Journal of the Statistical and Social Inquiry Society of Ireland.
Vol XXVI, Part III

R.A. FISHER AND THE DEVELOPMENT OF STATISTICS

- A VIEW IN HIS CENTENARY YEAR

DENIS CONNIFFE
Economic and Social Research Institute

(Read before the Society, 15 November 1990)

---

1. INTRODUCTION

It is certainly appropriate that The Statistical and Social Inquiry Society of Ireland should present a paper about the great statistician Fisher in this year, the centenary of his birth. I am very pleased to be giving the paper, but I must admit that writing it posed some problems. There have already been many other commemorations of Fisher this year: by statistical societies; institutions he worked in and various journals. I do not want to merely repeat what has been said elsewhere and what some of the audience may already know.

One approach would have been to concentrate on that portion of Fisher's work of most relevance to the themes arising in the meetings of this Society. But it is probably fair to say that the bulk of these themes are associated with issues in Irish economic policy and the provision and interpretation of official statistics. Although Fisher had interests in a huge range of applications of statistics he did not contribute directly to either econometrics or official statistics. This is not to say he had no influence - the average textbook on econometrics contains more mentions of maximum likelihood than do most statistics texts and books on the methodology of sample surveys draw heavily on his ideas and the formulations of those, like Yates and Cochran, who were very much his disciples. But while his influence ought to be mentioned - and I will be giving some examples of how Fisherian ideas keep popping up in econometrics (acknowledged or not) - to confine a paper to such topics would not do any justice to Fisher's thought.

Another way of differentiating this paper from others would have been to stress the Irish connections. There were the early communications between Fisher and "Student" (W.S. Gossett) of Guinness's brewery in 1912 and their later collaborations through the twenties led to important findings. Fisher stayed with Student when he visited Dublin in 1932, but eventually they quarrelled over the current approaches to certain problems in agricultural experimentation. There was also Fisher's influence on Roy Geary, who had such an impact on the Society and on statistics in Ireland generally. The connections do not end there. For example, the statistics department set up in the Agricultural Institute in 1959 was greatly influenced by, and in a miniature way, modelled upon, the Rothamsted statistics department where Fisher had worked.

However, some at least of these topics have been covered elsewhere. Student's interactions with Fisher and the Pearsons have been described by various historians of statistics and recent accounts include Boland (1984) and Plackett and Barnard (1990). Roy Geary has given his own account (Geary, 1983) of his dealings with Fisher and Spencer (1976, 1983) has analysed Geary's contributions to advances in methodology, so permitting assessment of Fisher's influence. I think the Irish connection in statistical developments is worth remembering and I have spoken on the theme myself (Conniffe, 1989) and will mention a few points again tonight. But a whole paper on this theme would not be justifiable.

I think I have to look at the broad sweep of Fisher's work, even if this is what has been done at other venues this year, often by very distinguished statisticians. But I still have scope to be different. Not all of Fisher's innovations met with unanimous approval and there are still issues that remain unsettled. Perhaps too, some fairly widely accepted methodologies due to Fisher are not beyond being questioned. Fisher contributed so much, so widely, that aspects of his work can be criticised, or reassessed, without taking from his preeminence. He himself might not have agreed. He took offence easily and could express himself most unpleasantly about people he perceived as obstructive, or pretentious, or plagiarists of his own work.

The audience at meetings of The Statistical and Social Inquiry Society are not usually very keen on detailed mathematical notation, or manipulation, and I can remember a few mathematical papers with embarrassingly small attendances. So I have been telling people that my presentation will not be mathematical and that what I have to say will be easily comprehended and even entertaining. I hope I can deliver on these promises without sacrificing content or trivialising issues. I will be supplementing the text by overheads at points where readers without much knowledge of statistical theory might find explanations insufficiently detailed. However, I should begin with some biographical material about Fisher that will not present any difficulties of that nature.

2. EDUCATION AND EARLY CAREER

Ronald Aylmer Fisher was born on 17th February 1890 into a well-to-do upper middle class family. His father was a partner in a prominent auctioneering firm (Robinson and Fisher) of St. James, London. He was educated at Harrow and his consciousness of being a member of an elite group stayed with him throughout a relatively impecunious period that followed the failure of his father's business. He obtained a scholarship to Gonville and Caius College, Cambridge; entering in 1909, graduating in mathematics and spending another year studying mathematical physics, specialising in quantum theory and statistical mechanics. He had actually hesitated between mathematics and biology for University studies. The story is that he then saw a cod's skull in a museum with its bones separated out and labelled; so he decided on mathematics. But he retained his interest in biology and was especially interested in Darwinian evolution and in the implications of Mendel's findings which were rediscovered about the turn of the century. He gave a talk to an undergraduate society in 1911 on "Mendelism and Biometry". He was interested in the implications of selection and heredity for humankind, as indeed were many of the established statisticians and biologists of that era. His views on eugenics interacted with and fortified his beliefs about the reality of the existence of an elite class to which he belonged.

Poor eyesight kept Fisher out of the army in WW1 and he spent from 1915 to 1919 teaching mathematics and science in various public schools. He married Ruth Guinness, a doctor's daughter, in 1917 and set about increasing the numbers of the elite. This may sound as if I am being cheaply facetious, but Fisher held very strong views based on his interpretation of genetics and his perception of British society. He believed that society was degenerating because the educated, intelligent and morally sound elite were having too few children, while the unintelligent and morally unsound were having too many. Fisher believed that not only physical qualities were heritable, but so also were socially useful traits.

> *Man's nature is not less governed by heredity than that of the rest of the animate world .... (heritabilities) of the same magnitude were obtained for the mental and moral qualities in man as for the physical measurements.*

The quotation is from Fisher (1925a) and many similar quotes could be made to other books and papers of his, especially Fisher (1930), which I will mention again later. His views are collated and discussed in his daughter's biography of him (Box, 1978). The problem, in Fisher's view, was that the cost of rearing and educating a child of the elite to the point where the child could compete successfully, had become so great that the elite were limiting the numbers of their children. But the non-elite, who did not intend to provide for their children similarly, were breeding away. Fisher believed that society, for its own good, should encourage the greater production of children by its more able members through such incentives as family subsidies that increased with income. He wrote a paper on this theme (Fisher, 1931) titled "The biological effects of family allowances".

This idea that higher income groups should be credited with incurring higher costs for their children than lower income groups is alive and well today. "Poverty" surveyors normally calculate their measure of average household income (which they compare with a poverty threshold income) by dividing actual household income by a measure of household size, usually by an "equivalence scale" of counting a child as half an adult, or something similar. This of course means that if two households have the same family composition, but one has ten times the income of the other, the children of the first are presumed to "cost" ten times as much. It is therefore also possible to show that large families with a high income are in "poverty" when small families with a lower income are not. The idea that if high income families spend above the norm on their children it is because they choose to do so and consequently derive more satisfaction than they would from alternative consumption, seems as unacceptable to poverty researchers as it was to Fisher. For an alternative approach to equivalence scales, see Conniffe and Keogh (1988).

I'm afraid I've begun digressing already, so I'd better return to Fisher. While teaching, he was continuing his statistical and genetical researches. He had published his first paper [Fisher (1912)] as an undergraduate and it dealt with maximum likelihood estimation, although he did not use that term then. Fisher (1915) gave the exact distribution of the sample correlation coefficient, which was quite complicated in the non-zero population correlation case, and three years later (Fisher, 1918) his study on genetic correlations based on Mendelian inheritance was published, partly at his own expense. He had previously submitted this to the Royal Society of London, but referees turned it down. One referee was the famous statistician Karl Pearson and it is suspected that Fisher eventually learned this, which may have laid the foundation for the animosity between the two.

It is anticipating a later section a little, but it is as well to say here that Fisher soon had his revenge. Pearson had considered Fisher's exact distribution for the correlation coefficient so important that he set the full resources of his statistical laboratory at University College London working for years tabulating or approximating it. But Fisher (1921) showed that the transformation

$$z = tanh\ r$$

converts the sample correlation r to a near normal variable z and made all Pearson's work on the topic redundant. One of the innovations of which Pearson was proudest (Pearson, 1906) had been the family of distributions bearing his name that he fitted by the method of moments: that is, equating the sample mean to the population mean, the sample second moment to the population second moment, etc., and solving for the unknown parameters. But Fisher (1922) showed the method of moments was highly inefficient compared to his own maximum likelihood.

In 1919 Fisher was appointed to a newly created post of statistician at

Rothamsted agricultural research station at Harpenden, outside London. He worked here until 1933 and his flood of publications revolutionised the subject of statistics. Most of the statistical establishment in Britain resisted this revolution for as long as they could. To appreciate where Fisher was starting from some account of pre-Fisherian statistics and statisticians is desirable.

3. PRE-FISHERIAN STATISTICS AND STATISTICIANS

Some statistical methods, associated with the normal distribution and with what would now be called least-squares, go back to Laplace and beyond, but most of the statistical procedures in use before Fisher's time had been developed by a group of extraordinarily energetic late Victorian polymaths, Galton, Edgeworth, Pearson, Weldon and Yule. In some ways, these were larger than life figures.

Galton was a first cousin of Charles Darwin and as a young man was an enthusiastic traveller. He explored remote regions of Africa and received a gold medal from the Royal Geographic Society for his discoveries. Eventually, adventures of another sort led to him to a period of convalescence in England where he took up scientific and statistical investigations. He developed many ideas other than in the fields of statistics. For example, he invented the system of identification by fingerprinting. But in statistics he is associated with the development and interpretation of correlation and regression methods which he applied to studies of heredity and biology. Galton founded the journal *Biometrika* and endowed a chair and laboratory of statistics at London University.

Edgeworth was another remarkably versatile man. He took a degree in classical literature, but later qualified as a barrister. Then he took up economics and mathematics. Unlike Galton he was not specially interested in biology (although he is supposed to have conducted much measurement of bees entering and leaving hives on the family estate in Longford). He applied Galton's methods in the social sciences, but also contributed greatly to general methodology. He formulated the multivariate normal distribution, developed many of the large sample statistical tests that (with somewhat different notations) are still in textbooks today, and studied the limiting distributions of sample means to higher than first order.

Karl Pearson was 36 when he wrote his first paper on statistics in 1893, but he had already over 100 publications to his name including nine books. His first subjects included religion (he wrote lives of Martin Luther and Spinoza), German folklore and history. Later he wrote about the education of women, sex and socialism. He became very interested in evolution, genetics and statistics through reading Galton's (1889) book on heredity. He obtained the Galton chair in London University and was occupying that position, as well as being editor of *Biometrika*, during Fisher's early career. He worked widely in statistics and biometry, modifying the research of his predecessors as well as making his own original contributions, including the famous chi-squared test. He played a prominent role in many scientific debates and controversies including a famous dispute on the validity of Mendelian inheritance , and another on alcoholism and inherited effects. This latter dispute arose from a paper by Elderton and Pearson (1910), but led to a two year debate drawing in many famous, or to-be-famous, names including Marshall and Keynes. By the time of his death in 1936, Pearson's count of publications had reached 648. He had immense prestige during the years Fisher was involved in disputes with him.

