International Statistical Review (2000), 68, 2, 155–172, Printed in Mexico

Fisher's "Inverse Probability" of 1930

John Aldrich

Department of Economics, University of Southampton, Southampton, SO17 1BJ, UK
E-mail: jca1@soton.ac.uk

Summary

In 1930 R.A. Fisher put forward the fiducial argument. This paper discusses the argument and its origins in Fisher's earlier work. It also emphasises the contribution of Mordecai Ezekiel to the 1930 publication.

Key words: Fiducial argument; Confidence intervals; Fisher; Ezekiel; Hotelling.

Introduction

R.A. Fisher's 'Inverse Probability' of 1930 has a place in history. Fisher's early papers on fiducial probability followed its reasoning and the first work on confidence intervals reflected its influence. Yet 'Inverse Probability' is something of a mystery—in how it relates to Fisher's earlier work and in how its different parts fit together. I draw on hitherto unused sources to try to get a better understanding of how it came to be written, of how it relates to Fisher's earlier work and of how it hangs—or does not hang—together.

Most accounts of the pre-history of the fiducial argument—see Zabell (1992), Barnard (1995) and Edwards (1995)—take off from what Fisher (1935a, §62, pp. 195–6) wrote about the role of the Rothamsted scientist E.J. Maskell in developing intervals based on the t-distribution in 1925/6. I follow a different line, one more closely linked to the 1930 publication and one for which there are contemporary records—correspondence and publications. It involves two of Fisher's American admirers—Mordecai Ezekiel and Harold Hotelling. In the 20s and 30s American agricultural economists produced a very impressive body of applied regression work; see Fox (1989) and Morgan (1990) for details. Fisher became the movement's authority on regression theory, thanks in part to Ezekiel and Hotelling's promotion. Ezekiel was in the forefront of the movement; Hotelling was a theorist who (1927b, p. 412) hailed Fisher's work as of 'revolutionary importance' and made the first American contribution to the new regression theory.

In 'Inverse Probability' Fisher (1930a p. 534) states: 'I have recently received from the American statistician, Dr. M. Ezekiel, graphs giving to a good approximation the fiducial 5 per cent. points of simple and multiple correlations for a wide range of cases.' The graphs were drawn for Ezekiel's book *Methods of Correlation Analysis* (1930) and Ezekiel considered them useful representations of the information in Fisher's tables. However something was added in the translation—a probability statement. Fisher did not reject the intrusion but elaborated it in his own way—as the fiducial argument.

Ezekiel was presenting intervals for correlations, regression coefficients and means but he had not written about probability intervals before. I suggest that Hotelling had a part in this change. Hotelling figures in accounts of pre-historic (i.e. pre-Neyman) confidence interval work—see Neyman (1952, pp. 221–2 & 1977, pp. 127–8)—as isolated from the events of 1930 but he seems to have been part of the background to those events. Hotelling and the agricultural economist Holbrook Working constructed a 'range of error'—a confidence interval—for the regression line. Their paper—Working & Hotelling (1929)—was presented together with one by Ezekiel (1929) and later Ezekiel put their construction in his book. It seems likely he got the idea of making probability statements about intervals from them. Hotelling also spent some time at Rothamsted and he may have influenced Fisher directly. The Rothamsted statistician J.O. Irwin recalled discussing interval inference with Fisher, Wishart and Hotelling—see Bartlett (1984, p. 109) but there are no details. The letters between Hotelling and Fisher preserved at Columbia University and at the University of Adelaide begin in 1928. They are not informative on any of the issues discussed in this paper—unless silence is informative.

These materials bear on how the paper came to be written but another 'new' source bears on what Fisher was saying. This is the abstract he wrote for the British Association. This was published but I reproduce it here for convenience—as Appendix A. 'Inverse Probability' is a poorly constructed paper and every scrap helps in understanding it. There are gaps in the argument and also contradictions. The best—or worst—example of the latter is the way the paper (pp. 532 & 4) negotiates the change in the domain of probability: 'We can state the relative likelihood that an unknown correlation is +0.6, but not the probability that it lies in the range .595 – .605' is first qualified 'There are, however, certain cases in which statements in terms of probability can be made with respect to the parameters of the population ... ' and then contradicted when correlation is used to illustrate the possibility. I suggest such infelicities reflect haste in putting the paper together.

Before 1930 Fisher rejected the possibility of probability statements about parameters along with inverse probability; Section 1 considers why. His way of constructing intervals without probabilities is examined in Section 2. Section 3 introduces Hotelling's probability intervals. Section 4 presents the Ezekiel/Fisher interaction. Sections 5 and 6 consider the new way of making probability statements—the fiducial argument—while Section 7 returns to inverse probability. Section 8 describes the status of the fiducial argument in 1930. Appendix A reprints the abstract of 'Inverse Probability' and Appendix B records Fisher's later transactions with Ezekiel.

1 Probability and Inverse Probability

Fisher's publications before 'Inverse Probability' show that, if he knew anything, it was that probability statements could *not* be made about parameters. His first publication (1912) picked over the standard theory of errors package, approving the least squares values—as an application of the 'absolute criterion'—but rejecting the possibility of obtaining 'an expression for the probability that the true values of the elements [parameter values] should lie within any given range' (p. 160).

Fisher (1921 & 1922a) re-constructed the absolute criterion as maximum likelihood and in the process set out the relationship between his method and Bayesian methods—'inverse probability' in the language of the day. Fisher originally used the expression 'inverse probability' in a peculiar way but from 1925—at least—he adhered to the conventional usage. Aldrich (1997) and Edwards (1997) discuss this change and other matters arising from Fisher's work in the period 1912–1922.

Fisher considered two forms of inverse probability: in one a uniform prior represented ignorance and in the other the prior reflected a process of sampling from a super-population. His conclusions about probability and inverse probability were summarised in the *Statistical Methods*[1] (1925a, pp. 9–11):

> The deduction of inferences respecting samples, from assumptions respecting the populations from which they are drawn, shows us the position in Statistics of the **Theory of Probability**. ... For many years, extending over a century and a half, attempts were made to extend the domain of the idea of probability to the deduction of inferences

[1] My shorthand for *Statistical Methods for Research Workers*.

> respecting populations from assumptions (or observations) respecting samples. Such inferences are usually distinguished under the heading of Inverse Probability, and have at times gained wide acceptance. ... Inferences respecting populations, from which samples have been drawn, cannot be expressed in terms of probability, except in the trivial case when the population is itself a sample from a super-population the specification of which is known with accuracy.

Likelihood provides the means of expressing sample to population inference. The division of uncertain inference into the domains of probability and likelihood was still in place—if not in force—in 1930: for an example see the second paragraph of the abstract in Appendix A.

The publication of 'Inverse Probability' seems to show that for Fisher this incapacity of probability was not a fundamental tenet—as it is, say, for Edwards (1972)—but a provisional conclusion based on the failure of the 'non-trivial' form of inverse probability—the form which involved maximising the posterior distribution derived from a uniform prior. Fisher (1922a, p. 325) had specific objections to this form of inference:

> Apart from evolving a vitally important piece of knowledge ... out of an assumption of complete ignorance, it is not even a unique solution.

