RONALD AYLMER FISHER

1890-1962

RONALD AYLMER FISHER was born on 17 February 1890, in East Finchley. He and his twin brother, who died in infancy, were the youngest of eight children. His father, George Fisher, was a member of the well-known firm of auctioneers, Robinson and Fisher, of King Street, St James's, London. His father's family were mostly business men, but an uncle, a younger brother of his father, was placed high as a Cambridge Wrangler and went into the church. His mother's father was a successful London solicitor noted for his social qualities. There was, however, an adventurous streak in the family, as his mother's only brother threw up excellent prospects in London to collect wild animals in Africa, and one of his own brothers returned from the Argentine to serve in the first world war, and was killed in 1915.

As with many mathematicians, Fisher's special ability showed at an early age. Before he was six, his mother read to him a popular book on astronomy, an interest which he followed eagerly through his boyhood, attending lectures by Sir Robert Ball at the age of seven or eight. Love of mathematics dominated his educational career. He was fortunate at Stanmore Park School in being taught by W. N. Roe, a brilliant mathematical teacher and a well-known cricketer, and at Harrow School by C. H. P. Mayo and W. N. Roseveare.

Even in his school days his eyesight was very poor—-he suffered from extreme myopia—and he was forbidden to work by electric light. In the evenings Roseveare would instruct him without pencil or paper or any visual aid. This gave him exceptional ability to solve mathematical problems entirely in his head; and also a strong geometrical sense, which stood him in good stead later in the derivation of exact distributions of many well-known statistics derived from small samples. Other mathematical statisticians, most of whom were not very skilled algebraists, consequently found his work difficult to follow, and often criticized him for inadequate proofs and the use of 'intuition'.

From Harrow he obtained a scholarship to Gonville and Caius College, Cambridge, which he entered in 1909. He became a Wrangler in 1912, with distinction in the optical papers in Schedule B. After graduating he spent a further year at Cambridge with a studentship in physics, studying statistical mechanics and quantum theory under James Jeans and the theory of errors under F. J. M. Stratton.

Although during his education Fisher specialized in mathematics, he had also from an early age strong biological interests. As a boy he was undecided

Reprinted from

Biographical Memoirs of Fellows of the Royal Society of London, 9*:* 91-120, (1963).

whether to concentrate on mathematics or biology. But on a chance visit to a museum he happened on a cod's skull with all its bones separated and labelled with their names; he decided on mathematics.

While at Cambridge, he came across Karl Pearson's *Mathematical contributions to the theory of evolution.* He became keenly interested in evolutionary and genetical problems, an interest which was to remain with him throughout his life.

On leaving Cambridge, he took up statistical work in the office of the Mercantile and General Investment Company. He was, however, clearly restless and for a time worked on a farm in Canada. In 1914 he had qualified himself for a commission, and did his best to join the army, but fortunately his defective eyesight made him ineligible for military service. Instead he spent the years 1915 to 1919 teaching mathematics and physics at various public schools, including Haileybury. In 1917 he married Ruth Eileen, daughter of H. Grattan Guinness, M.D. There were eight children of the marriage, two sons and six daughters. The elder son, George, broke off his medical education to join the Royal Air Force as a fighter pilot, and was killed in operations over Sicily in 1943. The younger, Harry, also joined the Royal Air Force, but was precluded from flying operations because of motion sickness.

During this period Fisher, although not engaged in work that itself led to research, was actively developing his ideas on both statistics and genetics. He had already published a paper on the fitting of frequency curves in 1912, and shortly before the war came into contact with Karl Pearson, and obtained 'in a week' the exact distribution of the correlation coefficient, a problem which had defeated Pearson and his colleagues. In 1918 he published a monumental study on correlation between relatives on the supposition of Mendelian inheritance. This was first submitted to the Royal Society but on the recommendation of the referees was withdrawn and was subsequently published by the Royal Society of Edinburgh, partly at the author's expense.

This early work led to simultaneous offers in 1919 of the post of chief statistician under Karl Pearson at the Galton Laboratory and of a newly-created post of statistician at Rothamsted Experimental Station under Sir John Russell. Neither post had much financial attraction, but in spite of the reputation of the Galton Laboratory at that time, Fisher unhesitatingly accepted the Rothamsted offer, which he thought would give considerably greater opportunities for independent research. Doubtless also the prospect of being able to pursue his genetical studies more actively at Rothamsted weighed with him.

Rothamsted proved a fortunate choice. It provided an exceptionally free atmosphere for the pursuit of research, and brought him into close contact with biological research workers of very varied disciplines and attainments. Although Russell had appointed him to re-examine by 'modern statistical methods' the mass of data that had accumulated from the long-term Rothamsted field trials, one dating from 1843—'raking over the muck heap', as Fisher later unkindly described it—Russell was too good a director to dictate to those of his staff who showed ability for original research and the will to pursue it. Fisher's appreciation of, and readiness to discuss, the practical needs and difficulties of other workers in the laboratory, and his own liking for numerical work—he was a fast and accurate computer—coupled with his great mathematical and logical ability, soon began to bear fruit. While at Rothamsted not only did he recast the whole theoretical basis of mathematical statistics, he also developed the modern techniques of the design and analysis of experiments, and was prolific in devising methods to deal with the many and varied problems with which he was confronted by research workers at Rothamsted and elsewhere. *Statistical methods for research workers*, which appeared in 1925, was essentially a practical handbook on these new methods. It made them generally available to biologists, who were not slow to take advantage of them.

By this time his work was becoming known to a wider circle of research workers. In addition to his own small department, an increasing number of visitors from other institutes came to work under him. He was elected a Fellow of the Royal Society in 1929, a fitting recognition of the great achievements of his first ten years' work at Rothamsted.

While at Rothamsted, he pursued his studies on genetics and evolution, and undertook a series of breeding experiments on mice, snails and poultry; from the last he confirmed his theory of the evolution of dominance. Rothamsted was never officially concerned with this work, as neither genetics nor plant breeding was done there, but the Station provided land for his poultry and accommodation for his snails. The mice he kept at his own home, helped by his wife and children.

The genetical theory of natural selection, which was published in 1930, completed the reconciliation of Darwinian ideas on natural selection with Mendelian theory. In this book he also developed his theories on the dysgenic effects on human ability of selection in civilized communities; the major factor responsible he believed to be the parallel advancement in the social scale of able and of relatively infertile individuals, the latter because of their advantages as members of small families, with consequent mating of ability with infertility.

He regarded the reversal of these trends as a matter of paramount long-term importance, and for a time played an active part in the affairs of the Eugenics Society. His immediate practical proposal of family allowances proportional to income was, however, so much at variance with the current thought of the time that he found few followers. His conviction of the power of genetic selection in moulding organisms to fit their environment also led him to belittle the importance of reforms, such as improvements in nutrition, designed to better the environment, and this enabled progressive thinkers to dismiss his theories on human evolution, which in any event they found distasteful, without serious thought.

In 1933 his genetical work led to his appointment at University College, London, as Galton Professor in succession to Karl Pearson. He did not, however, take over the whole of Karl Pearson's Department; Karl Pearson's son, E. S. Pearson, was put in charge of the statistical part, which was renamed the Department of Applied Statistics. As the accommodation was also split, the two departments lived cheek by jowl, sharing the same common room, which produced at times Gilbertian situations, and called for considerable tact by visitors with contacts with both camps. For by then Fisher's long-standing conflict with Karl Pearson and his followers (referred to below) had become completely irreconcilable.

The Galton Laboratory offered opportunities for the experimental breeding of animals which had not been available at Rothamsted. He brought to the laboratory the mouse colony which he had started at his home and expanded it, first to search for linkages and later for the study of modifiers in the expression of mutant genes. He continued with his snails, introduced grouse locusts and began a co-operative experiment in dog genetics. He even brought marsupials into the department, though they were soon abandoned as impracticable for breeding experiments.

Nor was the study of human genetics neglected. In association with G. L. Taylor and R. R. Race, he developed the study of the genetical aspects of blood groups, and in particular unravelled the complexities of the Rhesus system, which is responsible for erythro-blastosis foetalis. In addition to the investigation of statistical problems in genetics, he was also exerting his influence on the improvement of statistical methods in many other subjects. He took over *The annals of eugenics* from Karl Pearson. This journal had originally been founded for the publication of papers on eugenics and human genetics, so as to leave *Biometrika* (which went to the Department of Applied Statistics) free for papers on statistical methodology, but under Fisher's guidance it rapidly became a journal of importance in statistics.

In 1935 *The design of experiments* was published; this was the first book explicitly to be devoted to this subject, and amplified and extended the somewhat cursory and elementary exposition in *Statistical methods.* In 1938 he and Yates produced *Statistical tables for biological, agricultural and medical research.*

During these years he and his family continued to live in Harpenden. He kept close contact with his old friends at Rothamsted, and was a familiar figure in the laboratories on Saturday mornings. At the outbreak of war University College was evacuated. Fisher resisted this for as long as he could, but eventually he had to bow to the inevitable, and returned to Rothamsted, where Sir John Russell found him and his Department accommodation. He had hoped that his Department would have been useful, as a computing team, in the war effort. But nothing came of this, and many of his staff dispersed to other jobs.