I'll skip Weldon and Yule, because I'll run out of time if I do not. But I must say something about Student. W.S. Gosset was a brewer in the Guinness firm in Dublin from 1899 until 1935, when he moved to London to take charge of the new brewery at Park Royal. Because Guinness had a policy of recruiting scientists as "brewers" (senior posts) and encouraged research, quite a few ideas about the design and analysis of experimental data resulted, and not only from Gosset alone. However, for the present the important point is that Gosset often worked with small samples and was therefore bothered about the large sample (normal approximation) methods used by Pearson and all others. Edgeworth (1905) had extended normal approximations to higher order so that, for example, instead of having to treat terms in n^{-1} as negligible, they could be retained and terms in $n^{-3/2}$ treated as negligible, but this is no help with truly small samples.

The influence these men had on Fisher was very substantial, but cannot be entirely reliably assessed from Fisher's own writings. The foreword of Fisher (1956) makes clear his high opinion of Galton's contributions and he quotes a lot from Galton's analyses of some of Darwin's data in Fisher (1935). Since Edgeworth had developed much of the mathematical statistical equipment in use before Fisher's time, one might expect to trace clear influences. Unfortunately, Bowley (1935) made the nasty suggestion that Fisher had "borrowed" results from Edgeworth (1908) without acknowledgement and this was picked up and repeated by Fisher's enemies (e.g., Neyman, 1956). So Fisher's treatment of Edgeworth in his later writings may have understandably been constrained. His treatment of Pearson was hostile, of course; in 1956 he described Pearson's huge output of publications as "pretentious and erratic" and the man himself as suffering from "incapacity in self-criticism and willingness to admit the possibility that he had anything to learn from others". But there is evidence from letters reproduced in Plackett and Barnard (1990) that Fisher had felt differently initially and Kendall (1963) believed that Fisher was influenced by Pearson's books and papers.

As regards Student, the positive influences on Fisher are in no doubt. Some of his earliest published papers were extensions or more mathematically rigorous demonstrations of Student's results. In Chapter 1 of his 1925 book he wrote

> *One of the chief purposes of this book is to make better known the effect of his (Student's) researches and of mathematical work consequent upon them.*

He repeated comments on the significance of Student in his later books, although with decreasing emphasis.

I think another person who merits mention for his influence on Fisher is Keynes. This may seem surprising since Keynes (1921) argued vehemently against frequentist induction and ideas of probability and Fisher is often considered as the greatest of the frequentist statisticians. But actually frequentist and inverse probability (early Bayesianism) co-existed and many statisticians including Pearson, Student and even Edgeworth behaved like frequentists in analysing data, especially when testing hypotheses, but resorted to inverse probability when debating the theoretical basis of induction. Fisher himself used (I believe) the inverse probability framework in his first 1912 paper on the method he later called maximum likelihood. Most writers (e.g., Plackett and Barnard, 1990) suggest that Fisher's use of terms like "the most probable value" were mistakes in wording, but I'm not at all sure of that. I think Fisher had been very influenced by Keynes (1911), which was written in an inverse probability context, but is indexed in Kendall and Stuart (1967) as "characterisations of distribution by forms of ML estimators". I'm also not sure that the usual interpretations of early letters between Student and Fisher, where effectively ML estimation was being discussed, are correct. However, this is a peripheral point, which I'll only expand on if I have time.

Keynes influenced Fisher in other ways too, I think. The 1921 book also put up arguments against the widespread use of the Principle of Indifference (Bayes's Postulate) and very similar arguments appeared in later works by Fisher. It may be just coincidence, but Fisher (1921) was his first clear statement of disagreement with the inverse probability method. Keynes also believed in several different kinds of probabilities (only one of which would nowadays be termed such) that shared the property of being measures of rational, objective, degrees of belief about parameters, but differed in the extent to which they were numerically, or probabalistically, interpretable. Fisher also proposed a variety of such measures ranging through fiducial probability, likelihood ratios, and significance test sizes. There were possibly very analogous ideas held by both men. I say "possibly" because neither Fisher nor Keynes practised writing styles that were conducive to unambiguous interpretation.

4. FISHER'S CONTRIBUTIONS TO STATISTICS AND RELATED DISCIPLINES

I left my account of Fisher's life with his joining Rothamsted in 1919 and much of the research to be described in this section was conducted there. However, for some themes it will be convenient to include research he conducted later at other venues. I will return to a more biographical account in the next section.

Genetics

Statisticians often forget that Fisher was an important figure in the history of genetics. In fact, during his lifetime he published 140 papers on genetics compared to 129 on statistics. Of course, the dividing line is not absolute and many important statistical ideas first appeared in his genetics papers. The term "variance" was first used in his 1918 paper on the genetic correlation between relatives and maximum likelihood was first applied in genetics. Some subsequently famous statisticians commenced as geneticists either working directly with Fisher, or on his ideas. C.R. Rao is probably the best known. Does his name mean much in econometrics? It ought to, since he developed the ubiquitous score test (called Lagrange Multiplier tests by econometricians) and a test for parallelism of regression lines (called the "Chow" test by econometricians). Arguably though, both tests go back to Fisher, since the score test is virtually implied by Fisher's definition of the score vector and the "Chow" test is a rather obvious application of the F test.

Fisher's initial genetical work was on reconciling Mendelism and Darwinism. It had been believed that Mendelian genetics contradicted Darwin's basic idea that continually occurring small changes provide the engine of evolution. Fisher showed that in fact Mendelism provided a sounder explanation of natural selection that did Darwin's own notion of the environment stimulating production of new variation. Fisher's later work essentially created the whole subject of population, or quantitative, genetics, which has been immensely important in biology and agriculture. One statistical spin-off was the development of the theory of stochastic processes, which Fisher had required to explain gene survival and drift.

The study of human genetics, or eugenics, was of special interest to Fisher. What was probably his greatest direct contribution to human welfare came from his work (with colleagues) on blood groups, which was motivated by a desire to find common "marker" genes for use in human linkage studies. The result was the identification of "The Rhesus Factor" - which has led to saving hundreds of thousands of lives since then. Other work gave further statistical spin-offs. When he was trying to improve on a "coefficient of racial likeness", he had multiple measurements on a set of skulls and he devised the discriminant function, so creating a new branch of statistics.

Yet some of his work on human genetics and publications, especially in the *Annals of Eugenics*, can leave some readers uneasy. Not because of the quality of the analysis, but because of the world-view that might be thought to underlie it. His (1930) book on natural selection contains valuable scientific work, but the chapters dealing with human heredity reveal some of his opinions that I have already mentioned. Human fertility is shown to be a heritable characteristic and the existence of a negative correlation between fertility and human ability is argued. The reversal of this correlation is claimed to be the great problem facing society. Greater family allowances for the better off appear again.

In Rothamsted, Fisher was supposed to be, first and foremost, a statistician assisting the agricultural researchers. Anyone who has worked in a statistical service capacity in a research institute, knows how demanding on time that role can be. Fisher's genetic researches, as well as much of his statistical methodological research, was largely conducted in his "spare" time. Actually, much credit must be due to his wife. Fisher conducted genetic experiments in his home with mice and other small animals and a lot of the physical work associated with this fell to his wife. Because of his poor eyesight she also took down, in longhand, the scrips for papers and reports, including his first book Statistical Methods for Research Workers. All this and eight children! Also, Rothamsted did not pay statisticians particularly well, so perhaps Fisher's views on family allowances were related to personal circumstances.

Design of Experiments and Analysis of Variance

There were various ideas around on how agricultural field and laboratory experiments should be conducted before Fisher commenced work at Rothamsted. But it was Fisher who sorted out the issues and turned experimental design into a rational methodology. It is convenient to discuss experimental design in two parts: the allocation of treatments to experimental units and the structure of treatments.

The response or "yield" of an experimental unit, be it a plot of land or an animal, depends on many factors besides the experimental treatment it receives. Some of these factors may be identifiable and potentially controllable and others may be unknown. Fisher's idea was to control the identifiable factors in the design and to randomise over the others. So the randomised blocks design, or Latin square designs, could control for fertility gradients in field experiments or for varying initial weights or genetic effects in animal experiments. Some of these ideas had been around before and it could be argued that Fisher (1935b) may even have drawn a little on experimental expertise developed in experimentation in Ireland. Student (1911) had written

If comparing two varieties .. arrange plots .. so that yields of both shall be affected as far as possible by the same causes ...

Compare plots that lie side by side ...

But Fisher systematized a methodology and faced all the complications when the number of treatments was so large that complete blocks became excessively heterogeneous. He investigated the combinatorial properties of balanced incomplete block designs and suggested the confounding of certain treatment differences with blocks. The first such experiment was conducted in Rothamsted in 1927.

The idea of randomisation was, and to a degree still is, controversial. Some statisticians (the results of an opinion poll on the subject accompany the article by Barnard, 1990b) consider it the greatest of Fisher's contributions, a view that I'd feel is a little overstated. The arguments in favour of randomisation are that it guarantees the assumptions required for the validity of analysis of variance of the experimental results. In econometrics much effort is devoted to checking 'standard assumptions', uncorrelated errors, homoscedasticity etc. Randomisation has the effect of mimicking the second moment structure of an appropriate normal theory model. Indeed, hypothesis testing can, in principle, be based directly on the permutations of the possible treatment-to-plot allocations.

However, it has been argued that experimenters familiar with their materials could chose assignments that were more trustworthy than those resulting from randomisation. What was one to do if randomisation produced an apparently very systematic scheme? It was also argued that more efficient analyses could be possible taking account of the correlations between plots rather than removing them by randomisation. Fisher dealt fairly roughly with his early critics about this and even fell out with Student. Although Fisher turned randomisation into the 'done thing', a degree of scepticism has persisted through Papadakis (1937), Atkinson (1969) and Bartlett (1978).

This alternative analysis involves specifying the actual spatial correlation of plots.

For example, if a crop experiment comparing levels of a fertiliser consisted of a series of plots laid out along a drill in the form

Plot 1	Plot 2	Plot 3	Plot 4	. . .

the usual model for yields would be written

$$y_2 = a + bx_i + e_i$$

where x_i is the fertiliser level applied to the ith plot. Now instead of, or as well as, randomising levels of fertilisers to plots, one might model the actual correlation between adjacent plots as

$$e_i = \frac{1}{2}\delta(e_{i-1} + e_{i+1}) + u_i$$

where the u_i are truly independent. For a rectangular array,

Plot 1	Plot 2	Plot 3
Plot 4	Plot 5	Plot 6
Plot 7	Plot 8	Plot 9
.	.	.
.	.	.

a possible error model would be

$$e_5 = \frac{1}{2}\delta_1(e_4 + e_6) + \frac{1}{2}\delta_2(e_2 + e_8) + u_5$$

etc. Actual estimation might be by first obtaining residuals from an 'ordinary' analysis and then estimating the δ's.