The solution is not unique because the inference depends on the parametrisation chosen. Lack of invariance to reparametrisation was already Fisher's objection in 1912 and invariance would be important again in 1930—see Section 6 below. In the fiducial scheme of 1930 a prior is not needed and so is no longer a 'vitally important piece of knowledge'. Suppose one could find an invariant method which did not need a prior distribution ....

2 Probable Limits without Probability

Before 1930 Fisher did not attach probabilities to interval statements but he had a technique for interval inference. This section considers the evolution of this technique from 1925 to 1928—the final stop before 1930.

Statistical Methods (1925a) gives limits for the mean, standard deviation, difference of means and intra-class correlation, based on normality, exact or approximate; perhaps they were Fisher's way of updating the practice of supplying estimates and probable errors. The limits are presented without fuss and described in a few scattered sentences.

The limits are obtained by inverting a two-tailed significance test and so rest upon the significant/insignificant distinction. This is introduced in the book's section on the normal distribution; Fisher (p. 47) explains the 'significance' of a deviation:

> The value for which $P = .05$, or 1 in 20, is 1.96 or nearly 2; it is convenient to take this point as a limit in judging whether a deviation is significant or not. Deviations exceeding twice the standard deviation are thus formally regarded as significant. Using this criterion, we should be led to follow up a negative result only once in 22 trials, even if the statistics are the only guide available.

'To follow up a negative result' means taking a deviation as signifying an effect when it only represents a sampling fluctuation.

Fisher (pp. 51–2) explains the limits through an example, fitting a normal distribution from a large sample. The sample mean is 68.6435 with standard error .0792:

> From these values it may be seen that our sample shows significant aberration from any population whose mean lay outside the limits 68.48″–68.80″ and it is therefore likely that the mean of the population from which it was drawn lay between these limits; similarly

> it is likely that its standard deviation lay between 2.59' and 2.81'.

In the sixth edition of 1936 Fisher changed 'likely' to 'probable, in the fiducial sense,' without altering the argument.

A similar treatment followed the instruction, 'estimate the correlation in the population from which [the sample] was drawn, and find the limits within which it probably lies' (p. 184). Fisher's use of 'probably' here but 'likely' elsewhere does not seem to correspond to any difference in the situation though it indicates a remarkable casualness from someone who made so much of the probability/likelihood distinction. The glosses on limits include: 'we shall seldom be wrong in concluding ... ' (p. 185) and 'we can ... place its value with some confidence' (p. 104).

The second edition of 1928 added a chapter on estimation expounding maximum likelihood and describing its merits. 'Probable limits' are described (§55, pp. 249 & 253): they take the form, estimate $\pm$ 2 standard errors, and are based on the asymptotic normality results in (1922a & 1925b). With such limits 'we might judge roughly ... '. Here, as in the correlation example, testing is not mentioned.

The limit work in Fisher's 1928 paper on the multiple correlation coefficient deviated from this pattern. The correlation limits in the *Methods* depend on the approximate normality of the inverse *tanh* transformed correlation—a device Fisher first exploited in 1921. Although the main contribution of the 1928 paper was the exact distribution of the multiple correlation coefficient, the paper's table was based on a non-normal approximation using the transformed quantities

$$B = \sqrt{n_2}\,tanh^{-1}R \text{ and } \beta = \sqrt{n_2}\,tanh^{-1}\rho$$

where R and ρ are the sample and population multiple correlations respectively and n_2 is the number of degrees of freedom ($= n - n_1 - 1$, where n is the sample size and n_1 is the number of independent variables).

The table (pp. 665–6) gives the (one-sided) 5% points of B for given values of n_1 and β. The table can be used for testing: for a hypothesised value of β, one rejects for the reported value of B or less. However Fisher (p. 666) described the limits use of the table:

> The values tabulated are the values of B which will be exceeded by chance in 5 per cent. random trials, and which therefore give a presumption that β is really greater than the value postulated. Thus, when $n = 3$, it may be seen at a glance that a value $B = 5.7$ indicates that β probably exceeds 3.8.

What is 'seen' is the inversion of a significance test though, for the first time, a one-sided test: the value $B = 5.7$ is significant for $\beta = 3.8$ and larger values of β. The interval no longer involves a two-sided limit based on a normal approximation and the interval conclusion is the centre of attention—instead of being subordinated to the test. However once again there is a 'presumption', a 'probably exceeds' but no probability value. The probability value was added when Ezekiel charted the table.

This treatment of limits may have owed something to Maskell's 'application' of the t-distribution—as Fisher (1935a, p. 195) called it—for there also the limit technique was uncoupled from the normal distribution. Unfortunately we have no direct knowledge of Maskell's contribution; neither of the publications—Maskell (1929 and 1930)—that Fisher mentioned contains any limits analysis. As described in 1935, Maskell's limits were of the same kind as Fisher's—i.e. there was no probability statement attached. Later—see Bennett (1990, p. 212) and Edwards (1995, p. 800)—Fisher recalled Maskell proposing a probability distribution for the parameter. Fisher also recalled rejecting the proposal.

Of these pre-1930 intervals Fisher never says how 'seldom', how 'confident', how 'probable', or how 'likely'. The silence can bear two obvious but contradictory constructions: one does not need to say because the interval naturally inherits the number from the test or one cannot say. In 1936 he stamped the first interpretation onto the 1925 work but in 1925 he opposed probabilisation. Black became white without any grey or any wavering.

The 1928 paper is just an interpretation away from 'Inverse Probability'. In her account of the origins of the fiducial argument Box (1978, p. 254) reads into the paper the pivotal interpretation which only began to emerge in Fisher (1935b). She also brings forward by a year the date of the first airing of the fiducial argument—to the *1929* British Association meeting; Edwards (1995) gives the correct date. There is no direct evidence that anything happened in 1929 but it is likely that Hotelling probabilised Fisher's correlation interval.

3 The Range of Error of a Trend

Hotelling arrived in Rothamsted a practised probabiliser of intervals; for biographical information see Darnell (1988) and Smith (1978). He had published two papers which included interval calculations—one based on inverse probability, the second apparently not. The first paper (1927a) used a population growth equation to extrapolate and interpolate population size. The interpolation analysis (p. 312) gave the Bayes posterior for the logarithm of population size assuming a uniform prior and conditional on the known values of population at an earlier and a later time. His Figure IV (p. 311) has dotted lines giving 'limits between which at any time the population is as likely or not to lie', i.e. 50% probability intervals.

The technique of Working & Hotelling (1929) was different; it built on Fisher's (1922b and 1925a) application of the t-distribution to the intercept and slope of the normal regression model. They describe the objective:

> It is desirable to set limits concerning which we can say with a considerable probability that they will not be transgressed by any trend line, differing from that calculated, merely because of the fluctuations of sampling.

They went a stage or three further than Fisher, skipping the intervals for intercept and slope, to give (p. 81) first an interval for the expected value of Y (the 'trend') associated with a given year x. In modern dress, the interval is:

$$a + b(x - \bar{x}) \pm t_{0.025}^{(n-2)} \left(s_a^2 + (x - \bar{x})^2 s_b^2\right)^{\frac{1}{2}} .$$

They gloss the formula, 'for any particular year we may say that the chances are 19 to 1 that the point representing the value of the true trend function for that year will lie between the two branches of the hyperbola.' This sounds like inverse probability.