In 1943 he accepted the Arthur Balfour Chair of Genetics at Cambridge in succession to R. C. Punnett. He held this chair until his retirement in

1957, and remained in Cambridge until his successor was appointed in 1959. While in Cambridge he continued and extended his work with mice, particularly in the study of linkage and inbreeding. Also, for the first time, he had facilities for plant genetics in the gardens of Whittingehame Lodge where the department was housed. These he used for his analysis of tristyly in *Lythrum* and *Oxalis*. He initiated a small and short-lived, but successful, programme of work with bacteria and built up a flourishing school of mathematical genetics.

When he returned to Cambridge he was re-elected a Fellow of his old college, Gonville and Caius. For the most part he resided in College, though for a time he slept at Whittingehame Lodge, at some personal inconvenience, to register his disapproval of the University regulation which required a reduction in stipend of University staff residing in College! In his later years he became almost a legend to Caius undergraduates. His venerable figure was well known to them, for every morning he took breakfast in solitary state on High Table. He attended Chapel regularly (reputedly wearing the hood of a different honorary degree each Sunday) and even on occasion preached in his own characteristic style. His scriptural knowledge was as extensive and accurate as his erudition in many fields. At first sight he appeared reserved and aloof to his colleagues in College, but closer acquaintance revealed affections which, although closely guarded, could bestow warmth of heart and generosity of spirit to his friendships. In conversation he brought not only a vast store of knowledge but also an independent mind of great rigour and penetration to bear on almost any subject. He constantly questioned conventional assumptions. He added distinction to the society of the Senior Combination Room, in which he particularly enjoyed the company of the younger Research Fellows. He was addicted to the *Times* crossword, which he usually did alone, since he was accustomed only to fill in those letters where words crossed each other; indeed, he was engaged on such a crossword almost on his deathbed. In 1956 the Fellows did him the honour of electing him as their President; and after he had retired from his University Chair they continued his Fellowship for life.

On leaving Cambridge he paid a visit to E. A. Cornish, Head of the Division of Mathematical Statistics of the Commonwealth Scientific and Industrial Research Organization in Adelaide. He found the climate and intellectual atmosphere there to his liking, and was persuaded by Cornish to accept a Research Fellowship. Apart from short visits to other parts of the world he remained in Adelaide till his death, following an operation, on 28 July 1962. He was still, up to a week before his death, full of intellectual vigour, and actively engaged in statistical research.

During his later years Fisher was the recipient of many honours. He was awarded the Royal Medal of the Royal Society in 1938, the Darwin Medal in 1948, and the Copley Medal, the highest award of the Society, in 1955. He was knighted in 1952. He was an honorary member of the American Academy of Sciences, a foreign associate of the National Academy of Sciences of the

United States of America, a foreign member of the Royal Swedish Academy of Sciences and of the Royal Danish Academy of Sciences and Letters and a member of the Pontifical Academy of Sciences. He also received recognition from several foreign universities.

He was never much of an organizer; he believed in suggestion rather than direction, and was apt to wash his hands of assistants whom he found muddle-headed or lacking in initiative. Moreover he was impatient of administrators, and did not attempt to preserve good relations with them. As he said sadly after his return to Cambridge: 'I used to think only University College was difficult, but now I know all officials are obstructive.'

Fisher had a likeable but difficult character. He had many friends, and was a.charming and stimulating man to work with, and excellent company. He liked good food and wine, which he found gave an agreeable background to intellectual discussion. To human affairs in general he had a benign and tolerant attitude, and did not presume to sit in judgement on the personal conduct of his fellow men. His own values were never those of the middle-class society of his time. His large family, in particular, reared in conditions of great financial stringency, was a personal expression of his genetic and evolutionary convictions. He was fond of children, and enjoyed having them around, being completely indifferent to the disturbance they created.* Indeed he could work or discuss in surroundings that many found hopelessly distracting.

His respect for tradition, and his conviction that all men are not equal, inclined him politically towards conservatism, and made him an outspoken and lasting opponent of Marxism. Although he did not subscribe to the dogmas of religion, he saw no reason to abandon the faith in which he had been brought up, and believed that the practice of religion was a salutary and humbling human activity. As he said in a broadcast on Science and Christianity (1955):

'The custom of making abstract dogmatic assertions is not, certainly, derived from the teaching of Jesus, but has been a widespread weakness among religious teachers in subsequent centuries. I do not think that the word for the Christian virtue of faith should be prostituted to mean the credulous acceptance of all such piously intended assertions. Much self-deception in the young believer is needed to convince himself that he knows that of which in reality he knows himself to be ignorant. That surely is hypocrisy, against which we have been most conspicuously warned.'

His eccentricities, though sometimes embarrassing, were for the most part a source of entertainment to his friends, and provided an inexhaustible fund of 'Fisheriana'. In spite of his frail frame he was physically tough—at school he was keen on running—and enjoyed excellent health. His eyesight, though a

* Lady Fisher writes: 'He wasn't really. He liked two doors closed between him and the children when he was trying to concentrate. But he had the ability to jot down a mathematical train of thought in the intervals of doing something else. Possibly the family activities took the place of a wireless to him—he could "have it on" and work at the same time.'

constant source of anxiety, never seriously hindered his work or his enjoyment of life.

He liked the company of other scientists and was a familiar figure at scientific meetings and international gatherings; the latter he attended more for the opportunity of meeting his friends than to listen to scientific communications. He was largely instrumental in setting up the Biometric Society, and played a leading part in the affairs of many other societies.

He was always ready to discuss the statistical problems of others, and often came up with a solution at surprising speed. He was extremely generous with his scientific ideas; many novel methods evolved by him appeared under the names of other authors. C. I. Bliss tells a characteristic story of the origin of the maximum-likelihood method of fitting probit lines to dosage-mortality data. This occurred one Saturday in Harpenden, in 1934, when after some discussion on how to treat groups with zero or 100 per cent kill, which could not be dealt with by the approximate methods then current, Fisher remarked 'When a biologist believes there is information in an observation, it is up to the statistician to get it out'. After lunch he obtained the maximum likelihood solution, and evolved a practical arithmetical method for calculating it which rapidly became standard for all exact work.

These were the positive and very great virtues of the man. On the negative side must be mentioned his notoriously contentious spirit, his quick temper, which was sometimes provoked by trivialities, and his tendency, on occasion, to be coldly rude to those whom he regarded as misguided. Nor, when his temper seized him, was he discriminating on whom his wrath should fall, or a respecter of persons; many innocent people in authority, as well as minor officials and servants, must have been amazed to find themselves blamed for faults of others, or of Fisher himself. It was indeed often guilt by association, particularly in scientific matters.

The originality of his work inevitably resulted in conflicts with accepted authority, and this led to many controversies, which he entered into with vigour, but often in that indignant frame of mind that leads to a partial view of the problem and leaves unanswered objections that are obvious to the impartial observer. On scientific matters he was uncompromising, and was intolerant of scientific pretentiousness in all its forms, especially the pretentiousness of mathematicians. He could be unforgivingly hostile to those who in his opinion criticized his work unjustly, particularly if he suspected they were attempting to gain credit thereby. Nevertheless he undoubtedly enjoyed the cut and thrust of scientific controversy. His pungent verbal comments were well known; though frequently made without malice, they were nevertheless disconcerting to those of less robust temperament.

CONTRIBUTIONS TO MATHEMATICAL STATISTICS

To appreciate fully the revolutionary nature of Fisher's early contributions to mathematical statistics we must recall the state of the subject at the time. Karl Pearson was then at the height of his fame, and *Biometrika* the leading journal of mathematical statistics. It was the age of correlation and curve fitting; in *Tables for statisticians and biometricians* 37 per cent of the tabular matter was devoted to curve fitting, and a further 18 per cent to various forms of correlation. It was also the age of coefficients of all kinds. In attempts to assess the degree of association in 2 x 2 contingency tables, for example, such measures as the coefficient of association, the coefficient of mean square contingency, the coefficient of tetrachoric correlation, equiprobable tetrachoric correlation, and the coefficient of colligation, were proposed. The way in which these coefficients were used revealed considerable confusion between the problem of estimating the degree of association, and that of testing the significance of the existence of an association.

This confusion permeated the whole of the statistical writing and thinking of the Pearsonian school. It is perhaps not surprising, therefore, that the paramount need in statistics for methods which would provide the most accurate possible estimates from the available data had been largely lost sight of. This need had been very apparent to earlier thinkers, in particular Gauss, who had developed the method of least squares—accepted by astronomers and geodetic surveyors as the standard method for 'the adjustment of observations'—with just this end in view.

A further weakness of the Pearsonian school was their failure to consider the need of experimenters for methods appropriate to small samples involving quantitative observations. The exact solution for the test of the significance of the mean of a small sample of normally distributed material had been provided by Gossett ('Student') in 1908. It was published in *Biometrika*, and the relevant table was reproduced in *Tables for statisticians and biometricians.* (The rigorous proof of the solution was later provided by Fisher.) But as Fisher said in his obituary of 'Student' (1939), this new contribution to the theory of errors was received with 'weighty apathy'. It was left to Fisher himself to draw attention to its importance, and to extend the solution to more complex problems.