Clearly the procedure is rather similar to what econometricians do in a Cochran-Orcutt analysis or in a maximum-likelihood solution of a correlated errors case. It is interesting that Student (1914) proposed differencing experimental data by analogy with the treatment of time series. Fashionable econometricians difference a lot nowadays to counter suspected non-stationarity and possibly younger econometricians think 'spurious correlations' started with Granger and Newbold (1974). Actually, econometricians and statisticians were always aware of the issue and Student wrote a paper titled "The elimination of spurious correlation due to position in time or space". He quoted an example of correlation between female cancer death rate and imports of apples. Other exponents of differencing included Cave-Brown-Cave (1904), Hooker (1905) and Anderson (1914).

Returning to Fisher; his sometimes bitter attacks on opponents of randomisation are hard to reconcile with his own willingness to use arguments that were sometimes conditioned on observed, or "ancillary", statistics. I will give an example later under the heading of statistical inference, where he seems quite willing to abandon the frequentist type randomisation justification. Fisher could see clever ideas and implement them, but he did not seem to worry too much whether they were mutually consistent. Perhaps he was right not to slow himself, or the development of statistics, through trying to reconcile everything, but he did, I think, try to give the impression that everything worked out at some higher plane that he, but few others, could perceive.

The other great development in design of experiments was the introduction of factorial treatment structure. Instead of the former practice of testing one factor at a time, a treatment structure was adopted that enabled the researchers to see if factors interacted with each other, and that provided hidden replication if they did not. As Fisher put it "Nature will respond to a logical and carefully thought out questionnaire; indeed, if we ask her a single question, she will often refuse to answer until some other is asked". As in the case of randomisation, Fisher's arguments quickly made factorial design the norm. His successors at Rothamsted, who developed fractional replication, permitted ever larger numbers of factors to be investigated. These designs are once again extremely topical, as anyone here who might be familiar with "Taguchi Methods" (Taguchi, 1976, 1977) will well know.

Fisher had to fight against a statistical establishment to get his ideas accepted, but when they were accepted they gained a remarkably strong grip on the minds of a new establishment and on biological researchers. In fact, it became difficult to suggest that there might be times when a factorial design laid out in randomised blocks was not necessarily the only thing to do. I'll give two examples. When I started working as a statistician in An Foras Taluntais in the early sixties I repeatedly came across pseudo-factorials of the following type. Several forms of a nitrogenous fertiliser, say, urea, calcium ammonium nitrate etc. are applied each at, say, three levels; on the face of it a typical factorial. But the lowest level is a zero level. Unless form of nitrogen does not matter, such a design will produce a huge form x level interaction, which will hide any true interaction and make the interpretation of main effects very dubious. There is also the purely technical point that the standard factorial "degrees of freedom" of the analysis of variance table are incorrect. With great ingenuity such a design can be analysed in a factorial context without drawing false inferences, but it would be much more sensible not to formulate the design as a factorial to start with. In those days I thought that these designs resulted from "do-it-yourself" experimental design by researchers who imitated, perhaps with little depth of understanding, the experimental methods they had seen elsewhere. But last February I was at the Fisher commemoration at Rothamsted and learned that this was how such experiments were done there, and all that worried the statisticians was how to adjust the degrees of freedom!

My second example relates to grazing experiments. These took the form of treatments of grassland (often fertilisers) and herds of animals grazed at sets of stocking rates (areas per animal). A moment's thought indicates that if one treatment grows more grass than others, it can carry more animals per unit area before individual animal performance is depressed through food constraint. There will be an "optimum" stocking rate corresponding to each treatment and it is at this point that meaningful comparisons can be made. Clearly then, a standard factorial with the same set of stocking rates on each treatment would be a poor design, unless the treatments were not expected to differ much. Otherwise, a different design and appropriate analysis is very desirable. Developing these posed to be rather controversial and even after various publications (Conniffe et al. 1970, 1971, 1972) some experimenters remained very reluctant to use the designs, but clung to the standard factorial.

There are certain benefits, I think, to be obtained for econometricians from examination of experimental design and analysis. The first reason for this is because genuine experimentation is now taking place in some branches of economics. Another is that certain issues, long familiar in experimental situations, are being raised now in econometrics. At the recent Barcelona conference of the Econometric Society, a session was devoted to "multiple testing", that is, what are the true error rates and powers like when a lot of regression coefficients are being tested for significance? One speaker was discussing Fisher's test and least-significant-difference and was amazed to hear of honestly-significant-difference and of Tukey's, Scheffe's, Dunnett's and Duncan's tests. It is important that the biometrical literature be examined, not to give credit where it is due, but because there were a lot of controversial issues and repetition of errors should be avoided.

There is a general point about design and analysis that I think has relevance also. The whole main effect, two factor, three factor interactions - linear by linear, linear by quadratic etc. - is *not* meant to be a representation of a "real" model in the sense that econometricians used to regard a structural macro model as real. The actual biological mechanisms through which fertiliser leads to more yields of a crop are fearsomely complex, with a whole series of chemical and biological reactions extending to the bacterial level. The model that relates the final variables of interest to the manipulable set under the experimenters control is essentially a "black box". The factorial effects are convenient devices for the researcher to classify and summarise his findings about the consequences of manipulating factors. But it has worked well in the biological sciences.

Econometricians used to bemoan the fact that, unlike experimenters, they have to make do with whatever data is available and yet their models, which purported to represent the structure of the economy, were vastly more ambitious. Suggestions that only a "reduced form" should ever be considered were usually rejected, on the grounds that "black boxes" are unacceptable. It may well be that inadequate data needs to be supplemented by plausible assumptions drawn from theory and that these assumptions are more likely to be formulated in terms of a structural model. Whether that always justified trying to estimate a structural model is another matter.

However, the retreat from theoretically specified models to an almost atheoretic statistical time series approach is extreme. The ideas of Granger (1969) on causality and "explaining" economic variables by their own past behaviour are open to a lot of criticism. If X "causes" Y only if X has explanatory power over and above that provided by past values of Y, no steadily evolving variable can ever be "caused" by anything else. What would Fisher have thought of Granger causality? Imagine a heifer being fed regularly and consequently gaining weight steadily. If weight at time t (Y_t) is regressed on weight at time $t-1$ a strong relationship will be observed. It is very likely that adding food (X_t) to the regression would not provide a significant improvement. What is the conclusion? Yet the paper of Pierce (1977), showing that autoregression explained most variation, did great damage to the credibility of economic relationships.

The analysis of variance was one of the most attractively packaged of Fisher's ideas. Corresponding to each effect in the design, whether to control error variation or to partition treatment variation, was a line of the analysis of variance table that revealed the magnitude of the effect and tested its significance. Fisher's insistence on retaining as much orthogonality (or balance, in the incomplete blocks case) as possible, ensured that the computational formulae of ANOVA were not too onerous, so that the analyses were quite feasible even in the pre-computer age. I had not mentioned this feature of orthogonality previously, because nowadays computational complexity is not the barrier it once was. But Fisher's attention to keeping calculation reasonably simple was one of the reasons for the spreading popularity of his methods in his own time.

Some of the modern statisticians, (for example, Box and Meyer, 1986 or Leon, Shoemaker and Kacker, 1987) who criticise "Taguchi" methods because they claim they are just slickly packaged versions of classical experimental design and analysis, should remember that Fisher was good at packaging himself and, indeed, had a low opinion of unessential mathematical formulations of ideas and arguments. I will not linger over aspects of the analysis of variance, although I will return to some ideas that originated there at a later stage.

Statistical Inference

Fisher articulated the logic of small sample significance tests, at least to his own satisfaction. The inversion of statistical tests, to put it a little loosely, to give interval estimates gave rise to much dispute and to Fisher's development of fiducial inference. For the normal mean case, the confidence interval approach (Neyman, 1937) would interpret the statement

$$p(\bar{x} - r\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + r\frac{s}{\sqrt{n}}) = .95$$

as meaning that 95 per cent of the intervals obtained by repeated samples of size n (getting a different $\bar{x}$ and s each time) would contain μ . This, of course, provides no grounds for taking the lower and upper values of a *particular* interval

$$\bar{x} \pm t\frac{s}{\sqrt{n}}$$

as measures of upper and lower bounds for μ with probability .95. Fisher (1935c) had already proposed the same formulae for this case as Neyman but with a different argument and interpretation and he called it a "fiducial interval". Fisher regarded the confidence interval approach as rather useless, unless in a context of repeated sampling as, for example, in quality control acceptance sampling. Fisher's argument was that the starting statement

$$p(-t\frac{s}{\sqrt{n}} \leq |\bar{x} - \mu| \leq t\frac{s}{\sqrt{n}}) = .95$$

did not depend on the particular value of μ. If samples were taken repeatedly of pairs $\mu, \bar{x}_i$ from an infinite population of pairs, the probability statement would still be true. So, provided there was no other information about μ (neither another statistic besides $\bar{x}$ containing extra information, nor prior knowledge), the distribution of $\bar{x} - \mu$ could as validly be taken as describing knowledge about μ given $\bar{x}$ as about $\bar{x}$ given μ. This fiducial argument attaches meaning to the upper and lower bounds obtained for the particular case.

There is no doubt that a researcher usually would like upper and lower bounds for the parameter he is trying to estimate in his experiment and if the probabilities attaching to the statement refer to a hypothetical population of parameter values rather than being "real world" probabilities, he may not mind. The controversy that raged over whether fiducial probabilities could be manipulated and interpreted like "real" probabilities may not have been the most important issue. Certainly, Keynes would have accepted probabilities that were not manipulable provided they could measure rational degree of belief.

But besides the interpretation of Fiducial probability it proved difficult to implement in many cases. It required a pivotal quantity ($\bar{x} - \mu$ in the example) and a statistic summing up all the information about the parameter. Sufficient statistics do not always exist, even in the single parameter case, and difficulties multiply with nuisance parameters. Fisher (1956) undoubtedly believed that the maximum likelihood estimator could give an "exhaustive" statistic that could form the basis for fiducial formulations. The idea is that sufficient statistics exist for the exponential family, for which, with appropriate parametrisation

$$\frac{\partial log L}{\partial \Theta} = A[g(x) - \Theta]$$

where A is a constant and where $g(x) = \hat{\Theta}$, the ML estimator. For any one parameter distribution,

$$\frac{\partial log L}{\partial \Theta} \simeq (\hat{\Theta} - \Theta)\left\{\frac{\partial^2 log L}{\partial \Theta^2}\right\}\hat{\Theta}$$

So Fisher saw the deviation from sufficiency, or the "loss of information" as determined by the degree to which the second derivative of the likelihood deviated from constancy. Therefore he proposed taking the distribution of $\hat{\Theta}$ conditionally on the second derivative as a device to "recover" the lost information. Unfortunately, apart from suggestive examples, Fisher never clarified how a general method would work. The ideas underlie the papers by Efron (1975) and Efron and Hinkley (1978).

The idea of conditioning on an ancillary statistic is not at all free of controversy and Fisher seems to have chosen to condition when it suited him. A simple situation is the comparison of proportions in the 2 x 2 table. One might think this would have been sorted out long ago, but in fact the case is still a live issue. A full session at this year's conference of the American Statistical Association was devoted to it and recent published papers on the topic include D'Agostino et al. (1988), Barnard (1990a) and Little (1989). Fisher favoured his own "Fisher Exact Test" which involves conditioning on an ancillary, while the "other" test (essentially a comparison of binomial proportions) is unconditional.