They went on to construct a 'graphic representation of the range of error of a trend'—a simultaneous 'interval' for the entire line—by taking the envelope of the hyperbolae appropriate to each value of x. In the text they (p. 84) refer to 'the probability of the true line cutting this hyperbola' but a footnote comments on the probability language:

> This mode of speech is an ellipsis, unless one accepts inverse probability. What is meant is that if a certain line cutting the χ^2 hyperbola is the true line then the probability that the calculated trend line could be obtained by chance is less than P. A similar elliptical interpretation might be given to many statements involving probable errors.

Given Hotelling's past it was probably understood that one does *not* accept inverse probability. This note shows that, though Working and Hotelling knew how to construct confidence intervals, they were not so good at describing what they were doing: surely 'what is meant' is not a proposition about the probability distribution of the statistic but some deduction from that proposition. Hotelling was more articulate later. (See the account in Hotelling (1931b, pp. 377–8) which Neyman (1952, pp. 221–2) has discussed.)

A much more effective 'elliptical' point had been made earlier by E.B. Wilson (1927, p. 209). Referring to Bernoulli trials:

> Strictly speaking, the usual statement of probable inference ... is elliptical. Really the chance that the true probability p lies outside a specified range is either 0 or 1; for p actually lies within that range or not. It is the observed rate p_0 which has a greater or less chance of lying within a certain interval of the true rate p. If the observer has had the hard luck to have observed a relatively rare event and to have based his inference thereon, he may be fairly wide of the mark.

It may be a coincidence that Wilson had a paper (1929) in the same December 1928 *ASA* conference session as Working & Hotelling. (Neyman (1952, p. 222) and Hacking (1980) discuss Wilson's paper. Wilson later had plenty of contact with Fisher but there is nothing to suggest that Fisher knew this paper.)

Hotelling spent the second half of 1929 at Rothamsted. Irwin (1953, p. 223) recalled how he and Hotelling went through Fisher's 1928 paper. Hotelling seems to have left no record of what he did at Rothamsted but his account of 'recent improvements in statistical inference' (1931a, p. 83) has an interesting view of the significance of the 1928 paper:

> When a multiple correlation coefficient is calculated from a sample, limits can now be found between which the true correlation may with specified reliability be said to lie. Prior to 1929 [sic] this was not possible.

This perspective may have come later but it is the natural one for someone used to probabilising intervals. The difficult part was the distribution theory; once that was done the limits work was easy.

In 1940 Fisher told Fréchet (see Bennett (1990, pp. 118–134)) that

> many people had been arguing in this [fiducial] way from the moment that theoretical distributions, such as χ^2, t and z were first tabulated so as to show the values taken at different levels of significance (values of P) instead of showing the values of P for different levels of χ^2 etc.

Perhaps Fisher had already heard what Hotelling had to say. We turn to Ezekiel who seems to have been more visibly affected.

4 The Reliability of Correlation Conclusions

Ezekiel's *Methods of Correlation Analysis* was, as Fox (1989, p. 67) suggests, 'by far the most comprehensive work on applied regression analysis published up to that time'. As well as consolidating the work of fellow researchers at the Bureau of Agricultural Economics, it brought the work of 'Student' and Fisher to the attention of American readers: 'During the last two decades, the English statisticians 'Student' and R.A. Fisher have been developing more exact methods of judging the reliability of conclusions, particularly where those conclusions involve correlation or are based on small samples.' (1930, p. vi). Ezekiel was making Fisher's work 'more readily available for non-mathematical students'.

There were no probability intervals in Ezekiel's earlier publications and with one exception all those he now published were new. Ezekiel (pp. 254–5) describes the Working & Hotelling analysis and it seems likely that he worked from their paper, taking their notion of a probability interval, filling in the simpler t results they omitted and reading Fisher's interval statements—Section 2 above—through their eyes. However Ezekiel was a more carefree probabiliser with not even a footnote on elliptical language.

The treatment of the 'reliability' of statistical results in the *Methods of Correlation Analysis* is naturally much more extensive than the treatment in Working & Hotelling. It begins with an account of the standard error of the mean, including a limit analysis based on the t-distribution. The limits analysis used a chart (p. 392) representation of the t-table giving

> The probability that an average (or other constant) computed from a sample lies further from the true value than a stated number of times the computed standard error, for samples with [a varying number of] observations.

In true confidence interval style Ezekiel (p. 23) illustrates his limit propositions by a sampling experiment—from a real population–indicating how often the intervals are correct. His language is more equivocal: 'The true value yield is between ... and ... with one chance in ... of being wrong.'

Ezekiel consulted Fisher but only at a very late stage in the writing of his book and only in connection with correlation. On April 23rd 1930 Ezekiel wrote to Fisher asking about the interpretation of the notation in the 1928 paper, which was new to him. He wanted an early reply, 'As I would like to make use of this latest development of your methods for judging the reliability of observed multiple correlations in some material which I am about to publish, I am anxious to make exactly the right interpretation of your conclusions.' Fisher replied on May 2nd explaining the notation and giving advice on calculations Ezekiel might do—not on the interpretations he should draw. On June 15th Ezekiel wrote back

> Following out your suggestions, I have computed through the probable minimum value of the correlation coefficient for a series of values of n' and of the observed correlation. I have managed to work out calculating charts which show these results in very concise form and yet can be read accurately to the second decimal place.

He sent Fisher his chapter on the reliability of correlation conclusions and four correlation charts—for different numbers of variables—with each chart displaying curves for six sample sizes. The axes are the 'probable true correlation' and the 'correlation observed in sample' and the caption reads 'Under conditions of simple sampling [random sampling] the odds are 19 to 1 that the correlation in the universe would be at least as high as the 'probable true correlation'.' The charts—printed in Ezekiel (1930, pp. 393–6)—clearly descend from Fisher's 1928 table: the stated odds replace Fisher's 'presumption that β [= transformed ρ] is really greater than ... ' and the correlations ρ and r themselves figure in the charts.

Ezekiel (p. 257) says very little about the probability interpretation of the correlation interval

> [Figure B] is based upon the idea that, if the chances are 19 to 1 that the true correlation is at least a specified value, that value will be a reasonable one to use as the probable minimum correlation existing in the universe from which the sample was obtained.

The chapter has more discussion of the intervals for regression coefficients; the exposition mixes reports of sampling experiments with talk of parameter probability statements. (Thus on p. 253):

> The odds are thus better than 98 to 2 that the true value is between 0.14 and 0.86—if the sample was drawn under such conditions that the formulas of simple sampling hold true.

In the second edition of 1941—see Appendix B below—Ezekiel kept the examples and charts but rewrote the commentary.

On June 15th Ezekiel asked Fisher if he had given 'a fair statement of the methods which you have developed for judging the reliability of correlation results'. Of the graphs, he asked, 'How do you like this method of presentation?' On June 26th Fisher congratulated him 'on the skill with which you have dealt with a very difficult subject' adding 'I think the charts will be very useful'. On July 23rd the Cambridge Philosophical Society received 'Inverse Probability'. On the same day the CPS received another paper from Rothamsted by Frances Elizabeth (Betty) Allan. She was a research student who worked with Fisher from December 1929 to July 1930 and did the calculations for 'Inverse Probability'.