Much of Fisher's early work on statistical theory was concerned with the determination of the exact sampling distributions of statistics derived from small samples, and with the practical implications of these results. In 1915 he published in *Biometrika* the exact distribution of the value of the correlation coefficient r derived from a sample from a bivariate normal population with correlation p. He did this by defining the sample by the co-ordinates of a point in Euclidean hyperspace, a method which later provided simple solutions to other distribution problems which had proved intractable by purely algebraic methods. That such representation had considerable aesthetic appeal to him is indicated by the remark:

'The five quantities [the first and second moments of the bivariate sample] defined above have, in fact, an exceedingly beautiful interpretation in generalized space, which we may now examine.'

Pearson reacted curiously to this discovery. Without consulting Fisher he threw the whole resources of his laboratory into a 'Co-operative Study'

(Soper *et al.* 1917) designed to determine, from Fisher's exact distribution, the accuracy of Soper's previously published approximations (1913) and to give tables of the ordinates of Fisher's distribution for regions for which Soper's approximations were not adequate.

This was doubly unfortunate. Fisher, not unreasonably, was hurt by the lack of consultation; and had he been consulted he would doubtless have drawn Pearson's attention to the suggestion made at the end of his 1915 paper 'aimed at reducing the asymmetry of the curves, or at approximate constancy of the standard deviation' by the transformation

$$r = \tanh z.$$

By using this transformation Pearson and his colleagues could have avoided much of their labours; labours which were in any case of little value, since integrals rather than ordinates of the distribution were required.

In the event Fisher himself developed the suggestion in his 1921 paper on the intraclass correlation coefficient, showing that the sampling distribution of z approximates closely to normality. In 1924 he established the distribution of the partial correlation coefficient, and showed that the effect of the elimination of variates is simply to reduce the effective size of the sample by unity for each variate eliminated. An account of the use of the (r, z) transformation was later included in *Statistical methods for research workers.* The practical task of testing the significance of differences between correlation coefficients, and of the difference of a coefficient from some theoretical value, was thereby immensely simplified, and the misleading nature of the 'probable error' of statistics such as the correlation coefficient of which the distribution is far from normal was made apparent.

This incident marked the beginning of a conflict that lasted until Pearson's death. It has been represented that both men were equally at fault, but this is unfair to Fisher, who in the early years at least, wrote in a temperate and respectful manner about Pearson's work. Had Pearson possessed a more open mind he would have perceived that in Fisher there was a man who could resolve many of the tangles which the subject had got itself into. Instead he did his best to prove Fisher wrong, and to place difficulties in his way. Even publication in *Biometrika* was refused. Small wonder that Fisher was unforgiving.

The publication of the 'Co-operative Study' had a further important consequence, in that it raised the issue of inverse probability. In 1912, while still at Cambridge, Fisher had suggested the method of maximum likelihood for estimation. At the time this attracted him because it provided an 'absolute criterion' for the choice of estimates. In his 1915 paper he followed the proposal of his 1912 paper, of finding the value of p for which the probability of obtaining the observed r was maximized, i.e. maximizing the likelihood of r. This procedure was adversely criticized in the 'Co-operative study'. These criticisms caused Fisher in his 1921 paper to point out emphatically the absurdities to which the Baysian approach could lead:

'The writers of the Co-operative Study apparently imagine that my method depends upon "Bayes' Theorem", or upon an assumption that our experience of parental correlations is equally distributed on the *r* scale (and therefore not so on the scale of any of the innumerable functions of *r*, such as *z*, which might equally be used to measure correlation), and consequently alter my method by adopting what they consider to be a better *a priori* assumption as to the distribution of *p*. This they enforce with such rigour that a sample which expresses the value 0.6000 has its message so modified in transmission that it is finally reported as 0.462, at a distance of 0.002 only above that value which is assumed *a priori* to be most probable!'

The decisive step in the development of the new theory of estimation, however, was made in 1920 with the discovery of the property of *sufficiency*. As a result of an examination of a statement by R. S. Eddington in his book, *Stellar movements,* Fisher determined the relative precision of estimates, *s* and *s'*, of the standard deviation *o* of a normal distribution based on the mean square deviation and on the mean deviation, and showed that of all estimates based on the powers of deviations the square power has maximum precision. These results, though previously unknown to Fisher, were not new; what was new and important was the discovery that for a given value of *s* the distribution of *s'* is independent of *o*, and that this is true also if any other estimate is substituted for *s'*. This, as he perceived, implies that the statistic *s* contains all the information provided by the sample on the parameter *o*. Such statistics he later termed *sufficient.*

These ideas were further and much more extensively developed in his paper in the *Philosophical Transactions* (1922). The concepts of consistency, efficiency and sufficiency were clearly defined, and it was shown that the method of maximum likelihood provides a sufficient statistic where one exists, and always provides an efficient statistic; moreover that the invariance of the maximum likelihood statistic is given by the second differential of the likelihood function. Symbolically, if the probability of an observation falling in the range d*x* is

$$f\ (\theta,\ x)\ \mathrm{d}x$$

where θ is an unknown parameter, then if

$$L = \mathrm{S}\ (\log f)$$

where S denotes summation over the observed sample, the most likely value $\hat{\theta}$ of θ is given by

$$\frac{\partial L}{\partial \theta} = 0$$

and

$$\frac{\partial^2 L}{\partial \theta^2} = -\frac{1}{\sigma_{\hat{\theta}}^2}$$

This method, and the extension given in the paper to the simultaneous estimation of several parameters, was found because of its generality to be of immediate practical utility in a very wide range of problems. It provided, for non-normal (and correlated normal) material, and quantal material, efficient methods of estimation equivalent to those which had long been available for normally and independently distributed material by the Gaussian method of least squares, which is a special case of the method of maximum likelihood. It undoubtedly ranks as one of Fisher's greatest contributions to statistical methodology.

The results obtained in the 1922 paper were further elaborated and expanded in 1925. It was then recognized that for small samples the form of the curve of error of a statistic has to be taken into account, and that while in large samples all efficient statistics are equivalent in the amount of information they supply, this is not so in small samples; and that moreover where a sufficient statistic does not exist the loss of information appertaining to the method of maximum likelihood can be made good, provided a simple condition holds, by a series of ancillary statistics.

In parallel with his work on the theory of estimation Fisher continued his investigations of the sampling distributions of statistics then commonly used in tests of significance. In 1922, also by means of geometrical reasoning, he unravelled the confusion that existed in the use of the x^2 test for testing deviations from proportionality in contingency tables. In a contingency table the observed marginal totals are used for calculating the expectations in the cells on the assumption of proportionality, and the effect of this, as Fisher showed, could be allowed for by taking account of what he later called the degrees of freedom available for variation. Thus in a $p \times q$ table with given marginal totals only $(p—1)(q—1)$ of the pq cell frequencies are independent, and the x^2 distribution for $(p—1)(q—1)$ degrees of freedom is obtained, instead of, as Pearson thought, the x^2 distribution for $pq— 1$ degrees of freedom. This simple and commonsense modification Pearson and others found hard to accept, although Yule, Greenwood, Bowley and others had previously expressed doubts of the validity of the test in its original form, and Yule at least had given the correct solution (obtained by an approximate method) for 2×2 tables in the first edition of his book, *An introduction to the theory of statistics* (1911). In attempts to clear up the confusion, two further papers were published (1924), the second of which discusses the more general use of x^2 for testing agreement between observation and hypothesis, and makes the important point that the estimates of the fitted parameters must be efficient.

In addition to its uses for contingency tables and for testing the fit of distributions of given mathematical form to data grouped into classes, Slutsky (1913) had put forward a method based on the x^2 distribution for testing the goodness of fit of regression lines. Fisher re-examined this problem in 1922. The essential difference between this and the other situations in which x^2 was used is that in the regression problem the variability of the data is estimated from the observed variability within arrays. In investigating the effect of this Fisher was led to the z distribution, which he there treated as a modified x^2 distribution.

As he stated in his foreword to this paper in *Contributions to mathematical statistics* (a collection of reproductions of his more important papers), 'Before its general applicability was recognized the z distribution kept turning up unexpectedly'. The examination of the common features of the problems which gave rise to it led to the development of the analysis of variance. The uses of this analysis were first set out systematically in a paper presented to the International Mathematical Congress at Toronto in 1924 (unfortunately not published until 1928), though a simple example is given in an earlier paper (1923) on the results of an agricultural experiment. The Toronto paper also summarizes the relation between the z distribution and the t, χ^2, and normal distributions, which are special cases of the z distribution.