Suppose the numbers of successes and failures (S and F) in two groups of subjects (G and H) are as follows

	S	F
G	a	b
H	c	d

The unconditional test compares the proportion

$$\frac{a}{a+b} \quad with \quad \frac{c}{c+d},$$

using chi-squared or a normal approximation. But if the total of successes $a + c$ is uninformative about the difference between the proportions in the two groups (obviously it is informative about the overall proportion), Fisher felt it reasonable to condition on it in formulating a test. On the null hypothesis of equal probability of success in both groups the joint distribution of a and c is

$${}^{a+b}C_a \;\; {}^{c+d}C_a \;\; p^{a+c} \;\; (1-p)^{b+d}$$

the product of two binomials. But the marginal distribution of $(a+c)$ from $(a+b+c+d)$ is also binomial

$${}^{a+b+c+d}C_{a+c} \;\; p^{a+c} \;\; (1-p)^{b+d}$$

So the conditional distribution of a and c given $a + c$ is

$$\frac{{}^{a+b}C_a^{\,c+d}C_c}{{}^{a+b+c+d}C_{a+c}}$$

which is the hypergeometric distribution. But this can give different results to the unconditional test. Also Fisher never reconciled his argument with other prescriptions of his. A score test, for example, would lead to the unconditional test.

Inference based on Likelihood was a fallback position for Fisher from Fiducial Inference. The Likelihood Principle - that all the information that the data provides concerning the relative merits of two hypotheses is contained in the ratio of likelihoods - can be adopted to provide a whole system of inference, as has been done by Edwards (1972). The ideas are all in Fisher's work, but so are the basic notions for other systems of inference and ambiguities abound. Why, for example, does Fisher switch in his 1956 book, from stressing the relevance of values more extreme than observed for significance tests to their irrelevance for estimation? An example by Barnard (1972) can be modified to illustrate the issue.

Suppose a computer displays a five letter 'word'. An initial hypothesis is that the machine is programmed to give a random arrangement of five letters. Suppose 'HORSE' results. One would doubt the hypothesis because the number of possible arrangements (26^5) is large relative to the number of genuine words of five letters (W). It is not the probability of 'HORSE' itself, but of the class of which it is one representative, that matters to the significance test argument. The hypotheses would be rejected if $W/26^5$ was less than the significance level. In his 1956 book, Fisher argued strongly for this sort of procedure.

Yet suppose there are only three possible situations: the computer either gives a random arrangement (H_1), selects from the class of words (H_2) or can only give 'HORSE' (H_3). The probabilities of 'HORSE' given these hypotheses, which are the values of the likelihood, are

H_1	H_2	H_3
26^{-5}	W^{-1}	1

Clearly H_3 maximises the likelihood. For estimation, Fisher stressed considering only the probability of the particular observed event and yet if he was to be consistent with the idea underlying significance tests he should have looked for the probabilities of the class given the hypotheses. In this example these would be

H_1	H_2	H_3
$W/26^5$	1	1

So now hypotheses H_2 and H_3 seem equally supported by the evidence. It could be argued that in estimation, as in significance testing, evidence should be measured by an integral or expectation.

In regard to maximum likelihood estimation itself and associated methodologies like score and likelihood ratio tests, many statisticians claim to hold a pragmatic view. They accept the methods because they have good properties and work well, rather than because they favour a particular philosophy of inference. But MLE does not always work well, especially when there are large numbers of nuisance parameters. Methods can be developed that coincide with MLE when it works well and that improve on it in other circumstances. At the risk of seeming to stress my own work (although Roy Geary said - "I used to be modest, but I got that out of my system when I was young"), I have proposed methods (Conniffe, 1987, 1988, 1990) that not only do this, but are also, I think, more intuitively plausible.

The jury is still out on a lot of Fisher's ideas on inference, but that does not mean his contributions were not vital. Kendall (1963) put it well when he said Fisher's work on inference, while it might prove less permanent than his work on distribution theory or experimental design, had stimulated most of the thinking done throughout the world on inductive reasoning.

Some Selected Topics

The number of other aspects of Fisher's contributions that could be mentioned is huge, but the paper must be kept to a reasonable length. I will finish the overview of his work with a few examples of how ideas of his have surfaced in econometrics. Fisher (1925b) pointed out, in passing, in a paper on estimation theory, that the covariance between an efficient estimator and another estimator of a parameter must be equal to the variance of the efficient estimator. An intuitive explanation for the result is that if it were not true a still better estimator could be constructed as

$$[(V - C)\hat{\Theta}_E + (V_E - C)\hat{\Theta}]/(V + V_E - 2C)$$

where $\hat{\Theta}_E$ and $\hat{\Theta}$ are the efficient and other estimators respectively and V, V_E and C are the variances and covariance. But then there would be a contradiction so

$$V_E = C \quad \text{or} \quad Var(\hat{\Theta} - \hat{\Theta}_E) = V - V_E$$

So the variance of the difference between the estimators has a very simple form. Now very often it is obvious that two estimators have the same expectation if some hypothesis is true. For example, in linear regression OLS is unbiased even if there is heterogeneity of variance of some specified form. But a GLS taking account of the specification would give the efficient estimator. So a simple comparison can test the specification, at least on an asymptotic test. The whole simplicity of the procedure depends on $V_E = C$. What I've described are, of course, Hausman (1978) tests.

When Fisher fitted analysis of variance, or linear models, to experimental data there were usually several dependent variables. For example, in an experiment on barley the variables might be: the yield of grain, the protein content, the energy content and so on. Standard analysis was equivalent to what econometricians would call single equation OLS. But Fisher also used covariance analysis, where some dependent variables were also analysed conditionally on the values of others, usually when the latter are known not to be affected by the sub-set of explanatory variables corresponding to treatments. I showed that this is actually equivalent to a simple case of what would have been called a systems approach in econometrics and is much more convenient computationally (Conniffe, 1982a). More generally, systems methods are equivalent to augmenting the variables in single equations with the residuals from other equations. By iteration, the estimates can be made computationally identical, but in some cases an uniterated single equation method may be more efficient (for example, Conniffe, 1982b). This correspondence is now well known, but the converse does not seem to be appreciated.

A single equation, containing what are in effect residuals, can be represented as a system of equations. Currently, single econometric equations often contain an error correction term, or 'cointegrated' relationship as an explanatory variable. Now I'm not trying to say that error correction terms, or feedback control loops, may not sometimes make sense in a single equation, just as differenced, or lagged, variables may make sense too. But a quite complicated single equation with all these features may actually be just a result of an underlying system of quite simple equations displaying no explicit "dynamic" effects, but with a non-diagonal covariance matrix. It is easy to add "dynamics" to a single equation and there is nothing like a lagged dependent variable to improve the looks of an unpromising regression, but true causative variables can easily be displaced. The frequently seen procedure of initially including lots of lagged versions of all variables, plus a few error correction terms, and then selecting retained variables purely on statistical criteria, is too reminiscent of stepwise regression. Researchers with good judgement may be able to use it sensibly, but there is scope for self-deception.

5. FISHER'S LATER LIFE AND CAREER

As already mentioned, recognition did not come easily to Fisher, at least from his statistical colleagues. His ideas spread quickly among biological research workers and his 1925 book, which was more a practical handbook than a textbook, was reprinted again and again. The importance of his statistical research was accepted in America, before his UK colleagues commenced to take it seriously. When he was finally invited to describe his work to the Royal Statistical Society, the reception he received was (Fisher, 1935a) largely one of incomprehension.

The situation at Rothamsted had its drawbacks and he made several attempts to gain University Chairs. He applied to the LSE in 1929, but was unsuccessful. He hoped to get Karl Pearson's chair in statistics at University College London when that dominant figure retired in 1933, but that post went to Pearson's son, Egon. However, he was offered and accepted, a professorship of eugenics and with it the editorship of the *Annals of Eugenics*. Although this journal had been established for the publication of papers on eugenics and genetics, under Fisher's editorship it published some of the most important statistical research, which would probably otherwise have gone to *Biometrika*, now edited by Egon Pearson. While working in London he continued to live in Harpenden and had made arrangements with Rothamsted to continue to conduct genetic research there. In 1938 he and Yates published *Statistical Tables for Biological, Agricultural and Medical Research*, a work every statistician was extremely familiar with in the pre-computer age. The publication also undermined the previous importance of the *Biometrika Tables for Statisticians.*

Gradually Fisher achieved the recognition desired for so long. He had been elected a Fellow of the Royal Society in 1933, because of his biometrical and genetic work at Rothamsted, and was awarded its Royal Medal in 1938. A team of very able statisticians, largely spreading out from Rothamsted, was turning Fisherian statistics into the new orthodoxy. But he continued to work as hard as ever and certainly did not become anyway mellowed towards his statistical enemies.

Fisher's marriage broke down in 1943 when his wife took up a fiercely evangelical brand of religion, which he could not abide. But since Fisher was probably not the easiest man to live with, the religious conversion was possibly a symptom. Fisher, while regarding the practice of religion as a salutary exercise, was not over impressed by dogmatic Christianity. In a 1955 radio broadcast he said:

> *I do not think that the word for the Christian virtue of faith should be prostituted to mean the credulous acceptance of piously intended assertions. Much self deception in the believer is needed to convince himself that he knows that of which in reality he knows himself to be ignorant.*

Oddly enough, he had an enormous knowledge of scripture and biblical studies, but clearly his views were not sympathetic to evangelical religion.

Fisher never held a Chair in Statistics. He moved in 1943, alone except for his mice, to the Chair of Genetics at Cambridge where he remained until 1959, although he officially retired in 1957. He was a Fellow of his undergraduate college, Gonville and Caius, and was elected President of the College in 1956. However, this does not mean his time at Cambridge was relaxed and placid. He was almost continually involved in quarrels with the administration of the University and when he was knighted in 1952 he said "That's one in the eye for the University". He was now also acquiring honorary degrees and other forms of recognition from institutions around the world. But he still did not mellow and twenty years after Karl Pearson's death, Fisher was still writing angrily about him. He was still working too. When he had moved from London he had to give up editorship of the *Annals of Eugenics* and so he founded the journals *Heredity* and *Biometrics* to make up for it.

After retiring from Cambridge he continued with various research interests, including getting involved in the smoking and cancer controversy (Fisher, 1959). He visited Australia and liked the place, so he accepted a Research Fellowship at Adelaide. He was still engaged in statistical research when be became ill and died there in July 1962.

6. CONCLUDING REMARKS

The account I have given of Fisher's work is probably not entirely a balanced one, because I have deliberately placed emphasis on themes relevant to econometrics, or with some Irish connections, even if tenuous. I hope, though, that I have indicated the huge scope of his work and at least touched on some of the controversies that have been associated with the development of statistical theory. Fisher was a genius, but not infallible. However, if his treatment of some issues cannot be considered convincing, no one else has resolved the matters in the meantime.