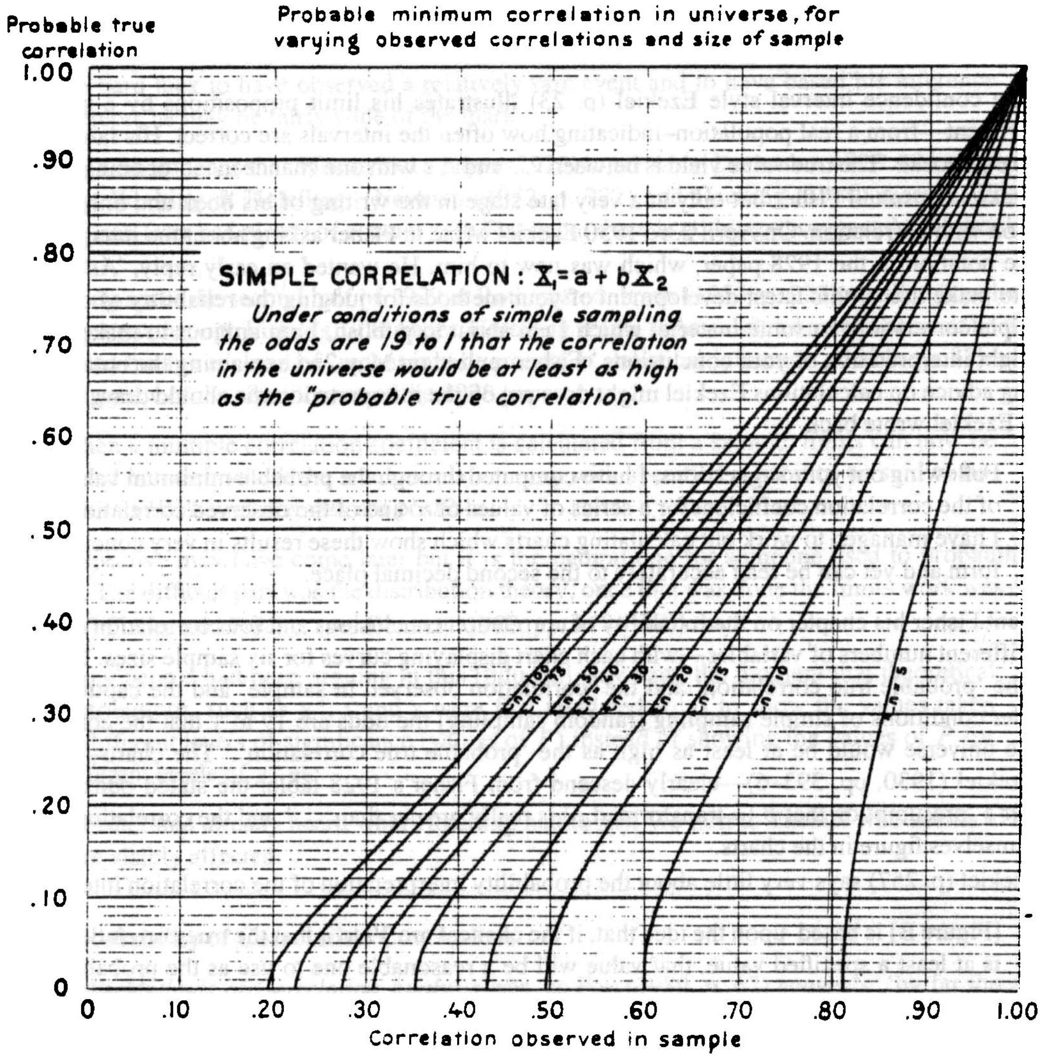

Figure 1. *The probable minimum correlation existing in the universe from which the sample was drawn, for samples of various sizes, and for varying observed correlations, for simple correlation. Figure B from Ezekiel (1930, p. 393) with curves for n =* 5, 10, 15, 20, 30, 40, 50, 75, 100.

Allan's (1930) two page piece, 'A Percentile Table of the Relation between the True and the Observed Correlation Coefficient from a Sample of 4', presents a table giving '95 per cent. values of the transformed correlation z, for different values of the [transformed] correlation ζ in the population sampled'. The point of her note is to give values for the exact distribution of z for a sample size of 4 so they can be compared with the normal approximation. Her table, analogous to the one in Fisher (1928), became the table in 'Inverse Probability' when re-worked to feature, not the transformed quantities z and ζ, but the correlations r and ρ. Allan's paper is a routine contribution to the Rothamsted correlation project. It was not particularly timely as it could have been written at any time from 1921 when Fisher first employed the inverse tanh transformation. However Fisher may have put Allan on to the topic when he encouraged Ezekiel to work on similar lines and then had the table reworked in response to Ezekiel's charts.

5 The Fiducial Argument

These three sections will examine 'Inverse Probability'. I treat the fiducial argument first as it is closely linked to the material in Sections 2–4 but it actually appears five pages into the eight page paper. Fisher's order of business is inverse probability, fiducial probability and a comparison of fiducial probability with inverse probability.

Fisher's formulation of the new kind of 'probability statement' is sharper and grander than anything in Ezekiel or Working & Hotelling. The argument (p. 532) starts out like an abstract restatement of the commentary on the 1928 table referring now to a single parameter θ with maximum likelihood estimator T:

> If T is a statistic of continuous variation and P the probability that T should be less than any specified value, we have then a relation of the form
>
> $$P = F(T, \theta)$$
>
> If now we give to P any particular value such as .95, we have a relationship between the statistic T and the parameter θ, such that T is the 95 per cent. value corresponding to a given θ, and this relationship implies the perfectly objective fact that in 5 per cent. of samples T will exceed the 95 per cent. value corresponding to the actual value of θ in the population from which it is drawn. To any value of T there will moreover be usually a particular value of θ to which it bears this relationship; we may call this the 'fiducial 5 per cent. value of θ' corresponding to a given T.

So far this is 1928 with a name, 'fiducial 5 per cent. value', for the parameter value tabulated, but probability is at hand:

> If, as usually if not always happens, T increases with θ for all possible values, we may express the relationship by saying that the true value of θ will be less than the fiducial 5 per cent. value corresponding to the observed value of T in exactly 5 trials in 100. By constructing a table of corresponding values, we may know as soon as T is calculated what is the fiducial 5 per cent. value of θ, and that the true value of θ will be less than this value in just 5 per cent. of trials.

Zabell (p. 371) comments

> Fisher not only gave a clear and succinct statement of (what later came to be called) the confidence interval approach but (and this appears almost universally unappreciated) he also gave a general method for obtaining such estimates in the one-dimensional case.

The second point was appreciated in the first publications to write of confidence intervals—Neyman (1934) and Clopper & E. S. Pearson (1934).

For Pearson—see Section 8 below—Fisher should have stopped right there. For this 'confidence' passage continues with a probability gloss on the clause, 'the true value of T will be less than this value in just 5 per cent. of trials', which changes it into something else:

> This then is a definite probability statement about the unknown parameter θ which is true irrespective of any assumption as to its *a priori* distribution.

Neyman (1941, p. 128) relates that he had thought such statements in Fisher were 'lapsus linguae'.

Fisher illustrates the argument with Allan's table re-expressed in terms of r and ρ, with the headings fiducial 5% values of ρ and 95% values of r.