The analysis of variance, although once described by Fisher as 'merely a convenient way of arranging the arithmetic', in fact serves to introduce logical clarity into the many types of statistical analysis to which it is applicable and thereby greatly facilitates their correct use in complex situations. The test of the goodness of fit of a regression line in which there are q fitted constants, a arrays, and N observations, for example, can be set out in analysis of variance form as follows:

	degrees of freedom	*sum of squares*
deviations of array means from estimated regression	$a-q$	$S_p N(\bar{y}_p - Y_p)^2$
deviations within arrays	$N-a$	$S(y-\bar{y}_p)^2$

where p represents a particular array, y an observed value, and Y the corresponding value calculated from the regression formula. This form of statement may be contrasted with that of the 1922 paper, where the expressions for the sums of squares occur on different pages, and of the two sets of degrees of freedom $a-q$ is verbally described, and $N-a$ occurs as a factor in the product $(N-a)s^2$.

In his 1922 paper Fisher also gave the exact distribution of estimates of regression coefficients, for the case in which the deviations from the regression line are used for estimating the residual variance. He showed that these coefficients, divided by their estimated standard errors, are distributed in what is now known as the t distribution, originally determined in slightly different form by Gossett in 1908 for testing the mean of a sample of a normal distribution of unknown variance, and tabulated by him in 1917. Fisher's treatment is noteworthy in that the extension to several coefficients, both of orthogonal and non-orthogonal form, is also presented.

By about 1925, therefore, the main structure of the modern theory of estimation and tests of significance had been completed. As was to be expected, however, the theory of inverse probability, in spite of Fisher's earlier criticisms, was not lightly abandoned by his contemporaries. Indeed the introduction, and immediate success, of the method of maximum likelihood, which had been accidentally associated with inverse probability, served to strengthen belief in the latter.

In 1930, in his search for better logical foundations for the process of inductive inference, Fisher introduced the concept of fiducial probability. Other mathematical statisticians have found great difficulty in accepting the later developments of this concept. In its original form, however, it was very simple. In the case of a correlation coefficient ρ estimated from four pairs of values, which Fisher took as his example, if ρ has a value of $0 \cdot 765$ then values of $r \geqslant 0 \cdot 99$ will be obtained in 5 per cent. of all samples. Therefore, Fisher argued, if a value $r = 0 \cdot 99$ is observed we may say that the fiducial 5 per cent value of ρ is $0 \cdot 765$. In other words if ρ were $0 \cdot 765$ the observed value would be just significant at the 5 per cent point. Other fiducial points can be found similarly. Generally, if the sampling distribution of a statistic T of continuous variation estimated by maximum likelihood is expressible solely in terms of a single parameter θ, then the probability that T should be less than any specified value is given by

$$P = F(T, \theta),$$

and the fiducial distribution of θ may be expressed as

$$\mathrm{d}f = -\frac{\partial}{\partial \theta} F(T,\theta)\mathrm{d}\theta,$$

while the distribution of the statistic for a given value of the parameter is

$$\mathrm{d}f = \frac{\partial}{\partial T} F(T,\theta)\mathrm{d}T.$$

So far, so good. However, in 1935 Fisher extended this argument to obtain the fiducial distribution of the mean μ of a normal distribution of which the standard deviation σ is unknown, given a sample of n with mean $\bar{x}$ and estimated standard deviation s. Observing that

$$t = (\bar{x}-\mu)\sqrt{n}/s$$

has a known distribution, independent of μ and σ, he stated that the fiducial distribution of μ is given by

$$\mu = \bar{x}-st/\sqrt{n}$$

since $\bar{x}$ and s are calculated from the sample and are therefore known.

In this fiducial distribution it is no longer true that if repeated samples are taken from a population with $\mu = \mu_{5\%}$, where $\mu_{5\%}$ is determined from the single sample actually observed, then in 5 per cent of these samples a value of $\bar{x}$ as large or larger than that observed will be obtained, whatever the value of σ. Nevertheless it is the distribution (or limits of error at any given probability level) of μ rather than that of t which concerns the practical worker. And since by fixing the limits in this manner he will be wrong with a frequency equal to the chosen probability level he is likely to rest content.

Fisher, however, went on to consider the difference of the means of two normally distributed populations, and put forward the solution, originally proposed by Behrens (1929), which depends on the simultaneous distribution of two *t*'s. The relevant fiducial distribution is in fact given by

$$\mu_1 - \mu_2 = \bar{x}_1 - \bar{x}_2 - s_1 t_1/\sqrt{n_1} + s_2 t_2/\sqrt{n_2}.$$

In this case it is *not* true, in repeated sampling from the same populations (or more specifically populations with a given σ_1/σ_2), that the limits will be exceeded with a frequency equal to the chosen probability level. If, for example, the fiducial distribution is used as a basis of a test of significance of the hypothesis $\mu_1 = \mu_2$, at say the 5 per cent level, significant results will be obtained in less than 5 per cent of the cases unless $\sigma_1/\sigma_2 = 0$ or ∞.

At that time the idea had been firmly implanted, not least by Fisher himself, that if the null hypothesis were true, significant results at the 5 per cent level would be obtained on the average with a frequency of 5 per cent, though Fisher had not, we believe, ever explicitly advanced this thesis. At any rate he wrote in 1945 *(Sankhya]:*

'In recent times one often-repeated exposition of the tests of significance, by J. Neyman, a writer not closely associated with the development of these tests, seems liable to lead mathematical readers astray, through laying down axiomatically, what is not agreed or generally true, that the level of significance must be equal to the frequency with which the hypothesis is rejected in repeated sampling of any fixed population allowed by hypothesis. This intrusive axiom, which is foreign to the reasoning on which the tests of significance were in fact based seems to be a real bar to progress. . . .'

Nevertheless, it is not surprising that when other statisticians became aware of this anomaly in the Behrens-Fisher test, they thought that this cast doubt on the fiducial argument by which it was established. This controversy, which is still not wholly resolved, caused considerable confusion in statistical circles, but the illogicalities that were revealed by the attempts to provide a test which satisfied the criterion of repeated sampling did serve to bring out the fallacy of the latter approach.

The Behrens—Fisher controversy led Fisher to consider much more explicitly the whole question of the meaning of probability statements in inductive inference. This is discussed at length in *Scientific methods and inductive inference,** where he put forward the idea of *recognizable sub-sets.* The argument runs as follows. For a correct statement of probability the following requirements must be satisfied:

(a) There is a measurable reference set (a well-defined set, perhaps of propositions, perhaps of events).

(b) The subject (that is, the subject of a statement of probability) belongs to the set.

(c) No relevant sub-set can be recognized.

* This book is full of good things. It should be read, marked, learnt and inwardly digested by all serious students of mathematical statistics.

If from the available information sub-sets can be recognized which are relevant in the sense that the probability differs in the different sub-sets, then the sub-set to which the subject belongs *must* be taken as the reference set, and no statement of probability concerning the subject based on the whole set is correct.

This clear, formal statement of what was previously intuitively and somewhat dimly perceived will we believe introduce much greater clarity of thought on the type of probability statement required in inductive inference. In the case of the Behrens-Fisher test the observed value of s_1/s_2 assigns the observed sample to a recognizable sub-set which is certainly relevant.

The clarification and improvement of the theory of statistical estimation and the logic of inductive inference are Fisher's fundamental contributions to statistical theory, but they do not by any means cover the whole of his achievements. Apart from the design of experiments, discussed separately below, he made many miscellaneous contributions; some of these are of historic interest only, others such as the application of maximum likelihood to biological assay now form the basis of standard methods in current use all over the world, and yet others such as the discriminant function have still to be exploited on any considerable scale. The introduction of the discriminant function was a major advance in the problem of dealing with populations of which the members are characterized by multiple measurements, a very common situation in biological material. Because of the heavy numerical computations required, however, the method was not of great practical utility before the introduction of electronic computers, and it is only now beginning to be seriously applied to the analysis of biological and other multivariate data.

A practising statistician, of course, is not only concerned with the development of theory. He is equally concerned with the analysis and interpretation of real data. It was part of Fisher's strength that he both liked and had an aptitude for numerical analysis; he was also keenly interested in the scientific content of the problems that came his way. Many of his contributions to statistical methodology arose out of practical problems of analysis with which he was confronted by those who sought his advice.

He took considerable trouble to expound the new methods for the benefit of practical workers who wished to use them. At the time it was written (1925) *Statistical methods for research workers* was a *tour de force.* It has been greatly expanded, but little altered, in subsequent editions. It is essentially a compilation of methods, without mathematical proofs. To the modern student it is in parts needlessly difficult; thus the analysis of variance is introduced *via* the intraclass correlation coefficient, a piece of statistical junk that could well be forgotten. Its main weakness, again attributable to historical causes and subsequent lack of major revision, is insufficient consideration of the problems and pitfalls of estimation—bias in a regression coefficient due to error in the independent variate, for example, is not discussed—with the result that practical workers using Fisherian methods have often tended to place excessive emphasis on tests of significance, without asking whether the estimates they are testing are the appropriate ones.