References

Anderson, O., 1914. Comments on "The Elimination of Spurious Correlation due to Position in Time and Space" *Biometrika*, Vol. 10, pp. 269-79.

Atkinson, A.C., 1969. "The Use of Residuals as a Concomitant Variable", *Biometrika*, Vol. 56, pp. 33-41.

Barnard, G.A., 1972. "The Unity of Statistics", *Journal of the Royal Statistical Society A*, Vol. 135, pp. 1-14.

Barnard, G.A., 1990a. " P Values and Power Functions for 2 x 2 tables" *Professional Statistician*, Vol. 9, No.2, pp. 4-5.

Barnard, G.A., 1990b. "Fisher: A Retrospective, *Chance*, Vol. 3, pp. 22-32.

Bartlett, M.S., 1978. "Nearest Neighbour Models in the Analysis of Field Experiments", *Journal of the Royal Statistical Society B*, Vol. 40, pp. 147-174.

Boland, P.J., 1984. "A Biographical Glimpse of William Sealy Gossett" *American Statisticians*, Vol. 38, pp. 179-183.

Bowley, A.L., 1935. "Discussion of Fisher's Paper" *Journal of the Royal Statistical Society*, Vol. 98, pp. 55-57.

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

Box, G.E.P. and Meyer, R.D., 1986. "Dispersion Effects from Fractional Design" *Technometrics*, Vol. 28, pp. 19-27.

Cave-Brown-Cave, F.E., 1904. "On the influence of the time factor and the correlations between barometric heights and stations more than 1,000 miles apart", *Proceedings of the Royal Society A*, Vol. 74, pp. 403-413.

Conniffe, D., Browne, D. and Walshe, M.J., 1970. "Experimental Design for Grazing Trials" *Journal of Agricultural Science, Cambridge*, Vol. 74, pp. 339-342.

Conniffe, D., 1971. "Treatment Comparisons in Grazing Trials using the Animal as Experimental Unit". *Journal of Agricultural Science, Cambridge,* Vol. 77, pp. 227-235.

Conniffe, D., Browne, D. and Walshe, M.J., 1972. "An example of a method of statistical analysis of a grazing experiment", *Journal of Agricultural Science, Cambridge,* Vol. 79, pp. 165-167.

Conniffe, D., 1982a. "Covariance Analysis and Seemingly Unrelated Regressions", *American Statistician,* Vol. 36, pp. 169-171.

Conniffe, D., 1982b. "A Note on Seemingly Unrelated Regressions", *Econometrica,* Vol. 50, pp. 229-233.

Conniffe, D., 1987. "Expected maximum log likelihood estimation", *The Statistician,* Vol. 36, pp. 317-329.

Conniffe, D., 1988. "Obtaining Expected Maximum Log Likelihood Estimators", *The Statistician,* Vol. 37, pp. 441-449, also Corrigendum, **1989** *The Statistician,* Vol. 38, p. 87.

Conniffe, D., and Keogh, G., 1988. *Equivalence Scales and Costs of Children,* Dublin, ESRI.

Conniffe, D., 1989. "Experimental Design, Ireland and Taguchi Methods" in E. Murphy, ed. *Proceedings of 1st Symposium on Taguchi Methods,* pp. 93-110, University of Limerick .

Conniffe, D., 1990. "Testing Hypotheses with Estimated Scores", *Biometrika,* Vol. 77, pp. 97-106.

D'Agostino, R.B., Chase, W. and Belanger, A., 1988. "The appropriateness of some common procedures for testing the equality of two independent binomial populations", *American Statistician,* Vol. 42, pp. 198-202.

Edgeworth, F.Y., 1885. "Observations and Statistics" *Transactions of the Cambridge Philosophical Society,* Vol. 14.

Edgeworth, F.Y., 1905. "The Law of Error" *Transactions of the Cambridge Philosophical Society,* Vol. 20, pp. 36-65.

Edgeworth, F.Y., 1908. "On the probable error of frequency constants" *Journal of Royal Statistical Society,* Vol. 71, pp. 381-397.

Edwards, A.W.F., 1972. *Likelihood,* Cambridge, Cup.

Efron, B., 1975. "Defining the Curvature of a Statistical Problem", *Annals of Statistics* Vol. 3, pp. 1189-1242.

Efron, B. and Hinkley, D.V., 1978. "Assessing the Accuracy of the ML Estimator" *Biometrika,* Vol. 65, pp. 457-488.

Elderton, E.M. and Pearson, K., 1910. "A first study of the influence of parental alcoholism on the physique and ability of the offspring", *Eugenics Laboratory Memoirs,* Vol. 10.

Fisher, R.A., 1912. "On an absolute criterion for fitting frequency curves", *Messenger of Mathematics,* Vol. 41, pp. 155-160.

Fisher, R.A., 1915. "Frequency distribution of values of the correlation coefficient in samples from an indefinitely large population", *Biometrika,* Vol. 10, pp. 507-521.

Fisher, R.A., 1918. "The correlation between relatives on supposition of Mendelian inheritance", *Transactions of the Royal Society of Edinburgh,* Vol. 52, pp. 399-433.

Fisher, R.A., 1921. "On the 'probable error' of the coefficient of correlation deduced from a small sample", *Metron,* Vol. 1, pp. 1-32.

Fisher, R.A., 1922. "On the mathematical foundations of theoretical statistics". *Philosophical Transactions of the Royal Society A,* Vol. 222, pp. 309-368.

Fisher, R.A., 1925a. *Statistical Methods for Research Workers,* London, Oliver and Boyd.

Fisher, R.A., 1925b. "Theory of Statistical Estimation", *Proceedings of Cambridge Philosophical Society,* Vol. 22, pp. 700-725.

Fisher, R.A., 1930. *The Genetical Theory of Natural Selection,* Oxford, OUP.

Fisher, R.A., 1931. "The biological effects of family allowances", *Family Endowment Chronicle,* Vol. 1, pp. 21-25.

Fisher, R.A., 1935a. "The Logic of Inductive Inference". *Journal of the Royal Statistical Society,* Vol. 98, pp. 39-82.

Fisher, R.A., 1935b. *The Design of Experiments,* London, Oliver and Boyd.

Fisher, R.A., 1935c. "The Fiducial Argument in Statistical Inference", *Annals of Eugenics,* Vol. 6, pp. 391-398.

Fisher, R.A. and Yates, F., 1938. *Statistical Tables for Biological, Agricultural and Medical Research,* Edinburgh, Oliver and Boyd.

Fisher, R.A., 1939. "Student", *Annals of Eugenics,* Vol. 9, pp. 1-9.

Fisher, R.A., 1956. *Statistical Methods and Scientific Inference,* London, Oliver and Boyd.

Fisher, R.A., 1959. *Smoking - The Cancer Controversy,* Edinburgh, Oliver and Boyd.

Galton, F., 1889. *Natural Inheritance,* London, MacMillan.

Geary, R.C., 1963. "R.A. Fisher: A Memoir", *The Economic and Social Review,* Vol. 14, pp. 167-171.

Granger, C.W.J., 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectoral Methods", *Econometrica,* Vol. 37, pp. 424-438.

Granger, C.W.J. and Newbold, P., 1974. "Spurious Regressions in Econometrics", *Journal of Econometrics,* Vol. 2, pp. 111-120.

Hausman, J.A., 1978. "Specification tests in Econometrics", *Econometrica,* Vol. 46, pp. 1211-1271.

Hooker, R.H., 1905. "On the correlation of successive observations" *Journal of the Royal Statistical Society,* Vol. 68, pp. 696-703.

Kendall, M.G., 1963. "Ronald Aylmer Fisher, 1890-1962", *Biometrika,* Vol. 50, pp. 1-15.

Kendall, M.G. and Stuart, A., 1967. *The Advanced Theory of Statistics,* Vol. 2, London, Griffin.

Keynes, J.M., 1911. "The Principal Averages and the Laws of Errors which lead to these". *Journal of the Royal Statistical Society,* Vol. 54, pp. 322-331.

Keynes, J.M., 1921. *A Treatise on Probability,* London, Macmillan.

Leon, R.V., Shoemaker, A.C. and Kacker, R.N., 1987. "Performance Measures Independent of Adjustment - An Expansion and Extension of Taguchi's signal-to-noise Ratio", *Technometrics,* Vol. 29, pp. 253-265.

Little, R.J.A., 1989. "Testing the equality of Two Independent Binomial Proportions", *American Statistician,* Vol. 43, pp. 283-288.

Neyman, J., 1937. "Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability" *Philosophical Transactions of the Royal Society,* Vol. A, pp. 333-380.

Neyman, J., 1956. "Note on an Article by Sir Ronald Fisher", *Journal of the Royal Statistical Society,* Vol. 18, pp. 288-294.

Papadakis, J.S., 1937. "Methode Statistique Pour Des Experiences Sur Champ," *Bull. Inst. Amel. Plantes a Salonique,* Vol. 23.

Pearson, K., 1906. "On the curves which are most suitable for describing the frequency of random samples of a population", *Biometrika,* Vol. 5, pp. 172-175.

Plackett, R.L. and Barnard, G.A., 1990. *'Student' (Based on Writings of E.S. Pearson),* Oxford, Clarendon Press.

Pierce, D.A., 1977. "Relationships - and the lack thereof - between economic time series, with special reference to money and interest rates". *Journal of the American Statistical Association,* Vol. 72, pp. 11-26.

Spencer, J.E., 1976. "The Scientific Work of Robert Charles Geary", *The Economic and Social Review,* Vol. 7, pp. 233-241.

Spencer, J.E., 1983. "Robert Charles Geary - An Appreciation", *The Economic and Social Review,* Vol. 14, pp. 161-164.

'Student', 1908a. "The Probable Error of a Mean", *Biometrika,* Vol. 6, pp. 1-25.

'Student', 1908b. "The Probable Error of a Correlation Coefficient", *Biometrika,* Vol. 6, pp. 302-310.

'Student', 1911. Appendix to Mercer and Hall. "The Experimental Error of Field Trials". *Journal of Agricultural Science Cambridge,* Vol. 6, pp. 128-131.

'Student', 1914. "The Elimination of Spurious Correlation due to Position in Time or Space", *Biometrika,* Vol. 10, pp. 179-181.

Taguchi, G., 1976. *Experimental Designs.* Vol. 1, Tokyo, Maruzen.

Taguchi, G., 1977. *Experimental Designs.* Vol. 2, Tokyo, Maruzen.

DISCUSSION

John E. Spencer: I am extremely pleased to propose the vote of thanks to Professor Denis Conniffe not only on account of my interest in the statistical work of Fisher but also because of my interest in much of Denis' own work. In these notes, I shall try to complement Denis' paper, by dwelling a little more on Fisher's influence on some later developments.