> Thus if a value $r = .99$ were obtained from a sample, we should have a fiducial 5 per cent. ρ equal to about .765. The value of ρ can then only be less than .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 .765.

The phrase *in this sense* is crucial. Does it mean, in this special sense, so that the clause following is taken as a short form of the preceding sentence, a controlled lapsus lingua, a deliberately elliptical form of expression? Or less naturally, does it mean in this same sense, so that sentence and clause are using probability in the same sense?

The same ambiguity invests the 'fiducial distribution' which Fisher now introduces. Here the natural reading is that the 'fiducial distribution' is a probability distribution not merely a formal device for generating all possible percentiles:

> In the same way ... any other percentile in the fiducial distribution of ρ could be found or, generally, the fiducial distribution of a parameter θ for a given statistic T may be expressed as
>
> $$df = -\frac{\partial}{\partial\theta}F(T,\theta)d\theta$$
>
> while the distribution of the statistic for a given value of the parameter is
>
> $$df = \frac{\partial}{\partial T}F(T,\theta)dT.$$

The later history of Fisher and the fiducial argument is that of his trying to make the probability interpretation stick but in 1930 nobody was pressing him to say what he meant.

6 A Remarkable Difference

The pages preceding the exposition of the fiducial argument contain material relating—none too clearly—to the conditions on which the argument rests. Fisher emphasised the difference between discrete and continuous random variables which led into an explanation (p. 534) of why the argument had been overlooked for so long. He held that modern distributional work was different in a way relevant to the kind of inference possible (p. 529):

> The introduction of quantitative variates, having continuous variation in place of simple frequencies as the observational basis, makes also a remarkable difference to the kind of inference which can be drawn.

Of course 'quantitative variates' had not just been introduced but those to which the 1930 form of inference could be applied were very special. Fisher does not list assumptions or remark on their restrictiveness but the continuous variate is a maximum likelihood estimator, the exact distribution of which is known and which involves one parameter—see also his (1935b, p. 391) and (1936, p. 253). Apart from the various correlations which he derived, the only important example was s^2, which he attributed to 'Student' (1908); this was the subject of his second fiducial paper (1933). The range of the argument was very limited: it did not extend to the regression coefficients and means treated by Ezekiel.

One of the eight pages of 'Inverse Probability' is spent on the properties of continuous variables. The importance of continuity is not explained but it is implicit that as the parameter space is continuous the sample space of the statistic has to be too; later Fisher explained that in the discrete sample space case there is only a collection of probability inequalities. The properties emphasised are those features of the density that are changed—or not—when the variable is transformed. Fisher described how the mean is 'not invariant': i.e. if x and ξ are random variables such that

$$x = \chi(\xi)$$

then 'we shall not in general find' their means $\bar{x}$ and $\bar{\xi}$ related as

$$\bar{x} = \chi(\bar{\xi}).$$

Nor is the mode invariant. However at last we find something invariant: 'the [percentiles] will be invariant for any transformation for which $d\chi/d\xi$ is always positive'.

All this seems to be preparation for a single point, a new gloss on the old (1921/2) matter of the identity of outcome from maximising the likelihood and maximising the posterior from a uniform prior. In the latter 'two wholly arbitrary elements ... have in fact cancelled each other out, the non-invariant process of taking the mode, and the arbitrary assumption that Ψ [the prior] is constant' (p. 531). Invariance is given no role in the fiducial argument but the abstract Fisher wrote for the British Association presentation—in Appendix A below—suggests a role and closes an important gap in the paper. The following statement is presumably linked to the assumption 'T increases with θ for all possible values':

> The invariant character of the percentile values does, however, make possible certain statements in terms of probability respecting the values of the parameters of populations.

This statement suggests a more connected argument and symmetrical structure than can be detected in the published paper: inverse probability does not respect invariance while the fiducial argument expresses it! The transition from statistic to parameter—treated in a non-confidence way as a random variable—is justified by the invariance of the percentiles and the invariance of the percentiles guarantees invariance to re-parametrisation—the failing of inverse probability. These features of the 'percentile method', the more informative name used in the abstract, are not at all evident in the paper.

The assumption that T is a maximum likelihood value had no clear role in the argument. In his 'theory of estimation' Fisher (1922a & 1925b) promoted maximum likelihood as the solution to 'the problem of estimation'. Later he (1934a, p. 617) objected to Neyman's approach that the statistics involved were not necessarily solutions to the 'problem' and he told Deming (Bennett 1990, p. 82) that 'the theory of fiducial probability is only an outgrowth or branch of the theory of estimation'. There were no similar statements in 1930 but Fisher probably saw no reason to emphasise the appropriateness of maximum likelihood. Whatever Fisher's view in 1934 of its logical status, genetically fiducial theory was an outgrowth of the practice of significance testing and the solution of exact distribution problems.

7 Inverse Probability Again

I have been considering the fiducial argument and how it developed out of earlier work on interval inference—Fisher's and others'. Yet the 1930a paper is called 'Inverse Probability' and mostly it is about inverse probability. It reads like an essay on inverse probability whose natural course has been diverted by the discovery of the fiducial argument.

The paper begins with an outline of the history of inverse probability—filling out the sentence or so in *Statistical Methods*. The first 'preliminary point' (p. 528) is that 'men of the mental calibre of Laplace and Gauss' would not have fallen into the error of accepting it without an 'uncommonly good reason':

> The underlying mental cause is ... [found] in the fact that we learn by experience that

> science has its inductive processes, so that it is naturally thought that such inductions, being uncertain, must be expressible in terms of probability. In fact, the argument runs somewhat as follows: a number of useful but uncertain judgements can be expressed with exactitude in terms of probability; our judgements respecting causes or hypotheses are uncertain, therefore our rational attitude towards them is expressible in terms of probability. The assumption was almost a necessary one seeing that no other mathematical apparatus existed for dealing with uncertainties.

This is part of a long overture to likelihood; it is strange introduction to a way of expressing inductions in terms of probability, to the fiducial argument.

This part of the 1930 paper is in the spirit of the work I discussed in Section 1 yet it is not a compilation of quotations; there are new points and the whole is well-constructed. It looks like a free-standing note or an introduction to a different kind of paper—perhaps one correcting Bayesian misunderstanding of Fisher's work. Such a piece was needed. Sheppard (1929, p. 145) confused maximum likelihood with maximum posterior estimation and Burnside's posthumously published *Theory of Probability* (1928) must have been disappointing. The terminology was none too happy—'likely' often means 'probable'—and, despite Fisher's efforts in correspondence and in print (1923), Burnside reproduced his derivation of the t-distribution via inverse probability without mentioning 'Student'.

There was nothing new about Fisher's attitude to the 'inverse type of argument'. He first 'cloaks its fallacy under a hypothesis' by supposing that the population from which the observations have been drawn 'has itself been drawn at random from a super-population of known specification'. Here is a 'perfectly direct argument' by which the probability that the parameters lie in any assigned limits can be calculated. Fisher moves on to 'inverse probability strictly speaking' which takes a constant prior density for the parameters. He reiterates that this inverse argument is 'devoid of foundation and incapable of consistent application', explaining its survival by the lack of an alternative. The scene set, likelihood and the probability/likelihood division make their expected appearance, to be overwhelmed by the new kind of 'probability statement'.