In his own work Fisher was at his best when confronted with small self-contained sets of data, and many of his solutions of such problems showed great elegance and originality. He was never much interested in the assembly and analysis of large amounts of data from varied sources bearing on a given issue. The analysis of a single experiment and the conclusions that could be drawn from it, for example, interested him greatly, the assembly and analysis of the results of a varied collection of experiments scarcely at all. This would not have mattered—it could well be left to others—had he not tended to brush aside these more laborious and pedestrian labours, while remembering and continuing to maintain his own first conclusions based on an examination of part of the data, conclusions which inevitably required re-examination in the light of subsequent work.

The smoking - lung-cancer controversy is a case in point. To those who mistrust the alarms and excursions of the medical world Fisher's scepticism of the evidence that cigarette smoking had been established as *the* causative agent of lung cancer was refreshing. His deduction from Doll and Hill's published results that inhalers were less subject to cancer than non-inhalers, and his subsequent confirmation of this from a more detailed breakdown of the results supplied to him by Doll and Hill, was a striking demonstration of the pitfalls attendant on the interpretation of survey data; and his demonstration, from data on monozygotic and dizygotic twins, that smoking habits are strongly genetically conditioned was delightful. Yet although for some years passionately interested in the controversy, he never attempted any review of the later evidence.*

Looking back on Fisher's statistical work, one of the most surprising things is the resistance that his ideas encountered, particularly from professional mathematical statisticians. This took two forms. The first was the plain statement: Fisher is wrong. The second, more subtle, was the apparent adoption of these ideas, with subsequent mutilation and distortion. Fisher himself made many efforts to correct this tendency, as for example in his 1955 paper on statistical methods and scientific induction, where he castigated 'repeated sampling from the same population', 'errors of the second kind', 'inductive behaviour' and 'the treatment of experimental design as part of the general decision problem'. Nevertheless the tendency persists, particularly in the United States of America; as is evinced by the concluding remarks of a lecture on the nature of probability he gave at Michigan State University in 1958:

'Of course, there is quite a lot of continental influence in favour of regarding probability theory as a self-supporting branch of mathematics, and

* It has been suggested that the fact that Fisher was employed as consultant by the tobacco firms in this controversy casts doubt on the value of his arguments. This is to misjudge the man. He was not above accepting financial reward for his labours, but the reason for his interest was undoubtedly his dislike and mistrust of puritanical tendencies of all kinds; and perhaps also the personal solace he had always found in tobacco.

treating it in the traditionally abstract and, I think, fruitless way. Perhaps that's why statistical science has been comparatively backward in many European countries. Perhaps we were lucky in England in having the whole mass of fallacious rubbish put out of sight until we had time to think about probability in concrete terms and in relation, above all, to the purposes for which we wanted the idea in the natural sciences. I am quite sure it is only personal contact with the business of the improvement of natural knowledge in the natural sciences that is capable to keep straight the thought of mathematically-minded people who have to grope their way through the complex entanglements of error, with which at present they are very much surrounded. I think it's worse in this country [the U.S.A.] than in most, though I may be wrong. Certainly there is grave confusion of thought. We are quite in danger of sending highly trained and highly intelligent young men out into the world with tables of erroneous numbers under their arms, and with a dense fog in the place where their brains ought to be. In this century, of course, they will be working on guided missiles and advising the medical profession on the control of disease, and there is no limit to the extent to which they could impede every sort of national effort.'

DESIGN AND ANALYSIS OF EXPERIMENTS

When Fisher came to Rothamsted in 1919 he was brought into direct contact with workers who were concerned with the interpretation of agricultural field experiments on crops, and with laboratory and greenhouse experiments. Ideas on the errors to which experimental results are subject were at that time confused. Knowledge of experimental error is required not only to give an idea of the general accuracy of the experimental results, but also to provide a basis for exact tests of significance. Although in the classical long-term experiments laid down by Lawes and Gilbert there was no replication other than that provided by the results of successive years, replication had for long been customary in many short-term field trials, and some agronomists had become aware that the differences between the replicates could be used to provide estimates of experimental error. The appropriate method of calculating such estimates when more than two treatments were under investigation was not known, however, and various alternative methods were in use, all of which were computationally laborious, and some grossly erroneous.

Fisher early perceived that the analysis of variance provided a powerful technique for the separation of sources of variation in agricultural field trials. The first published application (1923) well illustrates the tentative nature of these early analyses. The results discussed were from an experiment at Rothamsted on the effect of potash (sulphate, chloride and none) on 12 varieties of potatoes, with three replicates on each of two series, dunged and undunged respectively.

The experiment was essentially of the split-plot type, the varietal plots being split for fertilizers, but by modern standards the design left much to be desired: in particular, the varietal plots were not arranged in blocks and the fertilizer treatments were in the same order on each plot. Nevertheless, it would have been reasonable to take account of the split plot features of the design by partitioning the degrees of freedom for error into whole-plot and sub-plot components; it is with the latter that the interaction between varieties and potash fertilizers should be compared. This was not done, however, the partition of the degrees of freedom in the paper being:

Manuring	5
Variety	11
Deviations from summation formula	55
Variation between parallel plots ..	141 *

By 1925 sources of experimental error were better understood, and the need for the separate assessment of whole-plot and sub-plot errors had been recognized. A more elaborate analysis of data from part of this experiment is given in the first edition of *Statistical methods for research workers.* The analysis of variance is there divided into two parts, for whole-plots and sub-plots respectively, and the fertilizer effect is sub-divided into sulphate *v.* chloride and mean of sulphate and chloride *v.* none.

At about the same time, in correspondence with Gossett, Fisher proposed the use of the analysis of variance to provide a pooled estimate of error in experiments in which the replicates of treatments are arranged in blocks, thus resolving Gossett's search for a method of pooling the error derived from the differences of all possible pairs of comparisons. Two proofs provided by Fisher are reproduced in a footnote to Gossett's paper ('Student', 1923); the second is the standard derivation of the analysis of variance procedure by fitting constants for blocks and treatments by least squares.

The foundations of sound methods of analysis for replicated experiments were thus laid. It was doubtless through applying these to the results of experiments involving systematic features that Fisher perceived that the need for questionable (and often demonstrably false) assumptions regarding the independence of the errors of the separate plots could be avoided by randomization of the treatments, subject to appropriate restrictions such as arrangements in blocks or in the rows and columns of a Latin square.

The principles of randomization were first expounded in *Statistical methods for research workers* (1925). Points emphasized were (i) that only randomization can provide valid tests of significance, because then, and only then, is the expectation of the treatment mean square in the absence of treatment effect equal to the expectation of the error mean square when both are averaged over all possible random patterns; (ii) that certain restrictions may be imposed, e.g. arrangement in blocks or a Latin square, for which (i) still holds if account is taken of the restrictions in the analysis of variance; (iii) that

* One whole plot was missing—it is noteworthy that the orthogonality of the partition of the first three components is ensured by giving equal weight to all variety-manuring means.

not all types of restriction that might commend themselves to the experimenter, e.g. balancing of treatment order within blocks, are permissible,

The first edition of *Statistical methods for research workers,* by providing a table of the z distribution for $P = 0.05$ also made available exact tests of significance in the analysis of variance when the degrees of freedom for one or both the components are small.

At about this time also, Fisher clarified his ideas on factorial design. Factorial design, i.e. the inclusion in the same experiment of all combinations of the levels of a group of treatments or 'factors', was not by any means new, indeed some of the early fertilizer trials laid down by Lawes and Gilbert at Rothamsted were factorial, but its advantages had never been clearly recognized, and many agricultural research workers believed that the best course was the conceptually simple one of investigating one question at a time. In 1926 Sir John Russell published a paper on experimental design in the *Journal of the Ministry of Agriculture* which put forward this thesis, and showed how little the subject was then generally understood. It was of value, however, in that it stimulated Fisher to set out his own ideas, which characteristically he did not hesitate to publish in the same journal at the earliest possible moment. Equally characteristically, Sir John bore him no ill will for this.

In this paper he made a very strong recommendation in favour of factorial design:

'In most experiments involving manuring or cultural treatment, the comparisons involving single factors, e.g. with or without phosphate, are of far higher interest and practical importance than the much more numerous possible comparisons involving several factors. This circumstance, through a process of reasoning which can best be illustrated by a practical example, leads to the remarkable consequence that large and complex experiments have a much higher efficiency than simple ones. No aphorism is more frequently repeated in connection with field trials, than that we must ask Nature few questions, or, ideally, one question, at a time. The writer is convinced that this view is wholly mistaken. Nature, he suggests, will best 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 topic has been discussed.'

The main practical difficulty in factorial design is that the number of treatment combinations increases rapidly as additional factors are added. Even with few replicates, therefore, many plots are required, and if each replicate forms a single block the blocks will also be large, and therefore relatively ineffective in removing the variability in the experimental material. Fisher perceived that this difficulty could be overcome by including a selection only of all possible combinations in a block, so that each block no longer contains a complete replicate. This device he there termed *confounding* in that certain treatment contrasts, usually high order interactions between the various factors, are identified with (confounded with) block differences.