Fisher's contribution to statistics should be seen as a whole and, as once pointed out by Bartlett, a listing of some highlights risks devaluing his work's overall unity and impact. With this in mind, it is nonetheless worth reminding ourselves that his contributions included information theory, fiducial probability and inference, discriminant function, likelihood, maximum likelihood estimation, efficiency, sufficiency, consistency, Fisher consistency, replication, randomization and block division in experimental design, analysis of variance, k statistics, Fisher's Z transformation, the variance ratio distribution (F), the distributions of the t-ratio, the mean square error, the regression coefficient, the correlation and multiple correlation coefficients in various circumstances, Fisher's exact test in 2 x 2 tables and much else besides, including fundamental contributions to genetics and, independently of Von Neumann, the idea of mixed strategies in games. He was also interested in application and in the popularization of statistical methods and his early book *Statistical Methods for Research Workers*, published in 1925, went through many editions and motivated and influenced the practical use of statistics in many fields of study. His *Design of Experiments* (1935) was also crucially fundamental in the promotion of statistical technique and application. In that book he emphasised examples and how to design experiments systematically from a statistical point of view. The mathematical justification of the methods described was not stressed and, indeed, proofs were often barely sketched or omitted altogether (see below), a fact which led H B Mann to fill the gaps with a rigorous mathematical treatment in his well known treatise, Mann (1949).

Fisher's best period was probably during the 1920s and 1930s a period when he was revolutionizing the theory of statistics and setting the agenda for the next twenty years and more. According to Roy Geary, also active then, statisticians at that time needed only to read Fisher. Everything was there, and only had to be dug out a bit. At the same time, the foundations of modern probability theory were being laid, mainly by Russian and French probabilists. By the 1940s, the time was opportune to bring together these two lines of development, a task attempted very successfully by Harold Cramer in his highly influential book, Cramer (1946).

Perhaps Fisher's work and influence can be exemplified by one of his most famous contributions, maximum likelihood (ML) estimation. ML probably grew out of his notion of likelihood, a notion itself connected with his controversial theory of fiducial probability and fiducial inference. If X is $N(m,1)$, $\bar{X}$ is $N(m,n^{-1})$ where n is the sample size and $P(\bar{X}-1.96n^{-\frac{1}{2}} \leq m \leq \bar{X}+1.96n^{-\frac{1}{2}}) = .95$ where the limits in parentheses define a random interval and P denotes probability. Once a particular sample is taken, however, the expression in parentheses is either true or false, i.e. has probability 1 or 0, and is the Neyman-Pearson 95 per cent confidence interval. Fisher, however, speaks of likelihood as measuring our intensity of credence in a particular value of m and of the fiducial distribution of m, viz, $N(\bar{x},n^{-1})$, as valid and defined after a sample is taken. Thus, $\bar{x}$ is a known number and $P(\bar{x}-1.96n^{-\frac{1}{2}} \leq m \leq \bar{x}+1.96n^{-\frac{1}{2}})$ is asserted to be 0.95 in this particular case, i.e. the probability statement has post-sample validity. An interesting summary and interpretation is in Dempster (1964).

This theory was always puzzling and controversial and is less discussed nowadays. Savage ends his famous book (Savage, 1954, 2nd revised edition 1972) with an apology for not having a serious section on fiducial probability and describes it as "the most disputed technical concept of modern statistics". Cramer (1981) while admitting it is possible sometimes to regard an unknown parameter as determined itself by a random experiment, thereby allowing Bayes methods, regards the parameters as fixed but unknown in other cases, for Cramer the majority. Cramer found Fisher impossible to follow on this matter, considering him to have made a mathematical mistake and regarding his deduction of a fiducial distribution to be in contradiction with modern probability theory.

From fiducial theory it is not a large step to maximum likelihood estimation. Thus, the probability of 2 successes in 3 Bernoulli trials is $3p^2(1-p)$ in obvious notation and regarding this as a function of p, the likelihood function, ML estimation chooses as estimator of p that value which maximises this likelihood function. Thus, in the likelihood function, the sample data is taken as fixed (2 successes in this example) and the parameter treated as a variable. Differentiating the likelihood function and equating to zero yields the likelihood equation which is solved for the ML estimator (MLE).

A vast edifice has been built on this principle and it has been described as the most important general method of estimation so far known from a theoretical point of view (Cramer, 1946). While claims have been made for others including Gauss and Edgeworth, it seems fair to regard Fisher as its originator (Rao, 1962, Norden, 1972). Its importance derives, not from its philosophical underpinnings, but from the properties of estimators which result. Fisher showed several such properties including a large sample efficiency property, later generalised by Geary to the many parameter case. An interesting property, apparently valued by Fisher (Anderson, 1986), is the invariance of ML estimators under transformations so that if b is MLE of B, $f(b)$ is MLE of $f(B)$, under mild conditions on f.

Many writer have developed the theory since Fisher's time, including issues relating to local versus global maximisation of the likelihood equation, boundary values, multiple roots, nondifferentiability and the like. Cramer (1946) showed consistency of some root and asymptotic normality under regularity conditions and showed the attainment of the Cramer-Rao lower bound by the variance (an inequality missed by Fisher) for ML estimators under large samples and regularity conditions. Fisher himself had shown in a non-rigorous way that ML estimators were asymptotically efficient and Rao was to extend and refine these results (DeGroot, 1987, Norden, 1972, 1973).

C R Rao, perhaps Fisher's most famous student in the immediate postwar period when Fisher had been appointed to the Balfour chair of genetics at Cambridge, has described Fisher, with Mahalanobis, as one of the two most important influences on his life and career (DeGroot, 1987). Registering for a PhD with Fisher, he was told by Fisher that he had to work on his mice, whether it led to a PhD thesis or not! Geary (1983) also refers to these mice and their "very pungent (but not unpleasant) aroma" when he met Fisher in his Cambridge laboratory at some time around 1947. Fisher apparently was more interested in genetics than statistics at that time, an observation also made by T W Anderson who was then working on discriminant analysis but found Fisher more interested in showing him round his garden and talking about genetics (Anderson op cit). Rao's thesis was also on discriminant analysis, though he took courses in mathematical genetics and wrote a paper on that topic which Fisher refused to read unless data was gathered and computations done. This emphasis on practical problems influenced Rao, as did Fisher's emphasis on graphics, and on the importance of sample surveys and experimental design.

Harold Cramer, born just three years after Fisher and nearly thirty years before Rao, was less influenced directly by Fisher. He met him first in London in 1938 and a year later in Geneva. On the latter occasion, he complimented Fisher on his geometric intuition, to be told "I am sometimes accused of intuition as a crime" (Cramer, 1981).

Fisher, indeed, seems to have had great geometric vision and strong intuition. Geary (op cit) recalled meeting him and asking about his 1929 hugely complicated expressions on the cumulants of k statistics only to be told that the formulae came to him on a train to Edinburgh, a story confirmed by Wishart. Mahalanobis (1938) suggests that this extraordinary ability to calculate in his head may have derived from his poor eyesight. For as a child he had been forbidden to read by artificial light and was taught much mathematics orally. And it is surely also likely that his lack of attention to rigour and the gaps and flaws in his proofs came from his impatience and desire to explore new ground. This, with his strong intuition, makes him difficult to read. Similarly, as a teacher he was inspiring if one could see what he was saying - though not many could (DeGroot, 1987).

He does not seem to have been a charming figure personally and many people were willing to attack him. His conflict with Karl Pearson was open and bitter. When Cramer lectured in Paris in 1946 on fiducial probability and Neyman confidence intervals, he was somewhat dismayed to find Fisher there and at the lecture, as he sided with Neyman in the controversy. Afterwards Fisher claimed he was not able to follow the lecture with his limited French and suggested dinner and a private talk, which apparently worked out happily enough (Cramer 1976, 1981).

I have great pleasure in thanking Denis for his extremely entertaining and instructive paper.

References

Anderson, T.W., 1986. "The E T Interview: Professor T W Anderson, interviewed by P C B Phillips", *Econometric Theory*, Vol. 2, No. 2, pp. 249-288.

Cramer, H., 1946. *Mathematical Methods of Statistics*, Princeton, Princeton University Press.

Cramer, H., 1976. "Half a Century with Probability Theory: Some Personal Recollections", *Annals of Probability*, Vol. 4, No. 4, pp. 509-546.

Cramer, H., 1981. "Mathematical Probability and Statistical Inference", *International Statistical Review*, Vol. 49, No. 3, pp. 309-317.

Dempster, A.P., 1964. "On the Difficulties Inherent in Fisher's Fiducial Argument", *Journal of American Statistical Association*, Vol. 59, pp. 56-66.

DeGroot, M., 1987. "A Conversation with C R Rao", *Statistical Science*, Vol. 2, No. 1, pp. 53-67.

Geary, R.C., 1983. "R A Fisher: A Memoir", *Economic and Social Review*, Vol. 14, No. 3, pp. 167-171.

Mahalanobis, P.C., 1938. "Professor Ronald Aylmer Fisher", *Sankhya*, Vol. 4, Part 2, pp. 265-272 and reprinted in W A Shewhart, ed. *Contributions to Mathematical Statistics by R A Fisher*, Wiley, 1950.

Mann, H.B., 1949. *Analysis and Design of Experiments*, New York, Dover.

Norden, R.H., 1972. "A Survey of Maximum Likelihood Estimation", *International Statistical Review*, Vol. 40, No. 3, pp. 329-354.

Norden, R.H., 1973. "A Survey of Maximum Likelihood Estimation, Part 2", *International Statistical Review*, Vol. 41, No. 1, pp. 39-58.

Rao, C.R., 1962. "Apparent Anomalies and Irregularities in Maximum Likelihood Estimation (with discussion)", *Sankhya,* A, Vol. 24, pp. 73-101.

Savage, L.J., 1954. *The Foundations of Statistics,* Wiley, 2nd revised edition, 1972, New York, Dover.

P. Boland: I would like to congratulate Professor Conniffe on his versatile and wide ranging tribute to R.A. Fisher, a man of true genius who laid much of the foundations for modern statistical methodology. In some sense it is a difficult challenge to pay ***appropriate*** tribute to a man of Fisher's accomplishments, because he made so many diverse contributions to scientific research. In a tribute to R.A. Fisher in the Journal of the Royal Statistical Society in 1963 (just after the death of Fisher), George Barnard wrote -

> *"to attempt, in a short article, to assess the contributions to the subject by one largely responsible for its creation would be futile"*

his central contribution was

> *"his deepening of our understanding of uncertainty"*

and of

> *"the many types of measurable uncertainty".*

Professor Conniffe notes that Fisher had poor eyesight, and it is perhaps worth adding that this handicap influenced his early mathematical education to the extent that he was largely taught without the use of standard tools like pencil, paper and other visual aids. As a result he developed an uncanny geometrical insight which was to later greatly contribute to his discoveries regarding statistical distributions. In fact it was this insight which was to lead him to suggest to 'Student' in 1912, that by using n dimensions, the formula for the standard deviation should be

$$\sqrt{\sum \frac{(x-\bar{x})^2}{n-1}}$$

and not

$$\sqrt{\sum \frac{(x-\bar{x})^2}{n}}$$

This was also essentially the discovery by Fisher of the concept of "degrees of freedom" a tool which is so extensively used today.