The paper begins with an account of the ills of the invalid 'inverse probability strictly speaking' argument but ends with a discussion of the valid 'perfectly direct argument' based on sampling from a known super-population. The fiducial argument does what the inverse method always claimed to do—it 'supplies definite information as to the probability of causes' (p. 533). The paper ends with a comparison of two valid forms of probability statement and considers which to accept should they conflict—as they may:

> It would be perfectly possible, for example, to find an *a priori* frequency distribution for ρ such that the inverse probability that ρ is less than .765 when $r = .99$ is not 5 but 10 in 100.

This seems more a theoretical than a practical issue for only rarely will there be a known super-population yet for one concerned with lack of uniqueness of solutions it had to be faced.

The discrepancy is not a contradiction, for the 'logical content' of the two statements is different. Unpacking the inverse probability statement gives (p. 535):

> if we repeatedly selected a population at random, and from each population selected a sample of four pairs of observations, and rejected all cases in which the correlation as estimated from the sample (r) was not exactly .99, then of the remaining cases 10 per cent. would have ρ values less than .765.

Unpacking the fiducial statement gives:

> if we take a number of samples of 4, from the same or from different populations, and

> for each calculate the fiducial 5 per cent. value for ρ, then in 5 per cent. of cases the true value of ρ will be less than the value we have found.

Both statements, fiducial and inverse, are legitimate statements about frequencies; the difference in 'logical content' reflect differences in the sampling situation. Fisher (1933, pp. 347–8) re-states the point but not, I think, ever again.

Fisher ends with a comparative evaluation of the two approaches. He argues that the fiducial probability is to be preferred:

> The fiducial probability is more general and, I think, more useful in practice, for in practice our samples will all give different values, and therefore both different fiducial distributions and different inverse probability distributions. Whereas, however, the fiducial values are expected to be different in every case, and our probability statements are relative to such variability, the inverse probability statement is absolute in form and really means something different for each different sample, unless the observed statistic actually happens to be the same.

In *Statistical Methods and Scientific Inference* Fisher (1956, p. 56) criticised this treatment. He had 'failed' to perceive that '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'. He had a different theory of the fiducial method by then. See Stone (1983) and Zabell (1992) for successive reformulations.

8 The Jewel in the Crown?

Twenty years on, Fisher (1950b, p. **22**.572a) summarised 'Inverse Probability':

> This short paper ... was intended to introduce the notion of 'fiducial probability', and the type of inference which may be expressed in this measure. It opens with a discussion of the difficulties which had arisen from attempts to extend Bayes' theorem to problems in which the essential information on which Bayes' theorem is based is in reality absent, and passes on to relate the new measure to the likelihood function, ..., and to distinguish it from the Bayesian probability *a posteriori.*

The account of the contents is streamlined—to say the least—and, while 'Inverse Probability' *did* introduce the notion of fiducial probability, Fisher's intentions in 1930 seemed fixed more on inverse probability. Lane (1980, p. 148) has called Bayesian inference Fisher's intellectual bête noire and in 1930 Fisher was less concerned with presenting the fiducial argument than with killing the beast or with ensuring it was not resurrected by mistake. After reciting the familiar case against inverse probability Fisher identified the principle embodied in Ezekiel's charts which should not be confused with the old inverse probability; the most important point was that the principle was different. Fisher's second fiducial publication (1933) again criticised Bayesian analysis–Jeffreys's (1932) for σ in the normal distribution. He first used the fiducial argument in a routine case—away from inverse polemics—in his (1934b, pp. 292–3).

Fisher did not rush to spread the fiducial word; the post-fiducial disposition on probability and likelihood does not appear in his 1932 and 1934c pieces criticising the Bayesian views of Haldane (1932) and Jeffreys (1932). He eventually wrote the fiducial argument into his system as a qualification: likelihood is always available for making statements about parameters but sometimes probability statements are available too. In the 1932 edition of *Statistical Methods* (p. 11) the ban on probability for parameters is eased from all cases to 'most cases' and a paragraph, explaining how the probabilities 'established' by the t and z tests are 'free from the objections' to inverse probability, added.

Ezekiel and Hotelling did not fuss about probability intervals but to others the intervals came as a revelation. The fiducial argument led the parade of Irwin's (1931) 'recent advances' and in late 1933 E.S. Pearson was writing to Neyman in Poland 'I am frequently talking about it in my lectures, though I do not use the idea of fiducial probability, only of the limits' (quoted by Reid 1982, p. 111). Pearson's paper with Davies (1934, p. 80) discussed 'fiducial limits' for the standard deviation and the thesis by Garwood (1934), which he supervised, extended fiducial limits to bivariate and discrete situations. Garwood (p. 11) emphasised that the fiducial distribution *cannot* be interpreted as a probability distribution for the parameter; perhaps it was to this work that Neyman was referring when he (1934, p. 562) wrote that the validity of certain of Fisher's statements had been 'formally questioned'. In Clopper & Pearson (1934) the term 'fiducial limits' gives way to Neyman's 'confidence interval'.

Neyman eventually overshadowed all other contributors to interval inference. In the beginning, from 1930 to perhaps 1935, his concern was with a form of inference which could be applied *whatever* the form of prior distribution because this was usually 'unknown'. Neyman (1977, p. 128)—also Reid (p. 128)—dramatise the dropping of the prior in 1935/6 and by 1937 Neyman was expounding confidence intervals without any Bayesian preliminaries. The less deep Clopper & Pearson (1934) and the pre-1930 Wilson (1927) had already done this for particular cases. Neyman contributed a theory of optimal confidence intervals—and publicity. Neyman was the first to see that this was a big topic. In the way Colonel Parker discovered Presley, Neyman discovered the fiducial/confidence argument: he (1934, p. 563n) wrote, 'The solution of the problem ... of confidence intervals has been sought by the greatest minds since ... Bayes'.

Fisher may have thought discovery unnecessary; through the efforts of Ezekiel and others, with his blessing and clarification, the argument was taking its due—modest—place. Fisher became engaged when Neyman started developing the subject in a way that Fisher had not expected—with no regard for *the* theory of estimation—and Bartlett (1936) started questioning Fisher's arguments. Zabell & Stone (1983) describe the intricacies of this engagement, which lasted for the rest of Fisher's life and involved refounding the argument more than once. By the time of *Statistical Methods & Scientific Inference* the argument was, in Zabell's (p. 370) phrase, the 'jewel in the crown' of Fisher's system but in the beginning there had been no sense of the preciousness of the argument—of the fullness of the statistical inference it permitted, of the unusualness of the circumstances in which it could be used or of the need to safeguard it.

Sources

The unpublished sources referred to in the text are files in the Barr–Smith Library of the University of Adelaide and the library of Columbia University. I am very grateful to Susan Woodburn of the University of Adelaide for much help with the Fisher material there. At Columbia Paul Nascimento assisted.

Fisher's published papers are collected in Bennett (1971–4); in the list of references a CP number indicates the volume in which the paper appears. The collection also includes the prefatory notes Fisher wrote for the papers included in the earlier collection, Fisher (1950a).

Acknowledgements

For information I am grateful to P. Armitage, G.A. Barnard and A.C. Darnell. I have had helpful discussions with A.W.F. Edwards (1999) about this paper and about his notes for a forthcoming reissue of 'Inverse Probability'.