The earliest published account of an experiment in which certain interactions were confounded is by Eden & Fisher (1929). This experiment was done at Rothamsted in 1927 and contained three replicates of all combinations of three levels (0, 1, 2) of nitrogen (N), three levels (0, 1, 2) of potash (K), and three forms of K (which we may denote by Q), in blocks of 9 plots. Each block contained the nine combinations of the first two factors, with the three forms of K assigned at random separately to the single and to the double levels of K. The design is not, therefore, divisible by blocks into replicates and the interaction N x Q, cannot be made orthogonal with blocks however it is defined. In consequence of this latter fact N x Q, cannot be separated from error in the analysis of variance unless special steps are taken to eliminate the block components which it contains. This is recognized in the published analysis, where N x Q, and N x K x Q, are left in error. Unfortunately no comment was made on why this was done, with the result that other workers later included components of treatment effects which were not orthogonal with blocks in their analyses of variance, an error which often has very serious consequences.

Having provided the basic ideas of confounding, Fisher left its detailed development to others. Much of Yates's early work at Rothamsted, when he came to work there in 1931, was concerned with this development. By about 1935 the main framework was complete, and the concepts of orthogonality, partial confounding, estimation of error from high-order interactions, etc., were thoroughly understood.

An important extension of the analysis of variance was the introduction of the analysis of covariance, which, as Fisher wrote, 'combines the advantages and reconciles the requirements of the two very widely applicable procedures known as regression and analysis of variance'. This extension, first propounded by H. G. Sanders (1930), was, as Sanders indicated, wholly due to Fisher. An example showing how the procedure can be used to increase the precision of an experiment by means of data from a preliminary uniformity trial on the same plots was included in the fourth edition of *Statistical methods for research workers* (1932).

In addition to his contributions to the design of one-year experiments, Fisher, while at Rothamsted, devised some very elegant designs for long-term rotation experiments. These initiated a more enterprising approach to the design of such experiments. Amongst other things, he emphasized the importance of having all phases of the rotation represented each year, a point often overlooked by practical agronomists.

We have set out the early history of experimental design and analysis at Rothamsted in some detail because it illustrates very well the tentative and experimental way in which the subject developed. It was part of Fisher's strength that he did not believe in delaying the introduction of a new method until every *i* was dotted and every *t* crossed. Nor did he attempt to monitor in detail all the actions and publications of those working under or in association with him. The early use of confounding in actual experiments at Rothamsted, in particular, acted as a powerful stimulus to further research into the subject. Occasionally errors occurred, some of them bad ones, but on balance progress was immensely stimulated,

The new ideas on experimental design gave rise to a surprising amount of controversy. The validity of the *t* and *z* tests was questioned on the ground that they were based on normal theory and that in much experimental material the errors are not normally distributed. Randomization was attacked because it was claimed that experimenters who knew their material could choose arrangements which were more accurate than some of the arrangements that would be arrived at by random choice. Factorial design was criticized because of the low accuracy (due to low basic replication) of the comparisons between individual treatment combinations.

Fisher reacted to these criticisms with his customary vigour, in discussions at scientific meetings and in private, by his own writings, and by encouraging others to make further investigations, particularly on the data of uniformity trials. He did not, however, always choose his ground, or argue his case, very carefully. There was a particularly regrettable controversy with his old friend Gossett, which if properly conducted would have demonstrated very clearly the value of randomization in providing valid and unequivocal estimates of error *(see* Yates 1939). Unfortunately, stimulated by an ill-judged and out-of-date recommendation of the half-drill strip method for comparing two varieties by Gossett ('Student' 1936*a*), Barbacki & Fisher (1936) tested the method on a uniformity trial and found the non-existent 'varietal' comparison to be significant. This significance was only attained because the strips were split into parts, a procedure which Fisher himself would have roundly condemned if it were done to increase nominal replication in an actual experiment, and which Gossett had himself condemned many years earlier. This led to a well-merited protest from Gossett in a letter to *Nature* ('Student', 1936*b*). Fisher replied with a further letter that can only be described as a smoke-screen. A lengthy rejoinder by Gossett ('Student', 1937) was published posthumously, with an appendix by Neyman & Pearson that only added further confusion. These outbursts were then preserved for posterity, like the archaeological remains of ancient fortifications, Gossett's *in extenso* by the editors of his *Collected papers,* and the Barbacki & Fisher paper by Fisher himself in *Contributions to mathematical statistics.*

Earlier Fisher had encouraged Tedin to test out the 'knight's move' or Knut Vik square, which is a systematic 5 x 5 Latin square arrangement, on 91 5 x 5 squares taken from 8 uniformity trials, and compare its accuracy with that of all possible Latin squares on the same plots. As was to be expected (Tedin 1931), the Knut-Vik square turned out to be somewhat more accurate than the average of all possible squares. Equally Tedin showed that the so-called diagonal squares were less accurate than the average of all possible squares. This was regarded by Fisher as a clear demonstration of the value of randomization, in that if the actual accuracy was increased the apparent accuracy would be decreased, and vice versa. In a sense it was, but the case required much more careful arguing than it received at the time. It is perhaps not surprising that the results were equally regarded by his opponents as proof of their contention that randomization did not pay.

In his general contention that a valid estimate of error is of great importance, and worth a small reduction in accuracy, Fisher was undoubtedly right. He never, however, discussed, and we believe never faced up to, a problem which still concerns practical experimenters, namely what to do if a design arrived at by random choice exhibits systematic features. In fact, of course, to say a design is random means merely that it was arrived at by random choice; once the choice has been made there is knowledge of the actual design, which provides supplementary information that may occasionally be useful; the design may even, in Fisher's later terminology, be a member of a relevant sub-set, e.g. a Knut-Vik 5 x 5 square.

To set at rest the misgivings of those who thought that tests of significance on experimental data would be seriously upset by lack of normality in the distribution of the observed values Fisher introduced the randomization test, comparing the value of t or z actually obtained with the distribution of the t or z values when all possible random arrangements were imposed on the experimental data. Randomization tests were, as Fisher wrote, 'in no sense put forward to supersede the common and expeditious tests based on the Gaussian theory of errors'; unfortunately tests of this nature, under the name of 'non-parametric tests' later came to have a certain vogue, which is not yet ended, to the confusion of practical workers.

Fisher found the combinatorial problems raised by experimental design of great interest, and he made notable contributions to the subject. He early enumerated the 5 x 5 Latin squares, and found that there were 56 standard squares, in contrast to the number 52 given by MacMahon in his *Combinatory analysis.** This stimulated him to enumerate the 6 x 6 squares, using an ingenious method that took account of their diagonal structure. The 9408 standard 6 x 6 squares were shown (1934) to be derivable from as few as 12 structurally different types by permutations and interchange of rows, columns and letters. At the same time Euler's tentative *(plus que probable)* conclusion that no 6 X 6 Graeco-Latin square exists was confirmed. This work stimulated a long series of researches into the combinatorial properties of larger squares, culminating in the discovery by R. G. Bose *et al.* (1960, 1961) of Graeco-Latin squares of side $4m + 2$ for all $m > 1$, thus disproving Euler's further conjecture that such squares did not exist, the truth of which more than one mathematician had previously purported to have proved.

* On this Fisher wrote (1934): 'The corrected number was communicated by one of the authors in 1924 to Professor MacMahon in time to be incorporated in the copies of *Combinatory analysis* then unsold'; commenting to his co-author that the reason for this self-evident statement was that MacMahon had insisted that the Cambridge University Press should reprint and replace the offending pages! The correct number, 56, had actually been given 'd'apres un *denombrement exact'* by Euler in 1782, but MacMahon, in spite of extensive quotations from this paper (still in 1934 marked in pencil in the Royal Society copy), had overlooked this.

Fisher also initiated the investigations into the existence of combinatorial solutions required for balanced incomplete block designs, i.e. designs in which the number of units in a block is less than the number of treatments, and each pair of treatments occurs together in the same number of blocks. A catalogue was included in the first edition *of Statistical tables for biological, agricultural and medical research* of the known designs for ten or less replications, together with a list of all arithmetically possible designs the existence of which had not then been disproved. The history of the discovery of further designs and proof of the non-existence of others, mainly by the researches of a brilliant group of Indian mathematicians, is recorded in subsequent editions.

The design of experiments was published in 1935. This book, which has not been greatly expanded in subsequent editions, is in no sense a manual of instruction, and does not include worked numerical examples of analysis. What it does do is to discuss in considerable detail the basic logical principles of experimentation, and the various complexities such as factorial design and confounding which are used in practical experimental work. Much of it is stimulating and relatively light reading. The introductory experiment has become world famous:

'A lady declares that by tasting a cup of tea made with milk she can discriminate whether the milk or the tea infusion was first added to the cup. We will consider the problem of designing an experiment by means of which this assertion can be tested.'

The main weakness of the book is that, as in *Statistical methods for research workers,* there is excessive emphasis on tests of significance:

'Every experiment may be said to exist only in order to give the facts a chance of disproving the null hyposthesis.'

Considering the many experiments which are made to estimate the magnitude of effects known to exist, e.g. varietal differences, responses to fertilizers, this is surely a remarkable statement.