I enjoyed Professor Conniffe's discussion of Fisher's desire to increase the numbers of the 'elite'. Given Ireland's relatively high birth rate in recent generations and its contributions to the 'elite' of the English speaking world via large families and emigration, it could perhaps be suggested that this phenomena represents an Irish connection with Fisher.

The method of maximum likelihood for estimating unknown parameters of a distribution is certainly, as Professor Conniffe has pointed out, one of Fisher's more important contributions to Statistics. Professor Conniffe has himself recently made valuable contributions to the theory of this method in the *Statistician*. Fisher was able to show that the method of maximum likelihood is almost always considerably more efficient than the then commonly used method of moments. Students of Statistics are today taught that the method of maximum likelihood is a very intuitive method, and that asymptotically, that is when working with large samples, the method is unbeatable. In spite of this I do not think enough care is taken in using this method with small samples. An interesting example of the pitfalls of this method appears in a recent article of the Journal of the American Statistical Association entitled "Estimating the Number of Faults in a System" (1985). Here the object is to estimate the unknown number of faults in a piece of software, after observing the software for some time. It is shown that with substantial probability, the maximum likelihood estimator of the unknown number of faults does not exist!

Professor Conniffe has rightly emphasized the contributions Fisher made in reconciling Mendelism and Darwinism. As Mather wrote (*JRSS 1963*), Fisher

> *"went beyond merely harmonizing to fusing the principles of genetics and natural selection".*

'Student' concurred with Fisher in this area, and in fact he sought the support of Fisher and his mathematical ability in justifying some of his own arguments for the theory of evolution by selection. An interesting and humorous letter in this regard appears in "A Biographical Glimpse of William Sealy Gosset", (*American Statistician, 1983*).

As far reaching and diverse as Professor Conniffe's view of Fisher has been, it is inevitable that certain contributions of Fisher have not been covered in depth. We have been told tonight that "Fisher believed that not only physical qualities were heritable, but so also were socially useful traits". A social trait which he suspected was heritable was that of smoking habit. In his later years he became caught up in the extensive debate and controversy covering lung cancer and cigarettes.

For example, he wrote in Nature in 1958

> *"The association observable between the practice of cigarette smoking and the incidence of cancer of the lung, to which attention has been actively, or even vehemently, directed by the Medical Research Council Statistical Unit, has been interpreted, by that Unit, almost as though it demonstrated a causal connexion between these variables.*
>
> *The suggestion, among others that might be made on the present evidence, that without any direct causation being involved, both characteristics might be largely influenced by a common cause, in this case the individual genotype, was indeed rejected with some contempt by one writer, although I believe that no one doubts the importance of the genotype in predisposing to cancers of all types".*

The whole episode was in hindsight perhaps regrettable, given that he put his tremendous statistical reputation on the wrong side of the controversy.

In conclusion, I would like to second the vote of thanks for Professor Conniffe's stimulating and far reaching view of the greatest of statisticians, Ronald A. Fisher. I think the following story about Gosset ('Student') by Stella Cunliffe *(JRSS, 1976)* appropriately summarises the respect most statisticians have for Fisher -

"Once in 1937, a young statistician who went to consult him (Student) said pompously:

- On behalf of fellow statisticians, I would like to thank you for all that you have done for the advancement of statistics -

to which Gosset replied:

Oh that's nothing, Fisher would have discovered it all anyway."

G. MacKenzie: I would like to congratulate Professor Conniffe on a most stimulating paper. Put simply, Ronald Fisher was the founder of Theoretical Statistics. His inventions, the output of a super-intellect, are the most distinctive features on the statistical landscape today. In large measure Fisher's contribution was meta-science - it transcended mere application.

Although most of what I have to say tonight is supplementary, I am bound to comment that, on reflection, I found Professor Conniffe's 'Econometric Perspective' somewhat tangential. The forces which moulded Fisher, as I shall show below, were rather different. They were the great scientific issues of the day: Darwinism, Evolution, Inheritance, Genetics, the role of Biometry, and of Statistical Method. It was in developing the theory relevant to these latter areas that Fisher was to excel.

In passing, I may remark that I have ever found his theory of Maximum Likelihood compelling, and no less intuitive than many other optimality principles. I do not mean to suggest that the principle is necessarily obvious, though it is clever. Its connection with *Sufficiency* merits formal consideration. The concept of *Sufficiency*, one of Fisher's greatest insights, coupled with *Ancillarity*, another, have spawned methods of *Conditional, Marginal, Partial* and *Profile Likelihood* analysis. Such methods are based on *Inferential Separation* - of statistics and parameters in the *Likelihood.* It is these developments, and the near optimal results to which they usually lead, that have sustained *Frequentist Statistics* in the face of a concerted and cogent challenge form the *Bayesian School.* Today, it is David Roxbee Cox who bears Fisher's standard. But that is another story.

As a young medical statistician, working in a Queen's Department which was founded, originally, by a medical geneticist, I was soon engrossed in Fisher's early papers and his various exchanges with Pearson. Although my primary interest lay in the underlying theory, one was drawn, inevitably, not only by the historic nature of the theoretical debate, but also by the personalities involved. Accordingly, I would like to outline briefly some relevant history without which it is difficult, if not impossible, to appreciate fully Ronald Fisher's contribution.

R.A. Fisher and the influence of the Biometrical School 1888-1925

Although Modern Statistics is only a hundred years old it has a fascinating history, closely interwoven with that of Biometry, Genetics, and, of course, the development of Mathematical Statistics. It is very much the story of the achievements of two celebrated statisticians, Karl Pearson (1857-1936) and Ronald Fisher (1890-1962), but there are interesting earlier connections with William Farr, Florence Nightingale and Sir Francis Galton.

I do not have time tonight to cover all of these or Ronald Fisher in the detail he deserves - that treatise will appear elsewhere. Accordingly, in the interests of parsimony, I begin, in 1989, with Sir Francis Galton's, Natural Inheritance[1], and end, in 1925, with Fisher's Statistical Methods for Research Workers[2].

It was this publication in which Galton first formalised his views on the inheritance of *continuous* traits in a mathematical law. His law, later dubbed the *'Law of Ancestral Heredity'*, incorporated the genetic influence of previous generations by means of a waning geometric series in which the coefficients were respectively 1/2, 1/4, 1/8.... In the same text he published his newly invented theory of *regression*, but his discovery of correlation (r=the Galton function) came too late for inclusion[3].

However, the derivation of the Law was flawed by his rudimentary command of mathematics, and it fell to Karl Pearson, Goldsmith Professor of Applied Mathematics at University College London, a colleague and admirer, to place the work on a sounder theoretical footing.

In 1898, Pearson published a revised version - a multiple regression equation which retained Galton's original geometric formulation[4] and, some two years later, in 1900, he followed-up with the *'Law of Reversion'*[5], re-casting Galton's earlier work on the inheritance of qualitative traits.

Describing how Galton's vision in Natural Inheritance had fired his own imagination, Pearson later said[6]:

> *"I interpreted Galton to mean that there was a category broader than causation, namely correlation of which causation was only the limit, which brought psychology, anthropology, medicine and sociology, in large parts, into the field of mathematical*

treatment. It was Galton who first freed me of the prejudice that sound mathematics could only be applied to natural phenomena under the category of causation".

Largely as a result of his association with Galton and Wheldon (Jodrell Professor of Zoology at UCL), Pearson, in a prolific period between 1894 and 1901, succeeded in: (a) developing a system of curves capable of describing skew variation (The Pearson System), (b) formulating and solving the basic theorems of multiple linear and curvilinear regression, and those of multiple linear correlation, and (c) inventing the chi-squared test of association[7].

Tool-maker and intellectual giant that Pearson was, he could not save Galton's Law from ultimate extinction as the Mendelian hypothesis, rediscovered in 1900, gained ground. The Ancestral Law came under incessant attack from the Cambridge biologist, William Bateson, who had been responsible for having Mendel's original work, on the physical basis of inheritance, translated from the German[8].

The controversy between the Ancestrians and the Mendelians deepened. The manner of the rejection of Pearson's paper on Homotyposis, an abstract of which was read to the Royal Society on the 15th of November 1990 (Bateson was a referee), led the Ancestrians to believe that further 'Biometric' papers were likely to suffer a similar fate[9]. Wheldon wrote to Pearson suggesting that they found a journal of some kind. Pearson agreed and within a month he had chosen a title - *Biometrika*. Wheldon and Pearson were the principal editors and the first volume of *Biometrika* was published in October 1901[10]. With succeeding volumes the power and influence of Pearson and the Biometrical School grew.

Although, by 1906, the controversy had largely been resolved in favour of the Mendel's Hypothesis, it fell to another statistician, R.A. Fisher, to demonstrate mathematically that continuous variation would result from the action of many genes independently following Mendelian segregation[11].

Ronald Fisher was born in London in 1890. He was educated at Harrow and Cambridge where he graduated in 1912, a wrangler in the Mathematical Tripos, just as Pearson had done some 33 years earlier. By all accounts

Fisher's mathematical ability was remarkable.

That year, Fisher published his first paper on mathematical statistics in which he introduced the concept of *Maximum Likelihood.* He did not come down in 1912, but remained in Cambridge for another year studying the classical theory of errors, statistical mechanics and quantum theory[12].

As an undergraduate at Cambridge, Fisher read much of the output of Pearson's Biometrical Laboratory, became interested in distribution theory, and familiar with the 1908 paper of 'Student', William Sealy Gosset, who was working on the problem of small samples.

Gosset's contribution is well known to the Society. He conjectured, but did not formally prove, that the variable s^2 was distributed as a Pearson Type III. However, by assuming this result, he was first to make allowance for the variation of s^2 in small samples by means of his z statistic[13]. The importance of Gosset's work in this area was not lost on Fisher.

In 1915 Fisher made the first of many remarkable contributions to the theory of statistical distributions, by publishing the exact distribution of correlation coefficient, r. His paper[14] appeared in the 10th volume of Biometrika and it is not difficult to imagine how the result must have been received in the Biometrical Laboratory which had laboured long to investigate the coefficient which Pearson had all but invented.

Three years later Fisher submitted his synthesis supporting Mendelian inheritance to the Royal Society[11]. By an astonishing irony Pearson and Bateson were appointed referees. Both recommended rejection! Published privately in 1918, Fisher's paper demonstrated that the findings of the Biometrical School could only rationally be explained in terms of Mendelism. Galton would have approved, but it was another blow to Pearson.

In 1919 he moved to the Rothamsted Experimental Station where he was to spend the most prolific decade of his life. By 1920 he had abstracted the principle of *Sufficiency*[15] - perhaps his greatest insight - linking it later to the *Likelihood* function, and his concept of the amount of *Information* in a sample.

Notwithstanding this effort, he found time, in 1922, to correct Pearson's formula for the number of degrees in the x^2 test of 'Goodness of Fit', and it was this paper[16] more than any other which riled Pearson and heralded the decline of the Biometrical School.

Remarkably, in the same year he also published that marvellous treatise, *'On The Mathematical Foundations Of Theoretical Statistics'*[17]. Fisher was just 32 years of age, but controlled mastery leaps from every page of this great synthesis. Here was the architect, visionary, and genius at work.