References

Aldrich, J. (1997). R.A. Fisher and the Making of Maximum Likelihood 1912–22. *Statistical Science*, **12**, 162–176.

Allan, F.E. (1930). A Percentile Table of the Relation between the True and the Observed Correlation Coefficient from a Sample of 4. *Proceedings of the Cambridge Philosophical Society*, **26**, 536–537

Barnard, G.A. (1995). Pivotal Models and the Fiducial Argument. *International Statistical Review*, **63**, 309–323.

Bartlett, M.S. (1936). The information Available in Small Samples. *Proceedings of the Cambridge Philosophical Society*, **32**, 560–566.

Bartlett, M.S. (1984). J.O. Irwin (1898–1982). *International Statistical Review*, **52**, 109–114.

Bennett, J.H. (Ed.) (1971–4). *Collected Papers of R.A. Fisher*, 5 volumes. Adelaide: Adelaide University Press.

Bennett, J.H. (1990). *Statistical Inference and Analysis: Selected Correspondence of R.A. Fisher*. Oxford: Oxford University Press.

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

Burnside, W. (1928). *Theory of Probability*. Cambridge: Cambridge University Press.

Clopper C.J. & Pearson, E.S. (1934). The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. *Biometrika*, **26**, 404–413.

Darnell, A.C. (1988). Harold Hotelling (1895–1973). *Statistical Science*, **3**, 57–62.

Davies, O.L. & Pearson, E.S. (1934). Methods of Estimating from Samples the Population Standard Deviation *Supplement to the Journal of the Royal Statistical Association*, **1**, 76–93.

Du Val, P. (1931). The British Association Meeting, Bristol 1930. *Mathematical Gazette*, **15**, 292.

Edwards, A.W.F. (1972). *Likelihood: An Account of the Statistical Concept of Likelihood and its Application to Scientific Inference*. Cambridge: Cambridge University Press.

Edwards, A.W.F. (1995). Fiducial Inference and the Fundamental Theorem of Natural Selection. *Biometrics*, **51**, 799–809.

Edwards, A.W.F. (1997). What Did Fisher Mean by 'Inverse Probability' in 1912–22? *Statistical Science*, **12**, 177–184.

Edwards, A.W.F. (1999). The Origin of the Concept of Confidence Limits, Chapter 5 of the forthcoming volume edited by H. A. David & A. W. F. Edwards *Annotated Readings in the History of Statistics*. New York: Springer.

Ezekiel, M. (1929). The Application of the Theory of Error to Multiple and Curvilinear Correlation. *Journal of the American Statistical Association, Supplement: Proceedings of the American Statistical Association*, **24**, 99–104.

Ezekiel, M. (1930). *Methods of Correlation Analysis*. New York: Wiley.

Ezekiel, M. (1941). *Methods of Correlation Analysis*, 2nd edition. New York: Wiley.

Fisher, R.A. (1912). On an Absolute Criterion for Fitting Frequency Curves. *Messenger of Mathematics*, **41**, 155–160. CP1.

Fisher, R.A. (1921). On the 'Probable Error' of a Coefficient of Correlation Deduced from a Small Sample. *Metron*, **1**, 3–32. CP1.

Fisher, R.A. (1922a). On the Mathematical Foundations of Theoretical Statistics. *Philosophical Transactions of the Royal Society A*, **222**, 309–368. CP1.

Fisher, R.A. (1922b). The Goodness of Fit of Regression Formulae, and the Distribution of Regression Coefficients. *Journal of the Royal Statistical Society*, **85**, 597–612. CP1.

Fisher, R.A. (1923). Note on Dr. Burnside's Recent Paper on Errors of Observation. *Proceedings of the Cambridge Philosophical Society*, **21**, 655–658. CP1.

Fisher, R.A. (1925a). *Statistical Methods for Research Workers*, 1st edition 1925; 2nd edition 1928; 4th edition 1932; 6th edition 1936. Edinburgh: Oliver & Boyd.

Fisher, R.A. (1925b). Theory of Statistical Estimation. *Proceedings of the Cambridge Philosophical Society*, **22**, 700–725. CP2.

Fisher, R.A. (1928). The General Sampling Distribution of the Multiple Correlation Coefficient. *Proceedings of the Royal Society A*, **121**, 654–673. CP2.

Fisher, R.A. (1930a). Inverse Probability. *Proceedings of the Cambridge Philosophical Society*, **26**, 528–535. CP2.

Fisher, R.A. (1930b). Inverse Probability. Abstract published in Report of the British Association for the Advancement of Science, Bristol, p. 302.

Fisher, R.A. (1932). Inverse Probability and the Use of Likelihood. *Proceedings of the Cambridge Philosophical Society*, **28**, 257–261. CP3.

Fisher, R.A. (1933). The Concepts of Inverse Probability and Fiducial Probability Referring to Unknown Parameters. *Proceedings of the Royal Society A*, **139**, 343–348. CP3.

Fisher, R.A. (1934a). Contribution to the Discussion of Neyman (1934). *Journal of the Royal Statistical Society*, **97**, 614–619.

Fisher, R.A. (1934b). Two New Properties of Mathematical Likelihood. *Proceedings of the Royal Society A*, **144**, 285–307. CP3.

Fisher, R.A. (1934c). Probability, Likelihood and the Quantity of Information in the Logic of Uncertain Inference. *Proceedings of the Royal Society A*, **146**, 1–8. CP3.

Fisher, R.A. (1935a). *The Design of Experiments*. Edinburgh: Oliver & Boyd.

Fisher, R.A. (1935b). The Fiducial Argument in Statistical Inference. *Annals of Eugenics*, **6**, 391–398. CP3.

Fisher, R.A. (1936). Uncertain Inference. *Proceedings of the American Academy of Arts and Science*, **71**, 245–258, CP3.

Fisher, R.A. (1950a). *Contributions to Mathematical Statistics*. New York: Wiley.

Fisher, R.A. (1950b). Note to 'Inverse Probability' in Fisher (1950a, p. **22**.572a). CP2.

Fisher, R.A. (1956). *Statistical Methods and Scientific Inference*. Edinburgh: Oliver & Boyd.

Fox, K.A. (1989). Agricultural Economists in the Econometric Revolution: Institutional Background and Leading Figures. *Oxford Economic Papers*, **41**, 53–70.

Garwood, F. (1934). *Some Applications in Statistics of Fiducial Probability*, unpublished London University Ph.D. dissertation.

Hacking, I. (1980). The Theory of Probable Inference: Neyman, Peirce and Braithwaite. In *Science, Belief and Behaviour*, Ed. D.H. Mellor, pp. 141. Cambridge University Press.

Haldane, J.B.S. (1932). A Note on Inverse Probability. *Proceedings of the Cambridge Philosophical Society*, **28**, 55–61.

