CrossRef\]](#)
24. Zabell, S.L. R.A. Fisher and fiducial argument. *Stat. Sci.* **1992**, *7*, 369–387. [\[CrossRef\]](#)
25. Box, J.F. R. A. Fisher: *The Life of a Scientist*; Wiley: New York, NY, USA, 1978.
26. Edwards, A.W.F. Commentary on the arguments of Thomas Bayes. *Scand. J. Stat.* **1978**, *5*, 116–118.
27. Stigler, S.M. Thomas Bayes’ Bayesian inference. *J. R. Stat. Soc. Ser. A Gen.* **1982**, *145*, 250–258. [\[CrossRef\]](#)
28. Zabell, S.L. Philosophy of inductive logic: The Bayesian perspective. In *The Development of Modern Logic*; Haaparanta, L., Ed.; Oxford University Press: Oxford, UK, 2009.
29. Pearson, K. The fundamental problem of practical statistics. *Biometrika* **1920**, *13*, 1–16. [\[CrossRef\]](#)
30. Edgeworth, F.Y. Molecular statistics. *J. R. Stat. Soc.* **1921**, *84*, 71–89. [\[CrossRef\]](#)
31. Burnside, W. On Bayes’ formula. *Biometrika* **1924**, *16*, 189. [\[CrossRef\]](#)
32. Pearson, K. Note on the “fundamental problem of practical statistics”. *Biometrika* **1921**, *13*, 300–301.
33. Pearson, K. Note on Bayes’ theorem. *Biometrika* **1924**, *16*, 190–193. [\[CrossRef\]](#)
34. Fisher, R.A. The fiducial argument in statistical inference. *Ann. Eugen.* **1935**, *6*, 391–398. [\[CrossRef\]](#)
35. Lindley, D.V.; Novick, M.R. The role of exchangeability in inference. *Ann. Stat.* **1981**, *9*, 45–58. [\[CrossRef\]](#)
36. Savage, L.J. On rereading R. A. Fisher. *Ann. Stat.* **1976**, *4*, 441–500. [\[CrossRef\]](#)
37. Buehler, R.J. Some validity criteria for statistical inferences. *Ann. Math. Stat.* **1959**, *30*, 845–863. [\[CrossRef\]](#)
38. Buehler, R.J.; Feddersen, A.P. Note on a conditional property of student’s *t*. *Ann. Math. Stat.* **1963**, *34*, 1098–1100. [\[CrossRef\]](#)

39. Brown, L. The conditional level of Student's t test. *Ann. Math. Stat.* **1967**, *38*, 1068–1071. [[CrossRef](#)]
40. Lindley, D.V. Review: Statistical Methods and Scientific Inference. *Heredity* **1957**, *11*, 280–283. [[CrossRef](#)]
41. Barnard, G.A.R.A. Fisher—A true Bayesian? *Int. Stat. Rev.* **1987**, *55*, 183–189. [[CrossRef](#)]
42. Lindley, D.V. Fiducial distributions and Bayes theorem. *J. R. Stat. Soc. B* **1958**, *20*, 102–107. [[CrossRef](#)]
43. Pitman, E.J.P. Statistics and science. *J. Am. Stat. Assoc.* **1957**, *52*, 322–330. [[CrossRef](#)]
44. Tukey, J.W. Some examples with fiducial relevance. *Ann. Math. Stat.* **1957**, *28*, 687–695. [[CrossRef](#)]
45. Dempster, A.P. Further examples of inconsistencies in the fiducial argument. *Ann. Math. Stat.* **1963**, *34*, 884–891. [[CrossRef](#)]
46. Wallace, D.L. The Behrens-Fisher and Fieller-Creasy problems. In *R. A. Fisher: An Appreciation*; Lecture Notes in Statistics; Fienberg, S.E., Hinkley, D.V., Eds.; Springer: Berlin/Heidelberg, Germany, 1980; Volume 1; pp. 119–147.