More was to follow quickly, and by 1924 he had extended 'Student's' work to the comparison of means from two independent samples and, further, into the realm of multiple regression, Pearson's own invention! In 1924, at 67, Pearson, though still active, was too old to respond to the flood of theoretical developments which flowed from Fisher's pen.

While working at Rothamsted Fisher developed, with Frank Yates, the modern analysis of variance, covariance, and the principles of experimental design which govern so much medical research today[12]. These techniques, of which the comparative trial is but one, were only introduced into Medicine after the second world war. They were however developed by Fisher before 1925 and published in the first edition of his now classic text: 'Statistical Methods for Research Workers'. Fisher wrote in the preface:

> *"It was clear that the traditional machinery inculcated by the Biometrical School was wholly unsuited to the needs of practical research. The futile enumeration of innumerable measures of correlation, and the evasion of the real difficulties of sampling problems, under cover of contempt for small samples, were obviously beginning to make its pretensions ridiculous."*

Homage indeed to Gosset. But with so much achieved, that Fisher felt it necessary, in 1925, to openly criticise the Biometrical School thus, was a mark of Pearson's great influence, undamaged by the vitriol of their exchanges, and sustained for a generation in the pages of Biometrika.

Just as Pearson had been inspired by Galton's vision of the Law, so had

Fisher been inspired by Pearson's creation of Mathematical Statistics. Detailed comparisons, though inevitable, would occupy more space than I have available tonight. That both men were geniuses is clearly true. Pearson, the first Biometrician, reaching for a concise statistical expression of the natural laws of inheritance; and Fisher, the young contemporary, an abler mathematician, revising Pearson's work and abstracting new principles with which to lay the foundations of Theoretical Statistics.

That Pearson fully acknowledged his debt to Francis Galton is historical fact[9], but Fisher never recognised his debt to Pearson. Somewhere, perhaps in the cauldron of competition from which he was to emerge the victor, his sensitivity had been blunted.

Personalities, apart, it is appropriate to end this brief note, in the spirit of Gosset (courtesy of Professor Boland) by quoting Fisher:

> *"The statistician is no longer an alchemist expected to produce gold from any worthless material offered him. He is more like a chemist capable of assaying exactly how much of value it contains, and capable also of extracting this amount, and no more. In these circumstances, it would be foolish to commend a statistician because his results are precise or to reprove because they are not. If he is competent in his craft, the result follows solely from the value of the information given him. It contains so much information and no more. His job is only to produce what it contains".*

It is Fisher's lexicon we use today.

References

1. **Galton, F., 1889.** *Natural Inheritance,* London, Macmillan .

2. **Fisher, R.A., 1925.** *Statistical Methods for Research Workers,* London, Oliver & Boyd .

3. **Galton, F., 1888.** "Co-relations and their measurement, chiefly, from anthropometric data", *Proceedings of the Royal Society,* Vol. 45, pp. 135-145.

4. **Pearson, K., 1898.** "Mathematical contributions to the theory of evolution. On the Law of Ancestral Heredity", *Proceedings of the Royal Society,* Vol. 62, pp. 386-412.

5. **Pearson, K., 1900.** "Mathematical contributions to the theory of evolution. On the Law of Reversion", *Proceedings of the Royal Society,* Vol. 66, pp. 140-164.

6. **Pearson, K., 1934.** *Speeches delivered at a Dinner held, in University College London, in honour of Professor Karl Pearson,* University Press, Cambridge.

7. **Froggatt, P., 1970.** "Modern Epidemiology: The Pearsonian Legacy", *Inaugural Lecture,* QUB.

8. **Bateson, W., 1901.** "Introductory note to the translation of Experiments in Plant Hybridisation by Gregor Mendel", *Journal of the Royal Horticultural Society,* Vol. 26, pp. 1-3 by Bateson, Translation: pp. 4-32 by CT Druery.

9. **Pearson, K., 1930.** *"The life, letters, and labours of Francis Galton",* Vol. IIIA, p. 241, Cambridge.

10. **Froggatt, P. and Nevin, N., 1971.** "Galton's 'Law of Ancestral Heredity': Its influence on the early development of human genetics", *History of Science,* Vol. 10, p. 13.

11. **Fisher, R.A., 1918.** "The correlation between relatives on the supposition of Mendelian Inheritance", *Transactions of the Royal Society of Edinburgh,* pp. 399-433.

12. **Obituary, 1963.** "Sir Ronald Fisher 1890-1962", *Journal of the Royal Statistical Society,* Part 1, pp. 159-178.

13. **'Student' 1908.** "On the probable error of a mean", *Biometrika,* Vol. 6, pp. 1-25.

14. **Fisher, R.A., 1915.** "Frequency distribution of the values of the Correlation Coefficient in samples form an indefinitely large population", *Biometrika,* Vol. 10, pp. 507-521.

15. **Stigler, M., 1973.** "Laplace, Fisher and the discovery of the concept of sufficiency", *Biometrika,* Vol. 60, No. 3, pp. 439-445.

16. **Fisher, R.A., 1922.** "On the interpretation of x2 from contingency tables and the calculation of P", *Journal of the Royal Statistical Society,* Vol. 85, pp. 87-94.

17. **Fisher, R.A., 1922.** "On the mathematical foundations of theoretical statistics", *Philosophical Transactions of the Royal Society A,* Vol. 222, pp. 309-368.

Michael Stuart: In this very interesting paper, Denis Conniffe has demonstrated the breadth and depth of his knowledge of our subject and its history. I congratulate him. He has covered a lot of ground and so, inevitably, leaves considerable scope for comment.

First, I note that the promotion and study of jurisprudence is part of the object of our Society, as enshrined in Article 1 of our Laws and Constitution. This should ease some of Dr. Conniffe's worries about the relevance of his topic to our Society. Certain forms of identification evidence owe much to both Statistics and Genetics. Dr. Conniffe himself referred to Galton and fingerprinting, Fisher and blood grouping. Fisher's enormous contribution to Genetics played a role in the development of "genetic fingerprinting", fast becoming a key form of identification.

Dr. Conniffe suggests that critics of "Taguchi" methods claim that they are just slickly packaged versions of classical experimental design and analysis, and implies that the criticism consists of unessential mathematical formulations of ideas and arguments. In fact, George Box and other critics make the point that Taguchi methods are garbled and sometimes grossly inefficient versions of classical experimental design and analysis, to be implemented uncritically as 'black box' methods. It takes very little mathematics to demonstrate that Taguchi's approach to experimental design and analysis is not only inefficient but can be misleading. On the other hand, Box praises Taguchi for getting engineers to use any sort of systematic approach to experimental design, a major advance even in this day and age. In a way, Taguchi succeeded in this respect where Box had been rebuffed. Box, who had studied with Fisher and married his daughter, was one of a group of British statisticians who took Fisher's ideas on agricultural experimentation and applied them to industrial experimentation, as well as developing and extending them for industrial application. These contributions were well received by engineers, and particularly chemical engineers, in the 1940s and 50s. However, when British industry fell into the control of accountants and related disciplines in the 1960s, the return on investment in research and development was seen as too long term; statisticians were the first to go. It is ironic that Fisher's ideas, in the garbled version popularised by the Japanese, are now being put to work again, in the effort to rescue Western industry from Japanese domination.

Dr. Conniffe refers to the apparently never ending controversy about the correct way to analyse a 2 x 2 table of frequencies, a problem that might naively be regarded as far too trivial to merit such extensive attention. He may not have seen a paper in a recent issue of Biometrika, where the arguments for and against five different approaches are teased out in the context of data arising from a clinical trial which used a modified "play-the winner" strategy in allocating subjects to treatments. Unfortunately, the resulting data were both scarce and extremely unbalanced, with 11 subjects being allocated to the experimental treatment, all successes, and one control, a failure. Depending on the approach taken, the p-value ranged from .001 to .62. It strikes me that the reason that the approaches vary so much is that the data is grossly inadequate for deciding the substantive issue; a larger sample with a more statistically appropriate design would have shown much more unanimity among the different approaches. I suspect that Fisher would have been extremely annoyed to witness such a debate on theoretical issues relating to methods of data analysis, when the real issue in this case was the practical one (and Fisher was, above all, a practical statistician) of the adequacy of the data for the problem to hand.

I was not convinced by the intuition advanced to explain the result that the covariance between an efficient estimator and another estimator must be equal to the variance of the efficient estimator. In fact, I would venture to say that Fisher would not have been impressed either, but would have used a different intuition based on a geometric approach. If an estimator is represented as a vector, then the covariance of two estimators is the inner product of the corresponding vectors. The efficient estimator is got from another estimator by projecting it onto an appropriate vector subspace; in a regression problem, this is the subspace spanned by the columns of the X matrix. Representing the other estimator as the sum of the projection and its complement, the covariance of it and the efficient vector is the corresponding sum of covariances. The covariance of the efficient estimator with the projection, itself, is just the variance. The covariance of the efficient estimator with the orthogonal complement of the projection is O. Hence the result. Anyone who doesn't have a geometric intuition wouldn't understand what I have just been saying. Fisher did, and would. Fisher's keen geometric intuition led him to state many mathematical results without apparent proof; when viewed geometrically, the proofs were trivial. However, this frustrated many students of his work, for whom the geometric view itself was far from trivial, leading to either scepticism or exceedingly tortuous analytical proofs, which made the study of mathematical statistics appear to be very forbidding indeed.

I conclude by congratulating Dr. Conniffe on his excellent paper and his impressive delivery of it.

Reply by D. Conniffe: I am grateful to Professor Spencer for proposing the vote of thanks and to Professor Boland for seconding it. I also thank Dr. MacKenzie and Dr. Stuart for their remarks. I think most of the content of the contributions has taken the form of extra information about Fisher, his work and influence, rather than being criticisms, or comments, on points I made myself. I found them most interesting and I think they add valuable extra perspective to a picture of Fisher.

Both Professor Boland and Dr. Stuart mention Fisher's powers of geometrical insight. Boland refers to the $n - 1$ denominator in the estimate of the variance as one example of this. But I suspect that at the time of the initial correspondence between Fisher and Student, it might have been Student, rather than Fisher, who was arguing for the 'right' divisor. This is because Fisher had only just published his paper advocating maximum likelihood and, of course, the MLE of the variance gives n as denominator, so he might have been reluctant to abandon it. However, Professor Boland's view is the more usual interpretation and since the surviving correspondence is only partial, I doubt if it can ever be disproved.

Dr. Stuart feels Fisher would have given a geometrical intuitive justification for the key idea in 'Hausman' specification tests and outlines how it would have gone.

He's probably right and my own approach is not all that bright, but not everyone finds geometric interpretation specially revealing. Dr. Stuart also suggests, with the 2 x 2 conditional v unconditional test case in mind, that Fisher would react angrily to hearing continued debate about inferential issues that only matter in small samples, rather than witnessing applications of statistical methods to practical problems. I think Fisher would be angry, not because of the small sample emphasis, but because he would claim to have conclusively demonstrated the correct procedure. Fisher sometimes argued in a very small sample context and he was, perhaps with justification, an arrogant man.

May I thank all the contributors again.