In spite of the early controversies the new ideas on experimental design and analysis soon came to be accepted by practical research workers, and the methods have now been almost universally adopted, not only in agriculture but in all subjects which require investigation of highly variable material. The recent spectacular advances in agricultural production in many parts of the world owe much to their consistent use. They certainly rank as one of Fisher's greatest contributions to practical statistics, and have introduced a certainty of touch into well-designed experimental work that is the envy of statisticians faced with the interpretation of non-experimental data.

GENETICS

Fisher's interest in genetics began early, for the first—and one of the most important—of his genetical papers was published in 1918. When he graduated in 1912 one era in the development of genetical science was ending and another was beginning. In the first decade of the century Mendel's principles of inheritance had been shown to apply in all the groups of plants and animals that were investigated. Apparent anomalies had been tracked down to genie interaction and lethality of specific genotypes, and indeed Mendelian inheritance was proving not only to be ubiquitous but virtually exclusive also. Linkage had been discovered but was not understood. The new era was being ushered in by the early *Drosophila* experiments which were to lay bare the mechanism of linkage and to concentrate the major attention of geneticists on the chromosome system and the mechanics of crossing-over.

Fisher evidently found the new developments to be of but limited interest for he paid little attention to the *Drosophila* experiments on chromosome mechanics and crossing-over even when in later years he himself turned to consider the properties of recombination between linked genes. Rather, his interests traced to two features of the earlier genetical studies: the controversy that had raged (and indeed that was still unsettled) about the hereditary determination of metrical or continuous variation, and the tacit assumption of the early geneticists that their findings were in conflict with Darwin's notion of evolution by natural selection. We can hardly doubt that Fisher's early concentration on these two points sprang from his conviction that natural selection must be the agent of adaptation and evolution, and that Mendelism and Darwinism must fit together.

Darwin had seen the small differences of continuous variation as the raw material of adaptive change. Galton had shown such variation to be heritable. Yule and others had pointed out that, although in apparent contrast to the sharp differences from whose segregation Mendel's rules were inferred, the basic mechanism of this continuous variation need not be dissimilar provided it was assumed that the expression of the character depended on the simultaneous action of many genes whose effects were additive. Pearson and the biometricians disputed that the correlations observed between human relatives could be interpreted successfully on this basis. In his first genetical paper, 'The correlations to be expected between relatives on the supposition of Mendelian inheritance', published in the *Transactions of the Royal Society of Edinburgh* (1918), Fisher showed clearly not only that they could be so interpreted, but that Mendelian inheritance must in fact lead to just the kind of correlations observed. He showed how the correlations could be used to partition the variation into its heritable and non-heritable fractions, how the heritable fraction could itself be broken down into further fractions relatable to additive gene action, to dominance and to genie interaction (which he generalized in a way to make it manageable by biometrical methods), and how due allowance could be made for the correlation observed between spouses. Finally he pointed out that the excess of the sib correlation over that found between parent and offspring must follow from the Mendelian phenomena of dominance, whereas it was quite inexplicable on the opposing view. After this paper there could be little doubt that the inheritance of continuous variation provided no exception to the Mendelian rule. The first apparent conflict was resolved.

In this analysis, which laid the foundation for what has now come to be called biometrical genetics, Fisher paid little attention to the experimental work being done on continuous variation with several species, especially of plants. And only once, in a paper published jointly with Immer & Tedin in 1932, did he return to the subject. Even then he was concerned chiefly with the use of third degree statistics, which in fact have found but little place in the extensive studies of biometrical genetics during the last two decades. Nor did he ever discuss the biometrical consequences of linkage which, though of course negligible in his human correlations, could have profound effects on the relationships observable in experimental investigations of continuous variation. He nevertheless maintained an interest in experimental work in biometrical genetics and was always prepared to comment on the experiments, analyses and conclusions that appeared in the literature; but he, who above all could have furthered these studies in their statistical as well as their genetical aspects, never himself undertook experimental work in this field.

After dealing with the genetical mechanism underlying continuous variation Fisher turned his attention to the relations between Darwin's principle of natural selection and Mendel's principles of heredity. He delayed publication of his consideration until 1930, when it appeared as the first chapter of his great book *The genetical theory of natural selection.* His conclusions are clear and simple: by assuming, falsely, that inheritance was blending, Darwin placed himself in a dilemma, since he then had to find some agency for replacing the variation which the blending would eliminate in large amounts in each generation. This led him to postulate a stimulating action of the environment on the production of new variation, and so to destroy the whole simplicity of his notion of adaptive change resulting from natural selection. As Fisher pointed out, the essence of Mendelian inheritance is that it conserves variation, so that far from being incompatible with Darwinism it provides the very piece whose absence had led Darwin into his, as we could now see, unreal difficulties. Furthermore Johannsen's pure line experiments showed that new variation is *not* arising at the rate Darwin had been forced to postulate. He went on to show that the information already available about mutations showed them to be incapable, by virtue of both their rarity and their lack of direction, of themselves directing the course of evolution as early geneticists had supposed. Natural selection was thus displayed as the only directive agency of adaption and evolution. Darwinism and Mendelism were indeed complementary, each supplying what the other lacked, and the need was to direct genetics towards the study of selection.

His own first major study of the action of natural selection was in relation to dominance. He noted that the great majority of the mutant genes observed in *Drosophila* and other species were recessive and that few if any fully dominant mutations had been reported. This he accounted for by the hypothesis that mutants were in general harmful and that selection acted to modify the phenotype of the heterozygote (itself very much more common in wild populations than the mutant homozygote) towards the wild-type, and he adduced a great deal of evidence in favour of this process of evolution of dominance. Reactions to this hypothesis were interesting: some asserted that even the heterozygote would be too rare for the forces of selection to be effective and others argued that the modification would be achieved by selection among allelic wild-type genes rather than by selection of non-allelic modifiers as Fisher had supposed. But few, if any, denied that selection and its consequences had to be taken into account. Neo-Darwinism was on its way.

Fisher himself noted exceptions to his hypothesis in certain species. Most mutants in domestic poultry were not recessive but dominant and this he sought to explain by an ingenious theory about the early domestication of poultry. To test this theory he set up poultry breeding pens at Rothamsted and proceeded to backcross these mutants into jungle fowl, in which in his view their dominance should be incomplete. The results of these experiments, which he continued at Rothamsted after he had moved to the Galton Laboratory, were published in a series of papers in the mid-1930's and bore out his expectations. He also observed that polymorphism in wild populations commonly depends on dominant genes which, contrary to his basic hypothesis, were not more common than their alleles. Now he had already shown that if the heterozygote had a selective advantage over both the corresponding homozygotes, both alleles would be held in the population and polymorphism would result. He argues therefore that in cases of polymorphism depending on dominant genes, the variant homozygotes, though not phenotypically distinct from the heterozygote, should be less fit. Nabours's data on the grouse locusts *Apotettix* and *Paratettix* provided him with the evidence he needed to show that this was indeed the case and he was able to demonstrate differences in fitness of some 10 per cent. These differences were much larger than most geneticists were prepared to contemplate twenty-five or thirty years ago, but are in full keeping with other more recent findings. Fisher began breeding experiments, first at his home and later in the Galton Laboratory, with snails and native grouse locusts to pursue the matter further but nothing came of this work.

His theory of dominance was published in a series of papers between 1928 and 1932. It was also summarized in *The genetical theory of natural selection*. The book contained a great deal more besides. He formulated the Malthusian parameter, as he called it, for representing the fitness or reproductive value of populations, and he developed his fundamental theorem of natural selection, that 'The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time'—a finding that has been 'rediscovered' (using much more elaborate mathematics) at least once in the succeeding thirty years. He showed that even the smallest genetic changes can have no more than a half chance of being advantageous and that this chance falls off rapidly with the size of effect the change produces; that the environment must be constantly deteriorating from the organism's point of view and that this deterioration offsets the action of selection in raising fitness; that any net gain in fitness is expressed as increase in size of population; and that more numerous species carry relatively more genetic variation so that they have a greater prospect of adaptive change, and hence of survival, than their scarcer fellow species. He argued that since mutation maintains the variation in populations, its rate of occurrence will determine the speed of evolution just as selection determines the direction—a view which, however, he would unquestionably have modified, at least in relation to short-term change, if he had pursued further his studies of continuous variation. He discussed sexual selection, developing the view that natural selection will tend to equalize the parental expenditure devoted to the two sexes rather than to equalize the sex-ratio itself, and he considered the action of selection in Batesian and Mullerian mimicry. He also developed a comprehensive mathematical theory of gene survival and spread under selection, special aspects of which he extended in a number of papers over the next fifteen years or so.

This mathematical study led him to conclude, *inter alia,* that the initial establishment of a favourable mutation in any population, no matter how large, would depend on random survival; but that sooner or later the recurrence of mutation would ensure the establishment of a favourable change and that once established its fate would be governed by its consequences for selection. For a gene then to be selectively neutral, its effect on fitness must be inversely proportional to the size of the population so that, other than in special cases, it would behave as effectively neutral in selection only if its consequences for fitness were extremely small. No doubt for this reason, he would never countenance Wright's notion of gene fixation by random drift in sub-populations. In collaboration with E. B. Ford (1940-1947) he amply proved his point in the case of the *medionigra* gene in the moth *Panaxia dominula,* by measuring the size of a colony near Oxford (using a technique of marking, release and recapture which he elaborated for the purpose) and showing that this was too large to allow the changes in gene frequency, small as they were, to be accounted for by random processes.