Hotelling, H. (1927a). Differential Equations Subject to Error, and Population Estimates. *Journal of the American Statistical Association*, **22**, 283–314.
Hotelling, H. (1927b). Review of R.A. Fisher's *Statistical Methods for Research Workers*. *Journal of the American Statistical Association*, **22**, 411–412.
Hotelling, H. (1931a). Recent Improvements in Statistical Inference. *Journal of the American Statistical Association*, **26**, 79–89.
Hotelling, H. (1931b). The Generalization of Student's Ratio. *Annals of Mathematical Statistics*, **2**, 360–378.
Hotelling, H. (1953). New Light on the Correlation Coefficient and its Transforms. *Journal of the Royal Statistical Society B*, **15**, 193–232.
Irwin, J.O. (1931). Recent Advances in Mathematical Statistics. *Journal of the Royal Statistical Society*, **94**, 568–578.
Irwin, J.O. (1953). Contribution to the Discussion of Hotelling (1953). *Journal of the Royal Statistical Society B*, **15**, 223.
Jeffreys, H. (1932). On the Theory of Errors and Least Squares. *Proceedings of the Royal Society A*, **138**, 48–55.
Lane, D.A. (1980). Fisher, Jeffreys and the Nature of Probability. In *R.A. Fisher: An Appreciation*, Eds. S.E. Fienberg & D.V. Hinkley, pp. 148–160.New York: Springer.
Maskell, E.J. (1929). Experimental Error: A Survey of Recent Advances in Statistical Method (continued). *Journal of Tropical Agriculture*, **6**, 5–11.
Maskell, E.J. (1930). Field Experiments. *Journal of Tropical Agriculture*, **7**, 101–104.
Morgan, M.S. (1990). *The History of Econometric Ideas*. Cambridge: Cambridge University Press.
Neyman, J. (1934). On the Two Different Aspects of the Representative Method (with discussion). *Journal of the Royal Statistical Society*, **97**, 558–625.
Neyman, J. (1937). Outline of a Theory of Statistical Estimation based on the Classical Theory of Probability. *Philosophical Transactions of the Royal Society A*, **236**, 333–380.
Neyman, J. (1941). Fiducial Argument and the Theory of Confidence Intervals. *Biometrika*, **32**, 128–150.
Neyman, J. (1952). *Lectures and Conferences on Mathematical Statistics and Probability*, second edition. Washington: Graduate School USDA.
Neyman, J. (1977). Frequentist Probability and Frequentist Statistics. *Synthese*, **36**, 97–131.
Reid, C. (1982). *Neyman–from Life*, New York: Springer.
Sheppard, W.F. (1929). The Fit of a Formula for Discrepant Observations. *Philosophical Transactions of the Royal Society A*, **228**, 115–150.
Smith, W.L. (1978). Harold Hotelling 1895–1973. *Annals of Statistics*, **6**, 1173–1183.
Stone, M. (1983). Fiducial Probability. In *Encyclopedia of Statistical Science*, volume 3, Eds. S. Kotz & N.L. Johnson, pp. 81–85. New York: Wiley.
'Student' (1908). The Probable Error of a Mean. *Biometrika*, **6**, 1–25.
Wilson, E.B. (1927). Probable Inference, the Law of Succession and Statistical Inference. *Journal of the American Statistical Association*, **22**, 209–212.
Wilson, E.B. (1929). Probable Error of Correlation Results. *Journal of the American Statistical Association, Supplement: Proceedings of the American Statistical Association*, **24**, 90–93.
Working, H. & Hotelling, H. (1929). Applications of the Theory of Error to the Interpretation of Trends. *Journal of the American Statistical Association, Supplement: Proceedings of the American Statistical Association*, **24**, 73–85.
Zabell, S. (1992). R.A. Fisher and the Fiducial Argument. *Statistical Science*, **7**, 369–387.

Résumé

En 1930 R. A. Fisher a présenté l'argument fiduciare. Je discute l'argument et ses origines, soulignant le rôle de Mordecai Ezekiel dans le développement du 1930.

Appendix A: the Abstract of 'Inverse Probability'

The abstract of 'Inverse Probability' (1930b, p. 302) in the report of the British Assocation meeting, Bristol September 1930, seems not to have been noticed before and I reproduce it here for convenience.

> Dr. R. A. FISHER—Inverse Probability.
> The controversy over 'inverse probability' seems to be unique in the history of mathematics. The reasons for the rejection of the classical theory are obvious and need only to be stated. Its retention in mathematical text-books is to be explained by the fact that until recently no alternative method was available to give an account of inductive reasoning.
>
> The method of maximum likelihood has no logical connection with inverse probability, although it has been associated with it historically. Its derivation by this path involves the introduction of arbitrary functions at two distinct stages, which can be made

> to cancel each other. Likelihood is not a synonym for probability; it is a quantity, which, like probability, measures the degree of rational belief, but it does not obey the laws of probability. Statements about unknown samples of known populations are made in terms of probability, statements about the unknown populations from which known samples are drawn are made in terms of likelihood. Likelihood serves all the purposes necessary for the problem of statistical estimation.
>
> The invariant character of the percentile values does, however, make possible certain statements in terms of probability respecting the values of the parameters of populations. Statements of this type, which have very strangely been overlooked, are available only when the observations are of quantitative variates, and not merely of frequencies. They differ from the statements of inverse probability, both numerically and logically; the statements of inverse probability are absolute in the form, based on a hypothetical super-population of an absolute character, but can never be verified, for any further samples from the same population will alter the content of the statements. The statements of the percentile method are relative in form and rigorously demonstrable without any assumption as to the *a priori* distribution of the parameters.

In the Mathematical Gazette Du Val (1931, p. 292) reported Fisher's contribution to the session as follows: 'Dr Fisher expounded the somewhat neglected distinction between probability and likelihood, in the course of a paper on inverse probability'.

Appendix B: Ezekiel & Fisher after 'Inverse Probability'

'Your article interested me very much' wrote Ezekiel after visiting Fisher in October 1930. His letter included a 'brief and rough restatement of the main thesis, indicating the main drift of your argument, so far as I can grasp the mathematics you use.' The last paragraph of the restatement reads:

> We therefore take the 'fiduciary' value as the value which probably exists in the universe from which the specified sample r was obtained, with but one chance in 20 of the true ρ in the universe being below the fiduciary value.

Fisher had been satisfied with the exposition up to this point but he corrected the economist's slip, 'fiduciary', and rewrote the paragraph:

> We therefore take the fiducial 5 per cent. value as the parametric value (ρ) for which the probability of exceeding the observed statistic (r) is 5 per cent. The corresponding fiducial values for other probabilities jointly specify the fiducial distribution of ρ for a given value of r.

Ezekiel's words are those he had used in his book yet in the spirit too of the article's 'definite information as to the probability of causes'. Fisher's re-write gives very little away for it scarcely goes beyond defining terms.

A second edition of *Methods of Correlation Analysis* appeared in 1941. The probability statements were rewritten in confidence terms to 'bring them up' to what Ezekiel (p. v) called 'the modern interpretation'. The statement for the regression coefficient—quoted in Section 7 above—thus became (p. 315):

> if we say that the true value lies between 0.14 and 0.86, we are making a statement of the sort which is likely to be wrong only once or twice out of each hundred such statements . . . .

The book went to a third edition in 1959, with K. A. Fox as co-author, and stayed in print until 1987.

In 1958 Ezekiel wrote to Fisher from Rome on FAO business: his greeting was, 'It is now many years since my visit with you at Rothamsted, and since you helped me subsequently in adapting your 1928 paper to the needs of my book on *Methods of Correlation Analysis*.' After thirty years Ezekiel had confused the sequence of events.

[*Received July 1999, accepted February 2000*]