The last five chapters of *The genetical theory of natural selection* were devoted to man, his societies and the decline of his civilizations. Human fertility was shown to have a heritable component and the causes of the negative correlation between ability and fertility were analyzed in relation to the structure of society. Fisher saw the reversal, or at least abolition, of this correlation as the great problem facing our society and discussed the use of family allowances as a means to this end. With his appointment to the Galton Chair in 1933 he turned more specifically to consider the promotion of human genetics, setting out his conclusions in a paper entitled 'Eugenics, academic and practical' (1935). He believed that linkage studies could be of great value in man and published a series of papers setting out a statistical methodology for the detection and estimation of human linkage in a wide variety of circumstances. These methods have in fact had little impact, though his statistical treatment of linkage in experimental genetics, which began with the consideration of linkage estimation from F_2 data (1928) and culminated in the balanced experimental designs for detecting and estimating linkage developed twenty years later in Cambridge, is the basis of all modern practice. It was characteristic of Fisher, too, that almost as an incidental interest he first devised the discriminant function for the combination of multiple measurements in the comparison of skulls to replace the so-called coefficient of racial likeness which he found in use at the Galton Laboratory and to which he objected as statistically unsound. He later developed and generalized the new technique for use in a wider range of applications.

The great contribution which emerged from his promotion of human studies at the Galton Laboratory was, however, in serology. He enlisted G. L. Taylor to begin work on human blood groups, initially with the aim of finding common marker genes for use in linkage studies. The great development came however when Taylor, now joined by R. R. Race, turned to the Rhesus blood groups. The story as set out by Fisher in 'The Rhesus factor—a study in scientific method' (1947) is wholly fascinating. Seven alleles of the Rhesus 'gene' had been recognized using four antisera, two of which gave complementary reactions with the seven genetic types. From this, he worked out the CDE structure of the gene, predicting the existence of two further antisera and one further rare allele. These quickly came to light once the search was begun under the stimulus of his analysis. Race and others have since shown the system to be even more complicated, but these later elaborations all fit into the trinitarian structure which was due initially to Fisher.

The Galton Laboratory offered facilities for experimental as well as human genetics and two lines of work were begun there, which continued to occupy his interest until his retirement from the Chair of Genetics in Cambridge. He brought with him to University College a small breeding colony of mice which he expanded to undertake a search for linkages. The first test of seven genes yielded little except an indication of recombination values exceeding 50 per cent between the genes for dilute pigmentation *(d)* and wavy hair (wv_1). To those with a background of *Drosophila* and chromosome studies this was evidence of the non-random assortment of the four strands in crossing-over at successive chiasmata, but Fisher preferred to approach the problem in his own way. He devised his own mathematical approach to the theory of crossing-over and recombination, which led him to conclude that genes near the ends of long chromosome arms would characteristically show values in excess of 50 per cent recombination. Protests that such recombination values *must* imply chromatid interference which the available evidence suggested was not a common phenomenon, were received with little sympathy for, he said, his treatment was based on single strands, not on the relations of the four strands of the bivalent. The issue was never finally settled, but there can be little doubt that the assumptions in Fisher's mathematical treatment do contain—and conceal—the postulate of chromatid interference.

The mouse experiments were later directed towards the adjustment of the expression of genes and their dominance properties by selection of modifiers, and also towards building up inbred lines in which groups of genes were maintained segregating for the investigation of their expressions and their linkage relations. Characteristically, too, he made a thorough theoretical investigation of the progress of inbreeding under these circumstances. The results of this investigation and an account of the theory of junctions which he developed for it, are set out in *The theory of inbreeding* (1949).

At this period he and his students were engaged in extensive theoretical investigations, one important group of which sprang from the second line of experiment begun at the Galton Laboratory. He had become interested in a complex hypothesis proposed by East to account for the inheritance of the mid-styled form of flower in *Lythrum salicaria* and he set up crosses to test it. The experiments proved East's theory wrong and established that inheritance of the controlling gene was in fact tetrasomic. The occurrence of a triplex plant, which had arisen by double reduction, aroused Fisher's interest and set him off on the theoretical exploration of segregation and linkage in both tetrasomic and hexasomic inheritance. Here his refusal to pay serious attention to the complexities raised by cytological considerations proved to be a powerful advantage, for he was able to establish both the minimal parameters to be estimated if the linkage was to be understood and the types of experiment necessary to provide data for their estimation, without trammel from the intricacies of chromosome behaviour that had confused earlier attempts. The need then was for experimental data to underpin the theory, but neither he nor his students took up material more tractable than *Lythrum* from which such data might have been obtained.

Like his theoretical work, Fisher's genetical experiments bore a characteristic stamp. They always started off with a precisely defined objective and were designed specifically and carefully for their purpose. Some, like the poultry experiments, ceased when the objective was gained. Others, like those with *Lythrum,* went on to explore further the consequences of the situation as it had been revealed. Even then, however, there was no departure from the basic pattern, for theoretical analysis was always sufficiently ahead of experimental investigation to ensure that new objectives would be successively defined. Thus in their own way his genetical experiments illustrated his statistical creed. Yet his animals and plants were never to him merely prospective entries in a statistical table. He had a real feeling for them, and despite his extreme myopia he had an eye for their variation and classification that was a lesson to all who worked with him.

In a sense his interests were always more in working out consequences, especially selective consequences, than in exploring the heritable materials and their mechanisms. He never showed much interest in chromosomes or in genetic systems. Yet he was far from blind to novel developments, as witness his initiation (in Cambridge) of experimental work with bacteria in the very early days of bacterial genetics—work which, in the hands of L. L. Cavalli-Sforza, was fruitful despite the physical difficulties of the accommodation available for it, but which unfortunately came to an early end. He brought wide reading and, within the limits of a single-mindedness that at times appeared to amount almost to prejudice, a roving interest to his genetics. His study of Mendel's experiments (1936) was a delightful example of statistical analysis applied to the better understanding of an important chapter in the history of science. He brought too the belief explicitly stated in the preface to *The genetical theory of natural selection,* that the mathematician's approach and imagination were complementary to the biologist's, that each stood only to gain from learning something of the other's world, and he handed on this belief to his students. His own work attested its truth: few could bring the same power to bear on their problems.

Fisher never occupied the dominant position in genetics that he did in statistics. Nevertheless his position in genetics was unique. He was never concerned that his work should accord with the trend of the time, and he pursued it in his own way. His achievement was as characteristic as it was basic.

F. YATES
K. MATHER

REFERENCES

Behrens, W. V. 1929 Ein Beitrag zur Fehlerberechnung bei wenigen Beobachtungen. *Landw. Jb.* 68, 807-837.

Bose, R. C., Chakravarti, I. M., & Knuth, D. E. 1960, 1961 On methods of constructing mutually orthogonal Latin squares using an electronic computer. *Technometrics, 2,* 507-516; 3, 111-117.

Euler, L. 1782 Recherches sur une nouvelle espece de quarres magiques. *Ver. v.h. Zeeuwsch Genootsch. der Wetensch. Vlissingen, 9,* 85-239.

Russell, E. J. 1926 Field experiments: how they are made and what they are. *J. Minist. Agric.* 32, 989-1001.

Sanders, H. G. 1930 A note on the value of uniformity trials for subsequent experiments. *J. agric. Sci.* 20, 63-73.

Slutsky, E. 1913 On the criterion of goodness of fit of the regression lines, and on the best method of fitting them to the data. *J. R. Statist. Sac.* 77, 78-84.

Soper, H. E. 1913 On the probable error of the correlation coefficient to a second approximation. *Biometrika,* 9, 91-115.

Soper, H. E., Young, A. W., Cave, B. M., Lee, A., & Pearson, K. 1917 On the distribution of the correlation coefficient in small samples. Appendix II to the papers of 'Student' and R. A. Fisher. A co-operative study. *Biometrika,* 11, 328-413.

'Student' 1908 The probable error of a mean. *Biometrika,* 6, 1-25.

'Student' 1923 On testing varieties of cereals. *Biometrika,* 15, 271-293.

'Student' 1936*a* Co-operation in large-scale experiments. *Suppl. J. R. Statist. Soc.* 3, 115-122.

'Student' 1936*b* The half drill strip system in agricultural experiments. *Nature, Lond.* 138, 971-972.

'Student' 1937 Comparison between balanced and random arrangements of field plots. *Biometrika,* 29, 363-379.

Tedin, O. 1931 The influence of systematic plot arrangement upon the estimate of error in field experiments. J. *agric. Sci.* 21, 191-208.

Yates, F. 1939 The comparative advantages of systematic and randomized arrangements in the design of agricultural and biological experiments. *Biometrika,* 30, 440-466.